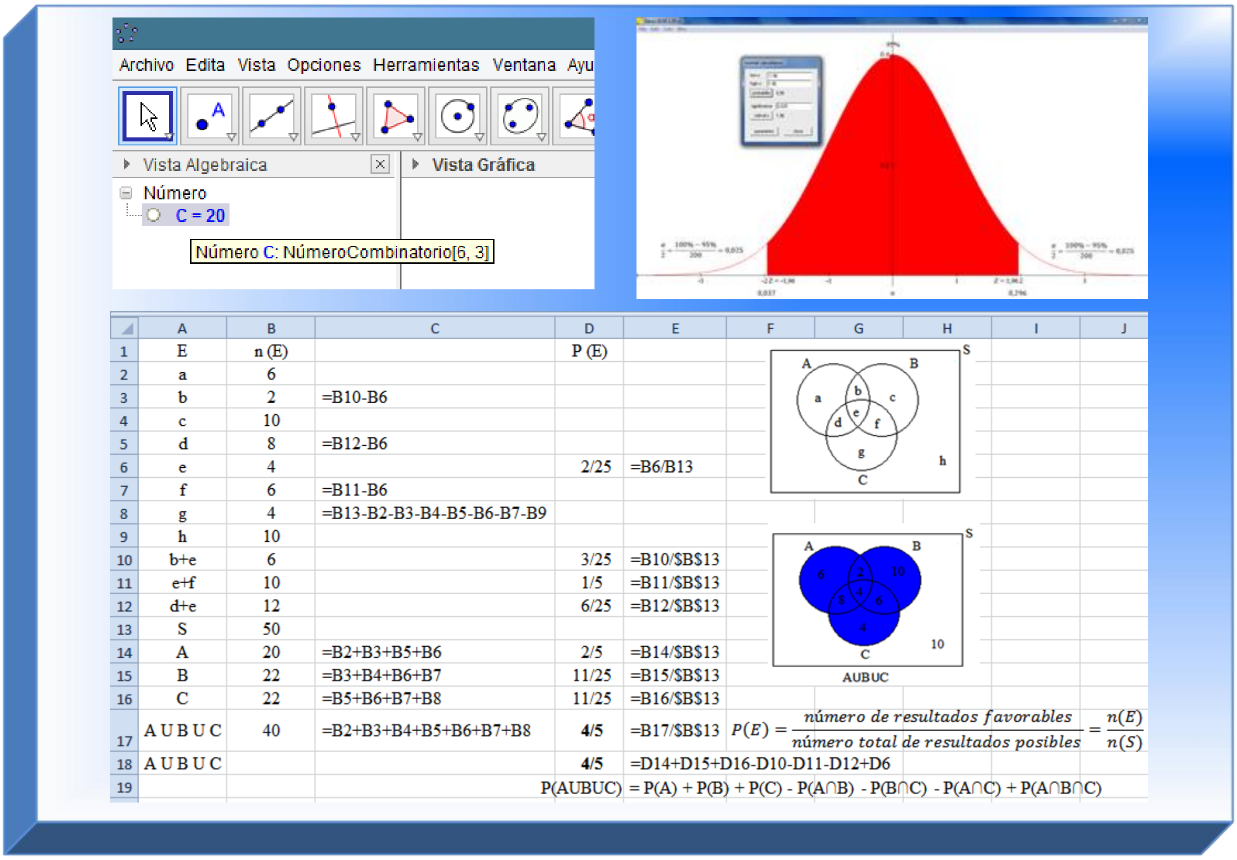
PROBABILIDADES Y ESTADÍSTICA EMPLEANDO LAS TIC



AUTOR:

Mgs. Mario Orlando Suárez Ibujés

IBARRA-ECUADOR

2014

AUTORIDADES DE LA UNIVERSIDAD TÉCNICA DEL NORTE

Dr. Miguel Naranjo
RECTOR

Dra. María de la Portilla
VICERRECTORA ACADÉMICA

Ing. Ney Mora
VICERRECTOR ADMINISTRATIVO

Dra. Soraya Rhea
DECANA FACAE

Ing. Edgar Monteros
SUBDECANO FACAE

AUTORIDADES DE LA UNIDAD EDUCATIVA IBARRA

Dra. Miriam Salgado
RECTORA

Dra. María Ruales
VICERRECTORA E.B.S.

Lic. Edwin Méndez
VICERRECTOR B.

Lic. Margarita Andrade
INSPECTORA GENERAL

Lic. José Reina
SUBINSPECTOR GENERAL

DERECHOS RESERVADOS DEL AUTOR:
Instituto Ecuatoriano de la Propiedad Intelectual
Dirección Nacional de Derecho de Autor y Derechos Conexos
Derecho de Autor Nº 044207
ISBN- 978-9942-20-184-3

Impreso de portada: Imprenta Graficolor
Impreso de páginas: Santiago Suárez y Mario Suárez

Primera Edición

Esta obra no puede ser reproducida total ni parcialmente por ningún medio sin expreso consentimiento previo y por escrito del autor.

Pedidos a los teléfonos:
Celular: 0985619601 Domicilio: 062632166
Correo:mosuarez@utn.edu.ec

DEDICATORIA

Con infinito amor en expansión
a mi esposa Dyanita Rivera (el amor de mi vida y de todas mis vidas),
a mi hija Emily y a mi hijo Mathías (la prolongación de mi existencia),
por ser mi fuente de inspiración y mi más anhelado sueño hecho realidad;
y a mis padres Bertha Ibujés y Segundo Suárez (los seres que más admiro)
por su ejemplo de sacrificio y lucha constante.

AGRADECIMIENTO

Mi gratitud y reconocimiento a las Autoridades
de la Universidad Técnica del Norte
y de la Unidad Educativa Ibarra
por el incondicional
apoyo brindado.

CONTENIDOS

PRESENTACIÓN

La Estadística en la antigüedad era empleada en asuntos del Estado tales como en los censos de población o bienes, organizados por el poder político con fines militares o fiscales. La Estadística en la actualidad es utilizada en todos los campos de saber humano, así por ejemplo, en los juegos de azar se emplea conocimientos de las probabilidades estadísticas, los investigadores utilizan conocimientos estadísticos para probar hipótesis, los gerentes de las empresas usan gráficos estadísticos para el control de la calidad de los servicios y productos que la empresa oferta, etc.

El objetivo del presente libro es incursionar a los lectores en la resolución de ejercicios y problemas de aplicación de las probabilidades y de la estadística en diversos casos de la vida cotidiana de manera manual y recurriendo al uso de Tecnologías de la Información y Comunicación (TIC) tales como Excel, Winstats, Graph y GeoGebra. Se presentan ejemplos ilustrativos prácticos que han sido cuidadosamente seleccionados y resueltos didácticamente, paso a paso, empleando un lenguaje matemático de fácil comprensión.

En cada capítulo constan los resultados de aprendizaje que se espera que el lector sea capaz de lograr al finalizar cada uno de los mismos, los contenidos a tratar y las tareas de interaprendizaje. Los contenidos y las tareas de interaprendizaje se han desarrollado de manera secuencial interrelacionadas entre sí. En el primer capítulo se desarrolla la introducción a la probabilidad (análisis combinatorio, conceptos básicos y reglas de la probabilidad), en el segundo capítulo se desarrollan las distribuciones de probabilidad discretas (binomial, Poisson e hipergeométrica) y continuas (exponencial, uniforme y normal), el tercer capítulo está dedicado a la estimación de intervalos de confianza (para la media, para la proporción y el tamaño de la muestra), en el cuarto capítulo se desarrolla la prueba de hipótesis (Z prueba, t prueba, Razón de F Fisher y Ji cuadrado), y en el quinto capítulo se desarrollan las aplicaciones de gráficas estadísticas en el control de la calidad (gráficas para variables y para atributos). Al final del libro se presentan tablas de probabilidades elaboradas en Excel y un formulario con ejemplos ilustrativos resueltos sobre conocimientos estadísticos básicos tales como el cálculo del tamaño de la muestra, regla de Esturges, medidas de tendecia central, medidas de dispersión y medidas de forma, correlación y regresión, y series crononógicas.

Los contenidos y procesos didácticos de interaprendizaje del presente libro son el fruto de la práctica en el aula durante algunos años de labor docente y que han sido publicados en http://es.scribd.com/mariosuarezibujes, https://docentesinnovadores.net/Usuarios/Ver/29591, http://repositorio.utn.edu.ec/handle/123456789/760, http://www.monografias.com/usuario/perfiles/mario_suarez_7/monografias, por lo que se infiere que el mismo tendrá la aceptación por parte de los lectores y contribuirá a mejorar significativamente el proceso de interaprendizaje de las probabilidades y de la estadística.

Convencido de que ninguna obra humana es perfecta, serán ustedes estimados lectores, los que con sus sugerencias contribuirán a mejorar la presente propuesta de interaprendizaje.

El Autor

EVALUACIÓN DIAGNÓSTICA

OBJETIVO: Verificar los resultados de aprendizaje desarrollados por l@s lectores a través del presente cuestionario para emitir suicios de valor y tomar decisiones.

INSTRUCCIONES:

Estimado lector:

- ✓ La evaluación tiene una duración de 2 horas.
- ✓ Cada pregunta tiene una valoración de dos puntos.
- ✓ No se ortorgará valoración a una respuesta correcta que no esté acompañada de un proceso de solución escrito.
- ✓ Use hojas adicionales y un esferográfico para resolver el presente cuestionario.
- ✓ Lea cuidadosamente el cuestionario y conteste según sus conocimientos previos.

¡Éxito!

CUESTIONARIO

1) Elabore un diagrama de caja y bigotes dada la siguiente distribución: 6, 9, 9, 12, 12, 12, 15 y 17

2) Calcule la moda en forma gráfica empleando un histograma con los siguientes datos:

Intervalo o Clase	f
10-19	3
20-29	7
30-39	15
40-49	12
50-59	8

3) Calcule el punto centroide con los datos de la siguiente tabla sobre la altura en centímetros (X) y los pesos en kilogramos (Y) de una muestra de estudiantes varones tomada al azar del segundo semestre de una universidad. Elabore las gráficas respectivas

X	152	157	162	167	173	178	182	188
Y	56	61	67	72	70	72	83	92

4) Calcule y grafique la ecuación de la parábola $Y = a_0 + a_1X + a_2X^2$ por el método de los mínimos cuadrados dada la siguiente tabla sobre la población de un país:

Año	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Población (millones)	4,52	5,18	6,25	7,42	8,16	9,12	10,92	11,62	12,68	13,12	13,97

5) Calcule la ecuación de la recta de tendencia por el método de los semipromedios, pronostique la tendencia de ventas para el 2011, elabore el diagrama de dispersión, y grafique la recta de tendencia con los siguientes datos sobre las ventas en millones de dólares de la Empresa D & M

Año(X)	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Ventas(Y)	1,5	1,8	2	1,5	2,2	2	3	2,8	2,4	2,9	3

Nota: *La solución de la presente prueba se encuentra al final del libro (en el formulario)*

CAPÍTULO I
INTRODUCCIÓN A LA PROBABILIDAD

RESULTADOS DE APRENDIZAJE DEL CAPÍTULO

Al finalizar el presente capítulo el lector podrá evidenciar que:

- ✓ Identifica las características, propiedades y aplicaciones del análisis combinatorio, de las probabilidades y de las posibilidades.
- ✓ Utiliza algoritmos del análisis combinatorio, de las probabilidades y de las posibilidades para resolver ejercicios y problemas de aplicación de manera manual y empleando Excel.
- ✓ Plantea y resuelve ejercicios y problemas de aplicación sobre análisis combinatorio, probabilidades y posibilidades de manera manual y utilizando Excel.

CONTENIDOS

- ✓ Análisis Combinatorio: Factorial, Permutaciones y Combinaciones
- ✓ Conceptos básicos: Experimento, Experimento Aleatorio, Espacio Muestral, Evento o Suceso, Probabilidad y Posibilidad.
- ✓ Reglas de la Probabilidad: Regla de la adición (para eventos no mutuamente excluyentes y para eventos mutuamente excluyentes) y regla de la multiplicación (para eventos dependientes y para eventos independientes)
- ✓ Probabilidad Total y Teorema de Bayes

1.1) ANÁLISIS COMBINATORIO

A) FACTORIAL

La factorial está relacionada con el cálculo del número de maneras en las que un conjunto de cosas puede arreglarse en orden.

El número de maneras en el que las n cosas pueden arreglarse en orden es:

$$n! = n(n-1)(n-2)\ldots\ldots..1$$

Donde n! se llama el factorial de n y 0! se define como 1

Ejemplos ilustrativos

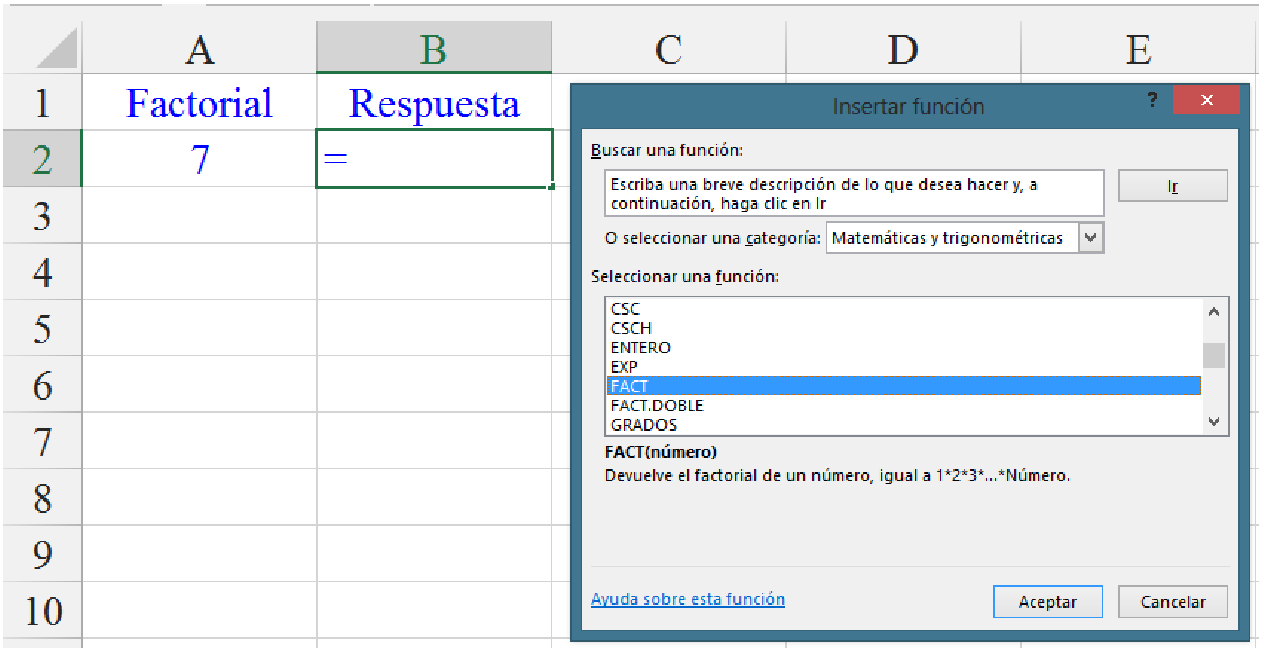
1) Calcular 7!

Solución:

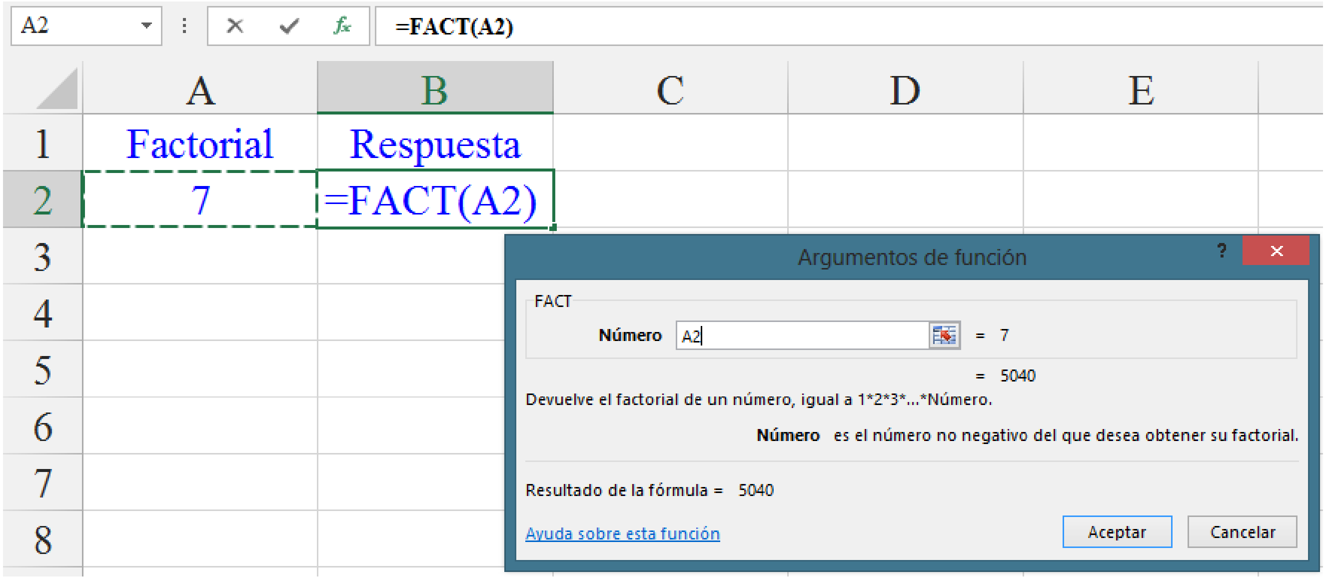
$n! = n(n-1)(n-2)\ldots\ldots..1$
$7! = 7(7-1)(7-2)(7-3)(7-4)(7-5)1$
$7! = 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1$
$7! = 5040$

En Excel se calcula de la siguiente manera:

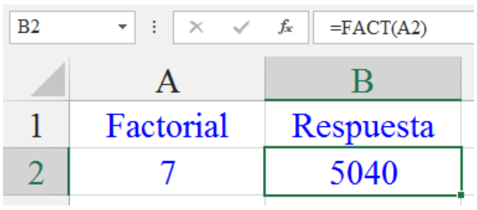
a) Insertar función. Seleccionar la categoría Matemáticas y trigonométricas. Seleccionar la función FACT



b) Clic en Aceptar. En el cuadro de diálogo de Argumentos de la función, en el recuadro correspondiente a Número seleccionar la celda correspondiente al factorial a calcular (A2).



c) Clic en Aceptar

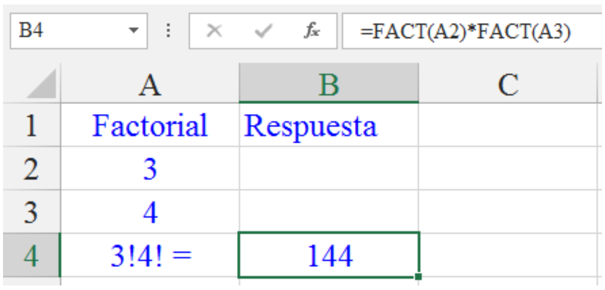


2) Calcular 3!4!

Solución:

$3!\,4! = (3 \cdot 2 \cdot 1)(4 \cdot 3 \cdot 2 \cdot 1) = 144$

En Excel se calcula como indica la siguiente figura:



3) Si un conjunto de 6 libros se colocan en un estante. ¿De cuántas formas es posible ordenar estos libros?

Solución:

$n! = n(n-1)(n-2)\,\ldots\ldots{}..1$
$6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$
$6! = 720$

B) PERMUTACIONES

En muchos casos se necesita saber el número de formas en las que un subconjunto de un grupo completo de cosas puede arreglarse en orden. Cada posible arreglo es llamado permutación. Si un orden es suficiente para construir otro subconjunto, entonces se trata de permutaciones.

El número de maneras para arreglar r objetos seleccionados a la vez de n objetos en orden, es decir, el número de permutaciones de n elementos tomados r a la vez es:

$$_nP_r = n(n-1)(n-2)\ldots(n-r+1) = \frac{n!}{(n-r)!}$$

Nota: si $n = r$, entonces la permutación se tranforma en factorial, es decir:

$$_nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

Ejemplos ilustrativos:

1) Calcular $_7P_3$

Solución:

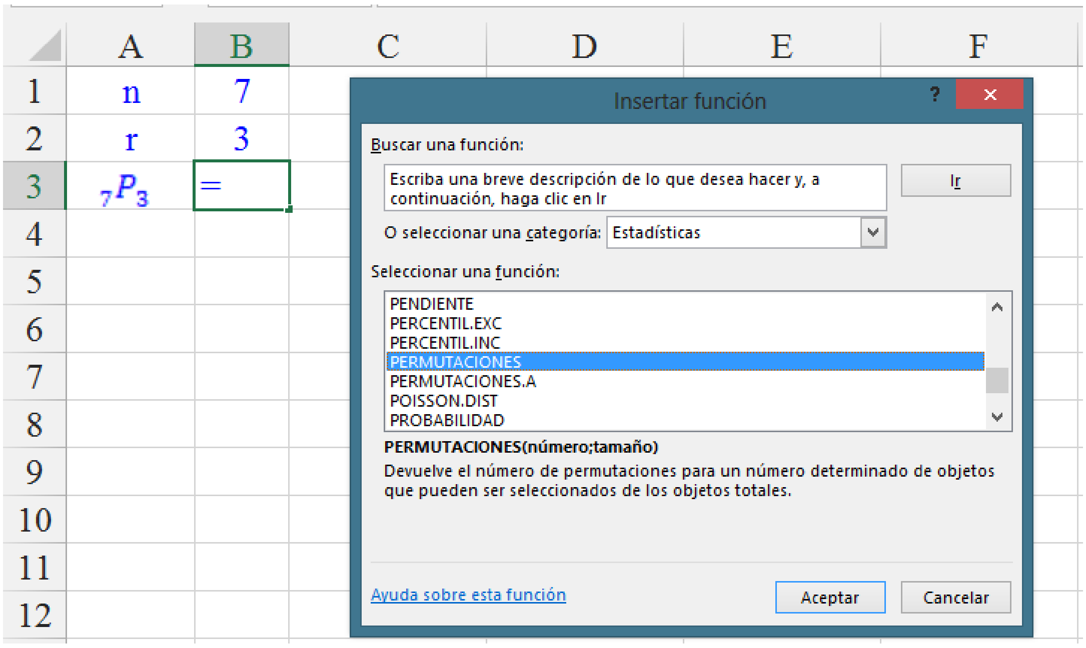
n = 7 y r = 3, entonces aplicando la fórmula se obtiene:

$$_nP_r = \frac{n!}{(n-r)!} \Rightarrow {}_7P_3 = \frac{7!}{(7-3)!} = \frac{7!}{4!} = \frac{7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{4\cdot3\cdot2\cdot1} = 7\cdot6\cdot5 = 210$$

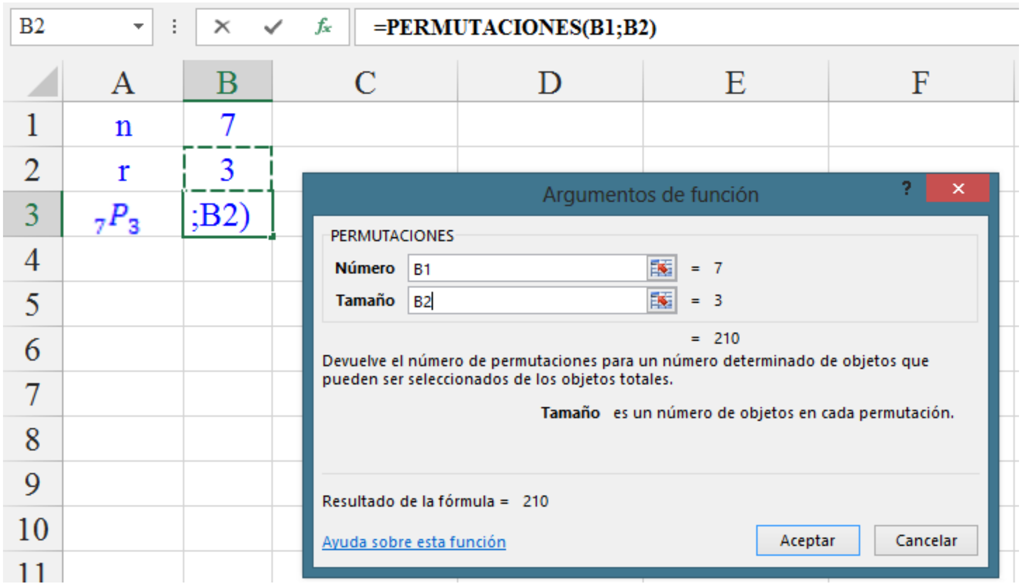
$$_7P_3$$

En Excel se calcula de la siguiente manera:

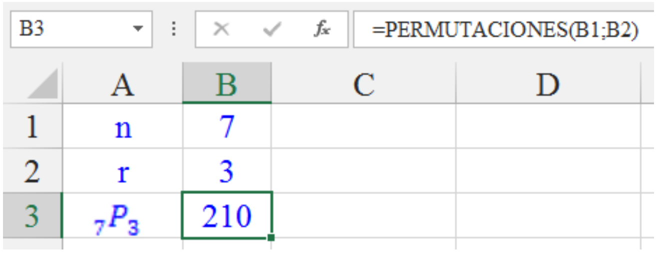
a) Insertar función. Seleccionar la categoría de Estadísticas. En función seleccionar la opción PERMUTACIONES.



b) Clic en Aceptar. En el cuadro de diálogo de Argumentos de la función, en el recuadro Número seleccionar la celda correspondiente a n (B1), en el recuadro de Tamaño seleccionar la celda correspondiente a r (B2).



c) Clic en Aceptar

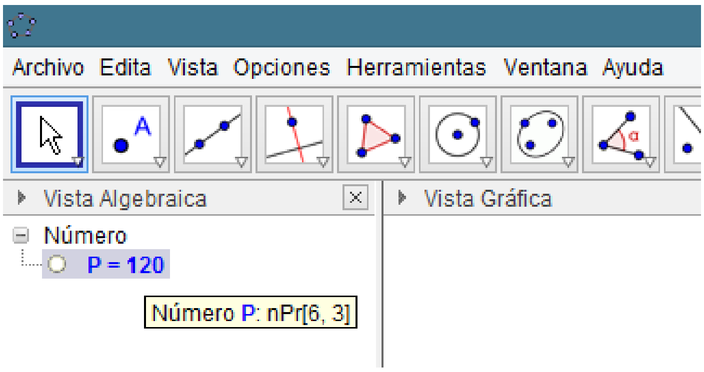


2) Si se desean ordenar 6 libros en un estante, pero sólo hay espacio para 3 libros. Calcular el número de resultados posibles de ordenar dichos libros

Solución: Como se pide calcular $_6P_3$, entonces,

$$_nP_r = \frac{n!}{(n-r)!} \Rightarrow {}_6P_3 = \frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{3\cdot 2\cdot 1} = 6\cdot 5\cdot 4 = 120$$

Los cálculos en GeoGebra se muestran en la siguientes figuras: En programa GeoGebra, en Entrada: se escribe nPr y aparece nPr[<Número>, <Número>]. Se escribe los números 6 y 3, quedando nPr[6,3]. Enter



C) COMBINACIONES

En muchos situaciones no interesa el orden de los resultados, sino sólo el número de maneras en las que r objetos pueden seleccionarse a partir de n cosas, sin consideración de orden. Si dos subconjntos se consideran iguales debido a que simplemente se han reordenado los mismos elementos, entonces se trata de combinaciones.

El número de maneras para arreglar r objetos seleccionados a la vez de n objetos, sin considerar el orden, es decir, el número de combinaciones de n elementos tomados r a la vez es:

$$ {}_{n}C_{r} = \binom{n}{r} = \frac{n(n-1)(n-2)\ldots(n-r+1)}{r!} = \frac{n!}{r!\,(n-r)!} $$

Ejemplos ilustrativos:

1) Calcular ${}_{7}C_{3}$

Solución:

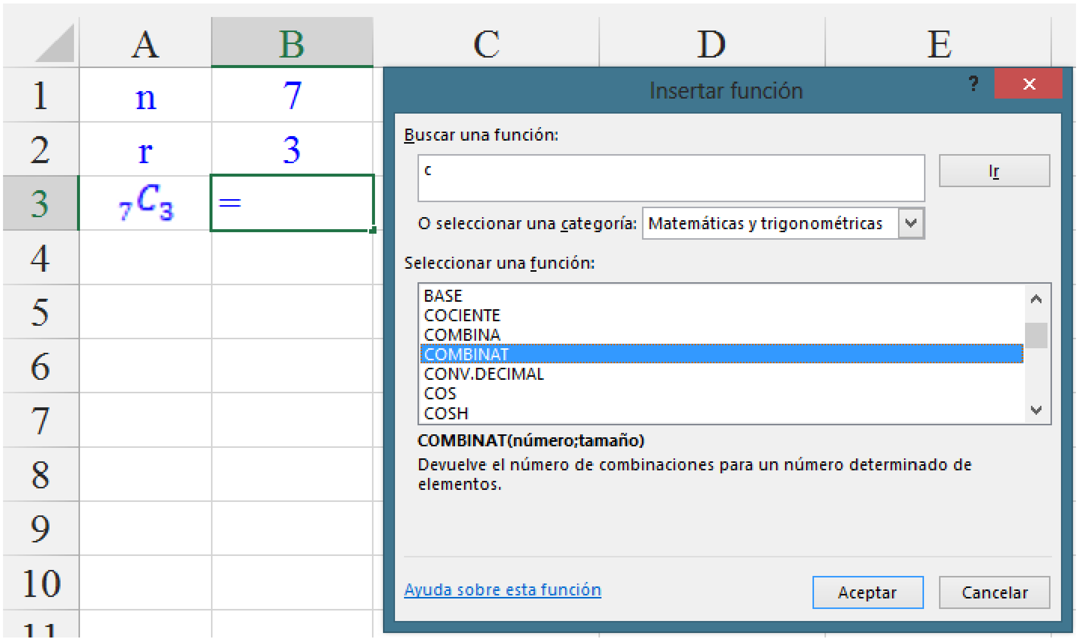
n = 7 y r = 3, entonces aplicando la fórmula se obtiene:

$$ {}_{n}C_{r} = \frac{n!}{r!\,(n-r)!} \Rightarrow {}_{7}C_{3} = \frac{7!}{3!\,(7-3)!} = \frac{7!}{3!\,4!} $$

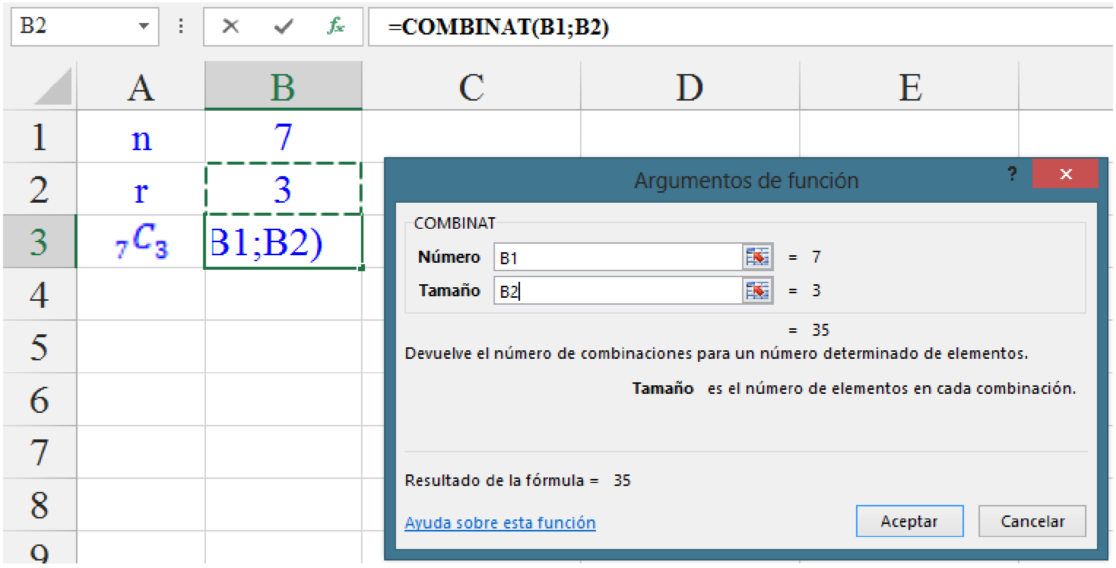
$$ {}_{7}C_{3} = \frac{7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{(3\cdot 2\cdot 1)(4\cdot 3\cdot 2\cdot 1)} = 7\cdot 5 = 35 $$

En Excel se calcula de la siguiente manera:

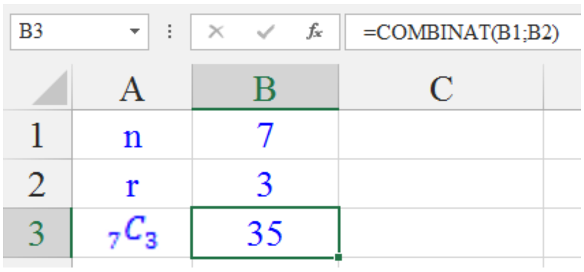
a) Insertar función. Seleccionar la categoría de Matemáticas y trigonométricas. En función seleccionar la opción COMBINAT



b) Clic en Aceptar. En el cuadro de diálogo de Argumentos de la función, en el recuadro Número seleccionar la celda correspondiente a n (B1), en el recuadro de Tamaño seleccionar la celda correspondiente a r (B2).



c) Clic en Aceptar



2) Si se desean ordenar 6 libros en un estante, pero sólo hay espacio para 3 libros. Calcular el número de resultados posibles de acomodar dichos libros sin importar el orden.

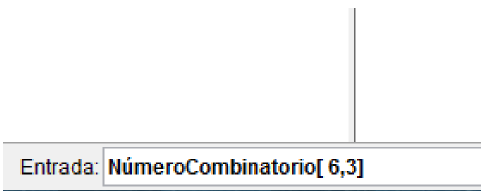
Solución:

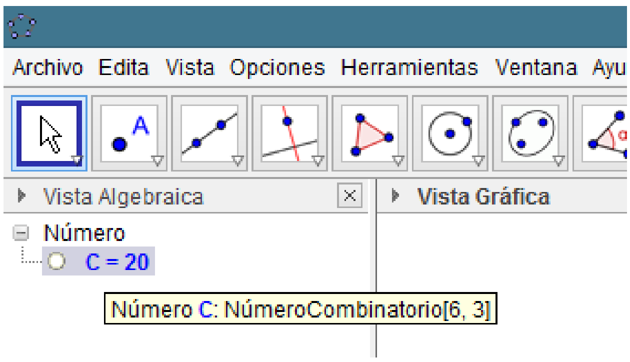
Como se pide calcular $_6C_3$, entonces,

$$_nC_r = \frac{n!}{r!\,(n-r)!} \Rightarrow \;_6C_3 = \frac{6!}{3!\,(6-3)!} = \frac{6!}{3!\,3!} = \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{(3\cdot2\cdot1)(3\cdot2\cdot1)} = 5\cdot4 = 20$$

Los cálculos en GeoGebra se muestran en la siguiente figura:

Se escribe en Entrada: Número y aparece NúmeroCombinatorio[<Número n (o valor numérico)>, <Número r (o valor numérico)>]. Digitar el 6 y el 3, y queda NúmeroCombinatorio[6,3]. Enter





TAREA DE INTERAPRENDIZAJE N° 1

1) Realice un organizador gráfico (mapa conceptual, organigrama, mentefacto, etc.) sobre el análisis combinatorio.

2) Resuelva de manera manual, empleando Excel y GeoGebra

2.1) 8!

40320

2.2) 10!

3628800

2.3) ${}_8P_3$

336

2.4) ${}_{10}P_3$

720

2.5) ${}_8C_3$

56

2.6) $\binom{10}{3}$

120

3) En la fórmula de la permutación, ¿qué valor debe tener r para que la permutación sea igual a la factorial?. Ilustre su respuesta con un ejemplo

n

4) Realice los cálculos de manera manual y empleando Excel para que compruebe las siguientes igualdades:

4.1) $5! = {}_5P_5$

4.2) $\binom{5}{0} = \binom{5}{5}$

4.3) $\binom{6}{3} - \binom{5}{3} = \binom{5}{2}$

5) Don Albertito desea parquear 3 automóviles en su garaje. Calcular el número de resultados posibles de parquear dichos automóviles. Realice los cálculos de manera manual y empleando GeoGebra

6

6) Se desea ordenar 4 libros en un estante. Calcular el número de resultados posibles de ordenar los mencionados libros. Realice los cálculos de manera manual y empleando Excel

24

7) Se desea ordenar 4 libros en un estante, pero solo hay espacio para 2 libros. Calcular el número de resultados posibles de ordenar los mencionados libros. Realice los cálculos de manera manual y empleando GeoGebra

12

8) ¿De cuántas maneras posibles se puede formar con 8 personas una comisión de 3 miembros?. Realice los cálculos de manera manual y empleando Excel

56

9) Consulte en la biblioteca o en el internet sobre la importancia de las permutaciones y combinaciones, y presente la consulta mediante un organizador gráfico.

1.2) CONCEPTOS BÁSICOS

A) EXPERIMENTO.- Es toda acción sobre la cual vamos a realizar una medición u observación, es decir cualquier proceso que genera un resultado definido.

B) EXPERIMENTO ALEATORIO.- Es toda actividad cuyos resultados no se determinan con certeza. Ejemplo: lanzar una moneda al aire. No podemos determinar con toda certeza ¿cuál será el resultado al lanzar una moneda al aire?, por lo tanto constituye un experimento aleatorio.

C) ESPACIO MUESTRAL (S).- Es un conjunto de *todos los resultados posibles* que se pueden obtener al realizar un experimento aleatorio. Ejemplo: sea el experimento E: lanzar un dado y el espacio muestral correspondiente a este experimento es: $S = \{1, 2, 3, 4, 5, 6\}$.

D) PUNTO MUESTRAL.- Es un elemento del espacio muestral de cualquier experimento dado.

E) EVENTO O SUCESO.- Es todo subconjunto de un espacio muestral. Se denotan con letras mayúsculas: A, B, etc. Los resultados que forman parte de este evento generalmente se conocen como *"resultados favorables"*. Cada vez que se observa un resultado favorable, se dice que "ocurrió" un evento. Ejemplo: Sea el experimento E: lanzar un dado. Un posible evento podría ser que salga número par. Definimos el evento de la siguiente manera: A = sale número par = $\{2, 4, 6\}$, resultados favorables $n(E) = 3$

Los eventos pueden ser:

i) Evento cierto.- Un evento es cierto o seguro si se realiza siempre. Ejemplo: Al introducirnos en el mar, en condiciones normales, es seguro que nos mojaremos.

ii) Evento imposible.- Un evento es imposible si nunca se realiza. Al lanzar un dado una sola vez, es imposible que salga un 10

iii) Evento probable o aleatorio.- Un evento es aleatorio si no se puede precisar de antemano el resultado. Ejemplo: ¿Al lanzar un dado, saldrá el número 3?

F) PROBABILIDAD.- Es el conjunto de posibilidades de que un evento ocurra o no en un momento y tiempo determinado. Dichos eventos pueden ser medibles a través de una escala de 0 a 1, donde el evento que no pueda ocurrir tiene una probabilidad de 0 (evento imposible) y un evento que ocurra con certeza es de 1 (evento cierto).

La probabilidad de que ocurra un evento, siendo ésta una medida de la posibilidad de que un suceso ocurra favorablemente, se determina principalmente de dos formas: empíricamente (de manera experimental) o teóricamente (de forma matemática).

i) Probabilidad empírica.- Si E es un evento que puede ocurrir cuando se realiza un experimento, entonces la probabilidad empírica del evento E, que a veces se le denomina *definición de frecuencia relativa de la probabilidad,* está dada por la siguiente fórmula:

$$P(E) = \frac{\textit{número de veces que ocurre el evento } E}{\textit{número de veces que se realizó el experimento}}$$

Nota: P(E), se lee probabilidad del evento E

Ejemplo ilustrativos

1) En el año 2010, nacieron en un hospital 100 hombres y 150 mujeres. Si una persona fue seleccionada aleatoriamente de los registros de nacimientos de ese año, ¿cuál es la probabilidad de que haya sido mujer?

Solución:

Ya que las probabilidades de que nazcan hombres o mujeres no son iguales, y por tener información específica experimental que respalda este hecho, se calcula empleando la fórmula de la probabilidad experimental

$$P(E) = \frac{\textit{número de veces que ocurre el evento } E}{\textit{número de veces que se realizó el experimento}}$$

$$P(Mujeres) = \frac{\textit{número de nacimientos de mujeres}}{\textit{número total de nacimientos}} = \frac{150}{100+150} = \frac{150}{250} = \frac{3}{5} = 0{,}6 = 60\%$$

Nota: la respuesta puede estar expresada como fracción, como un número decimal y como un porcentaje.

2) La siguiente tabla muestra el número de cajas y el número de artículos dañados por caja que un comerciante recibió. Calcular la probabilidad para cada resultado individual

N° de cajas	N° de artículos dañados
50	0
40	2
10	3

Solución:

Ya que las probabilidades de defectos por caja no son iguales, y por tener información específica experimental que respalda este hecho, se calcula empleando la definición de frecuencia relativa de la probabilidad.

N° de cajas	N° de artículos dañados	P(E)
50	0	P(0) = 50/100 = 1/2 = 0,5 = 50%
40	2	P(2) = 40/100 = 2/5 = 0,4 = 40%
10	3	P(3) = 10/100 = 1/10 = 0,1 = 10%
100		1 = 100%

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D
1	N° de cajas	N° de artículos dañados	P(E)	
2	50	0	0,5	=A2/$A$5
3	40	2	0,4	=A3/$A$5
4	10	3	0,1	=A4/$A$5
5	100	=SUMA(A2:A4)	1	=SUMA(C2:C4)

Nota:

La respuesta 0,5 significa que existe una probabilidad de 0,5 o del 50% de que 0 artículos en cualquier caja dada estuviera dañado

La respuesta 0,4 significa que existe una probabilidad de 0,4 o del 40% de que 2 artículos en cualquier caja dada estuviera dañado

La respuesta 0,1 significa que existe una probabilidad de 0,1 o del 10% de que 3 artículos en cualquier caja dada estuviera dañado

La suma de las probabilidades individuales siempre es igual a 1 que en porcentaje es igual al 100%

ii) Probabilidad teórica.- Si todos los resultados en un espacio muestral S finito son igualmente probables, y E es un evento en ese espacio muestral, entonces la probabilidad teórica del evento E está dada por la siguiente fórmula, que a veces se le denomina la *definición clásica de la probabilidad*, expuesta por Pierre Laplace en su famosa Teoría analítica de la probabilidad publicada en 1812:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ posibles\ resultados} = \frac{n(E)}{n(S)}$$

Ejemplos ilustrativos

1) En cierta rifa de un automóvil se venden 5000 boletos. Calcular la probabilidad de ganarse el automóvil

1.1) Si se compran 20 boletos.

1.2) Si se compran todos los boletos

1.3) Si no se compran boletos

Solución:

Ya que el espacio muestral S (5000 boletos) es finito, y los resultados de cada boleto son igualmente probables, se calcula empleando la fórmula de la definición clásica de la probabilidad

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

1.1) $P(20) = \frac{20}{5000} = \frac{1}{250} = 0{,}004 = 0{,}4\%$

1.2) $P(5000) = \frac{5000}{5000} = 1 = 100\%$

1.3) $P(0) = \frac{0}{5000} = 0 = 0\%$

2) Calcular la probabilidad de obtener un número impar en el lanzamiento de un dado

Solución:

Espacio muestral = S = {1, 2, 3, 4, 5, 6}, entonces, n(S) = 6

Resultados favorables = {1, 3, 5}, entonces, n(E) = 3

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} = \frac{3}{6} = \frac{1}{2} = 0{,}5 = 50\%$$

3) En un ánfora existe 10 fichas amarillas, 6 rojas y 4 azules.

3.1) ¿Qué probabilidad existe de sacar una ficha amarilla en un primer intento?
3.2) ¿Qué probabilidad existe de sacar una ficha no roja en un primer intento?

Solución:
n(S) = 10 + 6 + 4 = 20

3.1) n(E) = 10

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

$$P(A) = \frac{10}{20} = \frac{1}{2} = 0{,}5 = 50\%$$

3.2) Si P(E) es la probabilidad de que ocurra el evento E y $P(\bar{E})$ la probabilidad de que no ocurra el evento E. Debido a que la suma de las probabilidades siempre da como resultado 1, es decir, $P(E) + P(\bar{E}) = 1$, por lo que se tiene: $P(\bar{E}) = 1 - P(E)$,

Calculando la probabilidad de sacar una ficha roja se obtiene:

n(E) = 6
$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$
$$P(R) = \frac{6}{20} = 0{,}3$$

Calculando la probabilidad de sacar una ficha no roja se obtiene:

$$P(\bar{E}) = 1 - P(E)$$
$$P(\bar{R}) = 1 - P(R)$$
$$P(\bar{R}) = 1 - 0{,}3 = 0{,}7$$

4) En una urna existe 10 bolas numeradas con los números dígitos.

4.1) ¿Qué probabilidad existe de sacar una bola enumerada con un número múltiplo de 3?
4.2) ¿Qué probabilidad existe de sacar una bola enumerada con un número divisor de 6?

Solución:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

Espacio muestral = S = {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}, entonces, n(S) = 10

4.1)
Resultados favorables = {3, 6, 9}, entonces, n(E) = 3
$$P(Múltiplo\ de\ 3) = \frac{3}{10}$$

4.2)
Resultados favorables = {1, 2, 3, 6}, entonces, n(E) = 4

$$P(Divisor\ de\ 6) = \frac{4}{10} = \frac{2}{5} = 0{,}4 = 40\%$$

5) De una urna que contiene 2 bolas rojas y 3 azules

5.1) Se extrae una bola, calcular la probabilidad de que la bola sea
a) Roja
b) Azul

Solución:
$Número\ total\ de\ resultados\ posibles = n(S) = 2 + 3 = 5$

a) Roja (R)

$Número\ de\ resultados\ favorables = n(E) = 2$

Remplazando valores en la fórmula de la probabilidad teórica se tiene

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(R) = \frac{2}{5}$$

b) Azul (A)

$Número\ de\ resultados\ favorables = n(E) = 3$

Remplazando valores en la fórmula de la probabilidad teórica se tiene

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(A) = \frac{3}{5}$$

5.2) Se extraen simultáneamente dos bolas, calcular la probabilidad de que las dos sean

a) Azules
b) Rojas
c) Diferente color

Designando por R_1 , R_2, las bolas rojas y por A_1 , A_2, A_3 las azules se tiene el siguiente espacio muestral:

R_1R_2 , R_1A_1, R_1A_2, R_1A_3
R_2A_1, R_2A_2, R_2A_3
A_1A_2, A_1A_3
A_2A_3

Entonces, n(S) = 4 + 3+ 2+ 1 = 10

a) Azules

Resultados favorables = $\{A_1A_2,\ A_1A_3,\ A_2A_3\}$, entonces, n(E) = 3

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

$$P(AA) = \frac{3}{10}$$

Otra forma de resolver este ejercicio es la siguiente:

El espacio muestral se calcula aplicando la fórmula de la combinación, es decir,

$$n(S) = \ _nC_r = \frac{n!}{r!\,(n-r)!}$$

En donde:
n = número total de bolas = 2 + 3 = 5
r = número de bolas azules motivo de probabilidad = 2

Entonces, remplazando valores en la fórmula de la combinación se obtiene:

$$n(S) = \ _5C_2 = \frac{5!}{2!\,(5-2)!} = \frac{5!}{2!\,3!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{(2\cdot 1)(3\cdot 2\cdot 1)} = 5\cdot 2 = 10$$

El número de resultados favorables se calcula aplicando la fórmula de la combinación, es decir,

$$n(E) = \ _nC_r = \frac{n!}{r!\,(n-r)!}$$

En donde:
n = número total de bolas azules = 3
r = número de bolas azules motivo de probabilidad = 2

Entonces, remplazando valores en la fórmula de la combinación se obtiene:

$$n(E) = \ _3C_2 = \frac{3!}{2!\,(3-2)!} = \frac{3!}{2!\,1!} = \frac{3\cdot 2\cdot 1}{(2\cdot 1)(1\cdot 1)} = 3$$

Reemplazando valores en la fórmula de la probabilidad se tiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

$$P(AA) = \frac{n(E)}{n(S)} = \frac{_3C_2}{_5C_2} = \frac{3}{10}$$

Los cálculos en Excel aplicando combinaciones se muestran en la siguiente figura:

	A	B	C	D	E
1	P(AA)	3/10	=COMBINAT(3;2)/COMBINAT(5;2)		

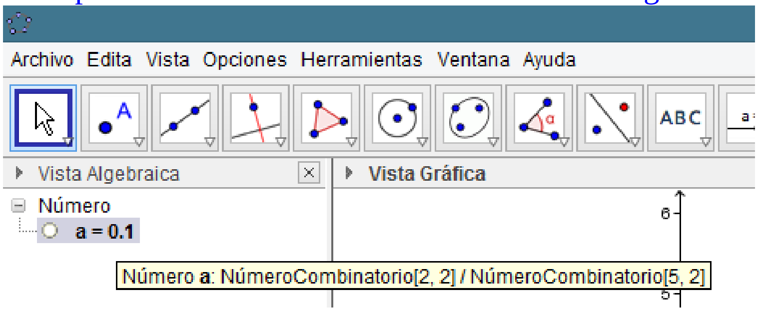
b) Rojas
Resultados favorables = $\{R_1R_2\}$, entonces, n(E) = 1

$$P(RR) = \frac{n(E)}{n(S)} = \frac{1}{10}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(RR) = \frac{n(E)}{n(S)} = \frac{{}_2C_2}{{}_5C_2} = \frac{1}{10} = 0{,}1$$

Los cálculos en GeoGebra aplicando combinaciones se muestran en la siguiente figura:



c) Diferente color
Resultados favorables = $\{R_1A_1,\ R_1A_2,\ R_1A_3,\ R_2A_1,\ R_2A_2,\ R_2A_3\}$, entonces, n(E) = 6

$$P(E) = \frac{n(E)}{n(S)} = \frac{6}{10} = \frac{3}{5}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_2C_1 \cdot {}_3C_1}{{}_5C_2} = \frac{6}{10} = \frac{3}{5}$$

Los cálculos en Excel aplicando combinaciones se muestran en la siguiente figura:

	A	B	C	D	E	F	G
1	P(E)	3/5	=COMBINAT(2;1)*COMBINAT(3;1)/COMBINAT(5;2)				

5.3) Se extraen simultáneamente tres bolas, calcular la probabilidad de que las tres sean

a) Dos rojas y una azul
b) Una roja y dos azules
c) Tres azules

Solución:
Designando por R_1, R_2, las bolas rojas y por A_1, A_2, A_3 las azules se tiene el siguiente espacio muestral:
$R_1R_2A_1, R_1R_2A_2, R_1R_2A_3$
$R_1A_1A_2, R_1A_1A_3$
$R_1A_2A_3$
$R_2A_1A_2, R_2A_1A_3$
$R_2A_2A_3$
$A_1\,A_2A_3$

Entonces, n(S) = 3 + 2 + 1 + 2 + 1 + 1 = 10

a) Dos rojas y una azul

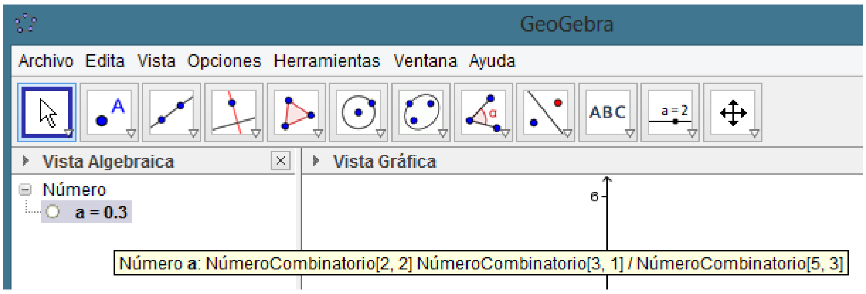
Resultados favorables = $\{R_1R_2A_1,\ R_1R_2A_2,\ R_1R_2A_3\}$, entonces, n(E) = 3

$$P(E) = \frac{n(E)}{n(S)} = \frac{3}{10}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_2C_2 \cdot {}_3C_1}{{}_5C_3} = \frac{3}{10} = 0{,}3$$

Los cálculos en GeoGebra aplicando NúmeroCombinatorio se muestran en la siguiente figura:



b) Una roja y dos azules

Resultados favorables = $\{R_1A_1A_2,\ R_1A_1A_3,\ R_1A_2A_3,\ R_2A_1A_2,\ R_2A_1A_3,\ R_2A_2A_3\}$, entonces, n(E) = 6

$$P(E) = \frac{n(E)}{n(S)} = \frac{6}{10} = \frac{3}{5}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_2C_1 \cdot {}_3C_2}{{}_5C_3} = \frac{6}{10} = \frac{3}{5}$$

c) Tres azules

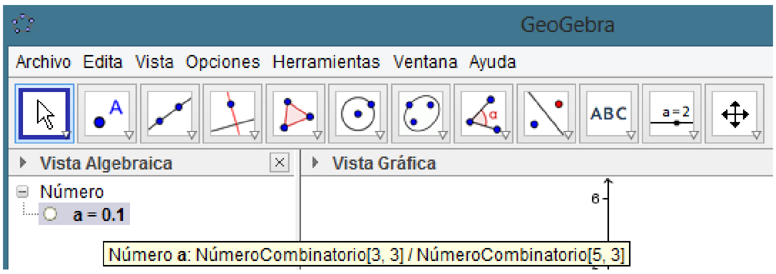
Resultados favorables = $\{A_1A_2A_3\}$, entonces, n(E) = 1

$$P(E) = \frac{n(E)}{n(S)} = \frac{1}{10}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_3C_3}{{}_5C_3} = \frac{1}{10} = 0{,}1$$

Los cálculos en GeoGebra se muestran en la siguiente figura:



5.4) Se extraen simultáneamente cuatro bolas, calcular la probabilidad de que las cuatro sean

a) Dos rojas y dos azules
b) Una roja y tres azules

Solución:

Designando por R_1, R_2, las bolas rojas y por A_1, A_2, A_3 las azules se tiene el siguiente espacio muestral:

$R_1R_2A_1A_2,\ R_1R_2A_1A_3,\ R_1R_2A_2A_3,\ R_1A_1A_2A_3,\ R_2A_1A_2A_3$

Entonces, n(S) = 5

a) Dos rojas y dos azules

Resultados favorables = $\{R_1R_2A_1A_2,\ R_1R_2A_1A_3,\ R_1R_2A_2A_3\}$, entonces, n(E) = 3

$$P(E) = \frac{n(E)}{n(S)} = \frac{3}{5}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_2C_2 \cdot {}_3C_2}{{}_5C_4} = \frac{3}{5}$$

b) Una roja y tres azules

Resultados favorables = $\{R_1A_1A_2A_3,\ R_2A_1A_2A_3\}$, entonces, n(E) = 2

$$P(E) = \frac{n(E)}{n(S)} = \frac{2}{5}$$

Otra forma de resolver este ejercicio es la siguiente:

$$P(E) = \frac{n(E)}{n(S)} = \frac{{}_2C_1 \cdot {}_3C_3}{{}_5C_4} = \frac{2}{5}$$

6) De una urna que contiene 6 bolas rojas y 5 negras se extraen simultáneamente dos bolas, calcular la probabilidad de que:

6.1) Las dos sean rojas
6.2) Las dos sean negras
6.3) De diferente color

Solución:

6.1)

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

$$P(RR)=\frac{n(E)}{n(S)}=\frac{{}_6C_2}{{}_{11}C_2}=\frac{\frac{6!}{2!\,(6-2)!}}{\frac{11!}{2!\,(11-2)!}}=\frac{\frac{6!}{2!\,4!}}{\frac{11!}{2!\,9!}}=\frac{\frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{(2\cdot1)(4\cdot3\cdot2\cdot1)}}{\frac{11\cdot10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{(2\cdot1)(9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1)}}=\frac{15}{55}=\frac{3}{11}$$

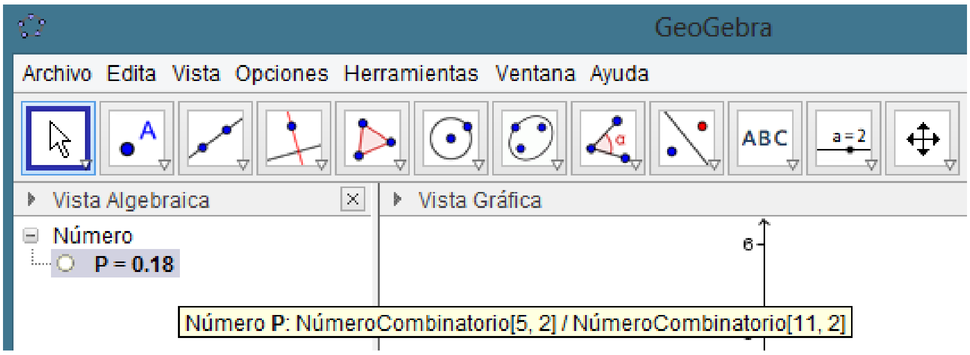
En Excel:

	A	B	C	D	E
1	P(RR)	3/11	=COMBINAT(6;2)/COMBINAT(11;2)		

6.2)

$$P(NN)=\frac{n(E)}{n(S)}=\frac{{}_5C_2}{{}_{11}C_2}=\frac{\frac{5!}{2!\,(5-2)!}}{\frac{11!}{2!\,(11-2)!}}=\frac{\frac{5!}{2!\,3!}}{\frac{11!}{2!\,9!}}=\frac{\frac{5\cdot4\cdot3\cdot2\cdot1}{(2\cdot1)(3\cdot2\cdot1)}}{\frac{11\cdot10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1}{(2\cdot1)(9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3\cdot2\cdot1)}}=\frac{10}{55}=\frac{2}{11}$$

En GeoGebra:



6.3)

$$P(E)=\frac{n(E)}{n(S)}=\frac{{}_6C_1\cdot{}_5C_1}{{}_{11}C_2}=\frac{\frac{6!}{1!\,(6-1)!}\cdot\frac{5!}{1!\,(5-1)!}}{\frac{11!}{2!\,(11-2)!}}=\frac{\frac{6!}{1!\,5!}\cdot\frac{5!}{1!\,4!}}{\frac{11!}{2!\,9!}}=\frac{30}{55}=\frac{6}{11}=0{,}55$$

7) De una urna que contiene 6 fichas rojas, 5 negras y 9 azules, Elizabeth extrae simultáneamente tres fichas, calcular la probabilidad de que las 3 fichas extraídas por Elizabeth sean:

7.1) Rojas
7.2) 2 rojas y una negra
7.3) De diferente color

Solución:

7.1) Rojas

$$P(E)=\frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}}=\frac{n(E)}{n(S)}$$

$$P(RRR)=\frac{n(E)}{n(S)}=\frac{{}_6C_3}{{}_{20}C_3}=\frac{20}{1140}=\frac{1}{57}$$

7.2) 2 rojas y una negra

$$P(E)=\frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles}=\frac{n(E)}{n(S)}$$

$$P(RRN)=\frac{n(E)}{n(S)}=\frac{{}_6C_2\cdot\ {}_5C_1}{{}_{20}C_3}=\frac{15\cdot 5}{1140}=\frac{75}{1140}=\frac{5}{76}$$

7.3) De diferente color

$$P(E)=\frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles}=\frac{n(E)}{n(S)}$$

$$P(E)=\frac{n(E)}{n(S)}=\frac{{}_6C_1\cdot\ {}_5C_1\cdot\ {}_9C_1}{{}_{20}C_3}=\frac{6\cdot 5\cdot 9}{1140}=\frac{270}{1140}=\frac{9}{38}$$

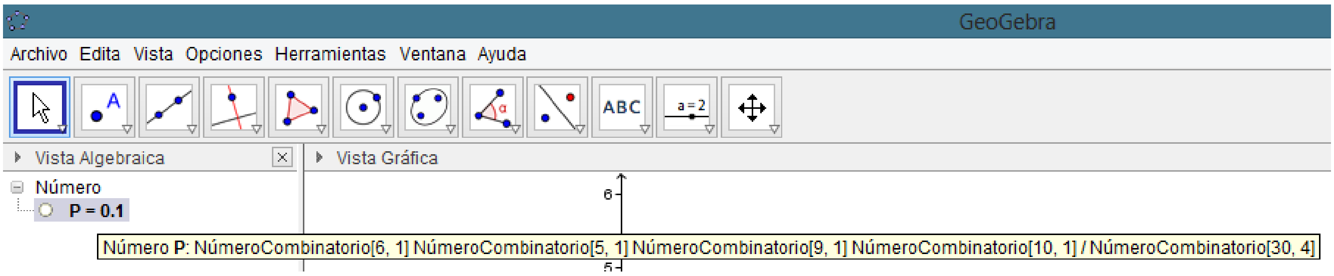
8) En una ferretería existen 6 galones de pintura roja, 5 de pintura naranja, 9 de pintura amarrillo y 10 de pintura blanca. Bertha compra aleatoriamente cuatro galones de pintura, calcular la probabilidad de que los galones comprados por Bertha sean de diferente color.

Solución:

$$P(E)=\frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles}=\frac{n(E)}{n(S)}$$

$$P(E)=\frac{n(E)}{n(S)}=\frac{{}_6C_1\cdot\ {}_5C_1\cdot\ {}_9C_1\cdot\ {}_{10}C_1}{{}_{30}C_4}=\frac{6\cdot 5\cdot 9\cdot 10}{27405}=\frac{2700}{27405}=\frac{20}{203}=0{,}09852=9{,}852\%$$

En GeoGebra:



9) Se lanzan simultáneamente tres monedas, calcular la probabilidad de que se obtengan dos caras y un sello.

Solución:

Designando por C = cara y por S = sello se tiene:

Espacio muestral = S = {CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS}, entonces, n(S) = 8

Resultados favorables = { CCS, CSC, SCC }, entonces, n(E) = 3

$$P(2C\ y\ Un\ S)=\frac{3}{8}$$

Todas las probabilidades individuales se representan en la siguiente tabla:

Monedas			n(E)	P(E)
1ra	2da	3ra		
C	C	C	1	1/8
C	C	S	3	3/8
C	S	S	3	3/8
S	S	S	1	1/8
Total			8	1

Interpretación:

La probabilidad de obtener 3 caras al lanzar simultáneamente tres monedas es de 1/8, es decir, P(CCC)= 1/8

La probabilidad de obtener 2 caras y un sello al lanzar simultáneamente tres monedas es de 3/8, es decir, P(CCS) = 3/8

La probabilidad de obtener una cara y 2 sellos al lanzar simultáneamente tres monedas es de 3/8, es decir, P(CSS) = 3/8

La probabilidad de obtener 3 sellos al lanzar simultáneamente tres monedas es de 1/8, es decir, P(SSS)= 1/8

Nota:

El número 8 (espacio muestral), se calcula empleando la ecuación 2^n

$$2^n = 2^3 = 8$$

En donde n es el número de monedas que se lanzan

Los números 1, 3, 3, 1 se calculan mediante el siguiente esquema conocido con el nombre de "Triángulo de Pascal", el cual está relacionado directamente con el Teorema del Binomio de Newton.

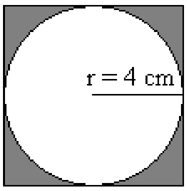
Este triángulo tiene como primera fila un 1, como segunda fila dos 1. Para las demás filas, la suma de cada par de números adyacentes de la fila anterior se ubica por debajo de ellos. Se añade un 1 en cada extremo.

Teorema del Binomio de Newton	Triángulo de Pascal
$(C+S)^0 = 1$	1
$(C+S)^1 = C + S$	1 1
$(C+S)^2 = C^2 + 2CS + S^2$	1 2 1
$\mathbf{(C+S)^3 = C^3 + 3C^2S + 3CS^2 + S^3}$	**1 3 3 1**

En donde:

$C^3 = CCC$; $3C^2S = CCS + CSC + SCC$; $3CS^2 = CSS + SCS + SSC$; $S^3 = SSS$

10) Si un dardo se clava de manera aleatoria en el objeto cuadrado que se muestra en la siguiente figura, ¿cuál es la probabilidad de que caiga en la región sombreada?



Solución:
Calculando el área del círculo:

$AO = \pi r^2 \Rightarrow AO = \pi(4cm)^2 = 3{,}14 \cdot 16cm^2 = 50{,}24cm^2$

Calculando el área del cuadrado:

Si el radio de la circunferencia es 4cm, entonces el lado del cuadrado es 8 cm, es decir,

Si $rO = 4\text{cm} \Longrightarrow \ell\square = 8\text{cm}$

Por lo tanto, el área del cuadrado es:

$A\square = \ell^2 = (8\text{cm})^2 = 64\ \text{cm}^2$

Calculando el área de la región sombreada:

Se obtiene al restar el área del círculo de la del cuadrado

$A\blacksquare = A\square - A\bigcirc$

$64cm^2 - 50{,}24cm^2 = 13{,}76\ cm^2$

Calculando la probabilidad:

$$P(E) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{n(E)}{n(S)}$$

$$P(E) = \frac{\text{área sombreda}}{\text{área total}} = \frac{13{,}76\ cm^2}{64\ cm^2} = 0{,}215 = 21{,}5\%$$

G) POSIBILIDADES

Las posibilidades comparan el número de resultados favorables con el número de resultados desfavorables. Si todos los resultados de un espacio muestral son igualmente probables, y un número n de ellos son favorables al evento E, y los restantes m son desfavorables a E, entonces las *posibilidades a favor* de E sonde de n(E) a m(E), y las *posibilidades en contra* de E son de m(E) a n(E)

Ejemplos ilustrativos:

1) A Mathías se le prometió comprar 6 libros, tres de los cuales son de Matemática. Si tiene las mismas oportunidades de obtener cualquiera de los 6 libros, determinar las posibilidades de que le compren uno de Matemática.

Solución:

Número de resultados favorables = n(E) = 3
Número de resultados desfavorables = m(E) = 3
Posibilidades a favor son n(E) a m(E), entonces,
Posibilidades a favor = 3 a 3, y simplificando 1 a 1.

Nota: A las posibilidades de 1 a 1 se les conoce como “igualdad de posibilidades” o “posibilidades de 50-50”

2) Dyanita compró 5 boletos para una rifa de su lugar de trabajo en la que el ganador recibirá un computador. Si en total se vendieron 1000 boletos y cada uno tiene la misma oportunidad de salir ganador, ¿cuáles son las posibilidades que Dyanita tiene en contra de ganarse el computador?

Solución:

Número de resultados favorables = n(E) = 5
Número de resultados desfavorables = m(E) = 1000-5 = 995
Posibilidades en contra son m(E) a n(E) , entonces,
Posibilidades en contra = 995 a 5, o de 199 a 1.

3) Emily participará en una lotería, en donde las posibilidades de ganar son de 1 a 999. ¿Cuál es la probabilidad que tiene Emily de ganar la lotería?

Solución:

Como las posibilidades a favor = 1 a 999 y se sabe que las posibilidades a favor son n(E) a m(E), entonces,
Número de resultados favorables = n(E) = 1
Número de resultados desfavorables = m(E) = 999

Como el número total de resultados posibles = n(S) = n(E) + m(E) = 1 + 999 = 1000, y aplicando la fórmula de la probabilidad:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$$

Se obtiene:

$$P(Ganar) = \frac{1}{1000} = 0{,}003 = 0{,}1\%$$

TAREA DE INTERAPRENDIZAJE N° 2

1) Realice un organizador gráfico sobre los conceptos básicos
2) Consulte sobre la biografía de Laplace y realice un organizador gráfico de la misma.
3) Consulte sobre la biografía de Blaise Pascal y realice un organizador gráfico de la misma.
4) Consulte sobre la biografía de Newton y realice un organizador gráfico de la misma.

5) Calcular las siguientes probabilidades

5.1) Obtener 4 lanzando un solo dado

1/6

5.2) Obtener una reina al extraer una carta de una baraja estándar de 52 cartas

1/13

5.3) Obtener cara al lanzar una moneda

1/2

5.4) No obtener un rey al extraer una carta de una baraja estándar de 52 cartas

12/13

6) Se tiene la información acerca de los ingresos mensuales por venta de material didáctico de la papelería D & M en los últimos 12 meses. Calcular la probabilidad para cada resultado individual de manera manual y empleando Excel

Meses	Ingresos ($)
1	500
2	600
2	700
4	800
3	900

P(500) = 1/12 ; P(600) = 1/6; P(700) = 1/6; P(800) = 1/3; P(900) = 1/4

7) En una urna existe 10 bolas numeradas con los números dígitos. ¿Qué probabilidad existe de sacar una bola enumerada con un número primo?

40%

8) En una ánfora existe fichas numeradas del 0 al 30. ¿Qué probabilidad existe de sacar una ficha enumerada con un número perfecto?

6,45%

9) Sea una caja que contiene 3 bolas rojas, 5 blancas y 4 azules, hallar la probabilidad de que al sacar una bola ésta sea:

9.1) Roja

1/4

9.2) Blanca

5/12

9.3) No Azul

2/3

10) En una mesa existen 5 cartas, de las cuales solo 2 son reyes. Se escogen simultáneamente dos cartas. Calcular la probabilidad de obtener

10.1) Dos reyes

1/10

10.2) Al menos un rey

7/10

11) La empresa D & M desea contratar 2 nuevos empleados. Si existen candidatos para los cargos 3 mujeres y 5 hombres, y a cada candidato se le da igual consideración, calcular la probabilidad de manera manual y con GeoGebra de que los 2 nuevos empleados sean:

11.1) Mujeres

10,714%

11.2) Hombres

35,714%

11.3) Diferente género

53,571 %

12) Un recipiente contiene esferográficos, 3 de color rojo, 2 de color negro y uno de color azul. Se extrae simultáneamente 3 esferográficos, calcular la probabilidad de que los 3 esferográficos sean:

12.1) Dos rojos y un azul

3/20

12.2) Dos negros y un azul

1/20

12.3) Diferente color

3/10

13) De una caja que contiene 3 bolas rojas, 5 blancas y 4 azules se extraen simultáneamente cinco bolas, calcular la probabilidad aplicando combinaciones de manera manual y con GeoGebra de que las cinco bolas sean:

13.1) Dos rojas, dos blancas y una azul

5/33

13.2) Una roja, dos blancas y dos azules

5/22

14) En un ramo de rosas existen 8 rosas de color rojo, 4 de color blanco, 5 de color amarrillo y 3 de color naranja. Mathías escoge aleatoriamente siete rosas, calcular la probabilidad de manera manual y con Excel de que las rosas escogidas por Mathías sean:

14.1) Siete rojas

0,01%

14.2) Cuatro rojas y tres amarrillas

0,903 %

14.3) Dos rojas, dos blancas, dos amarrillas y una naranja

6,502%

15) Se lanzan simultáneamente dos dados, calcular la probabilidad de que se obtenga:

15.1) Dos 4

1/36

15.2) Ningún 4

25/36

15.3) Al menos un 4

11/36

16) Se lanzan simultáneamente cuatro monedas, calcular la probabilidad de que se obtengan

16.1) 4 caras

1/16

16.2) Una cara y 3 sellos

1/4

16.3) El mismo evento

3/8

17) Se lanzan simultáneamente cinco monedas, calcular la probabilidad de que se obtengan

17.1) 2 caras y 3 sellos

5/16

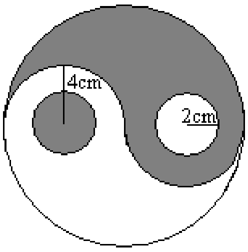
17.2) 4 caras y un sello

5/32

17.3) El mismo evento

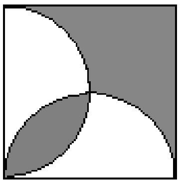
0

18) Si un dardo se clava de manera aleatoria en el objeto circular que se muestra en la siguiente figura, ¿cuál es la probabilidad de que caiga en la región sombreada?



0,5

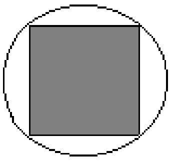
19) Si un dardo se clava de manera aleatoria en el objeto cuadrado de lado 4 cm que se muestra en la siguiente figura, ¿cuál es la probabilidad de que caiga en la región sombreada?



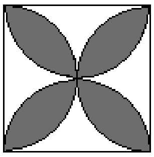
50%

20) Si un dardo se clava de manera aleatoria en el objeto circular de diámetro 4 cm que se muestra en la siguiente figura, ¿cuál es la probabilidad de que caiga en la región sombreada?

$$\frac{2}{\pi}$$



21) Si un dardo se clava de manera aleatoria en el objeto cuadrado de lado 4 cm que se muestra en la siguiente figura, ¿cuál es la probabilidad de que caiga en la región sombreada?



$$\frac{\pi - 2}{2}$$

22) Dyanita tiene una probabilidad de 0,1% de ganar una lotería. ¿Qué posibilidades de ganar la lotería tiene Dyanita?

1 a 999

23) Mathías dice a Emily: Voy a arrojar seis monedas al aire. Si todas caen cara, te daré diez centavos. Si todas caen cruz, te daré diez centavos. Pero si caen de alguna otra manera, tú me das dos centavos a mí. Calcule las posibilidades de que pierda Emily.

31 a 1

24) Plantee y resuelva 3 problemas de aplicación sobre probabilidad teórica.

25) Consulte en la biblioteca o en el internet sobre la importancia de las probabilidades, y presente la consulte a través de un organizador gráfico.

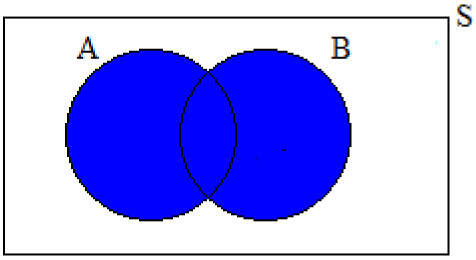
1.3) REGLAS DE LA PROBABILIDAD

A) REGLA DE LA ADICIÓN DE PROBABILIDADES

i) REGLA GENERAL PARA EVENTOS NO MUTUAMENTE EXCLUYENTES

Si A y B son dos eventos no mutuamente excluyentes (eventos intersecantes), es decir, de modo que ocurra A o bien B o ambos a la vez (al mismo tiempo), entonces se aplica la siguiente regla para calcular dicha probabilidad:

$$P(AoB) = P(A) + P(B) - P(AyB)$$
$$o$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



A U B

En donde:

El conectivo lógico "*o*" corresponde a la "*unión*" en la teoría de conjuntos (o =∪)
El conectivo "*y*" corresponde a la "*intersección*" en la teoría de conjuntos (y =∩)
El espacio muestral (S) corresponde al conjunto universo en la teoría de conjuntos

Ejemplos ilustrativos

1) Sea A el suceso de sacar un As de una baraja estándar de 52 cartas y B sacar una carta con corazón rojo. Calcular la probabilidad de sacar un As o un corazón rojo o ambos en una sola extracción.

Solución:

A y B son sucesos no mutuamente excluyentes porque puede sacarse el as de corazón rojo.
Las probabilidades son:

$$P(A) = \frac{4}{52}$$

$$P(B) = \frac{13}{52}$$

$$P(A \cap B) = \frac{1}{52}$$

Reemplazando los anteriores valores en la regla general de la adición de probabilidades para eventos no mutuamente excluyentes se obtiene:

$$P(AoB) = P(A) + P(B) - P(A \cap B)$$

$$P(AoB) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{4}{13}$$

2) En una urna existe 10 bolas numeradas del 1 al 10. ¿Qué probabilidad existe de sacar en una sola extracción una bola enumerada con un número par o con un número primo?

Solución:

$Espacio\ muestral = S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \Longrightarrow n(S) = 10$

A = número par = {2, 4, 6, 8, 10}

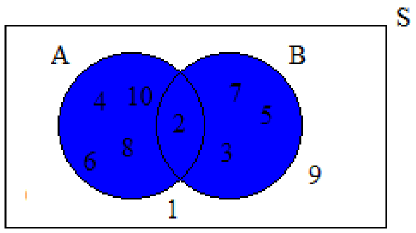
B = número primo = {2, 3, 5, 7}

$Resultados\ favorables = A \cup B = \{2\ ,4, 6, 8, 10, 3, 5, 7\} \Rightarrow n(E) = 8$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(A\ o\ B) = \frac{8}{10} = \frac{4}{5}$$

O también, realizando un diagrama de Venn-Euler se obtiene:



A U B

$$P(A) = \frac{5}{10}$$
$$P(B) = \frac{4}{10}$$
$$P(A \cap B) = \frac{1}{10}$$

Entonces, aplicando la regla para eventos no mutuamente excluyentes se obtiene:

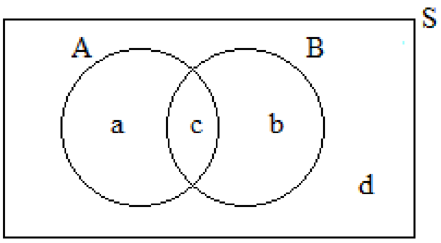
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

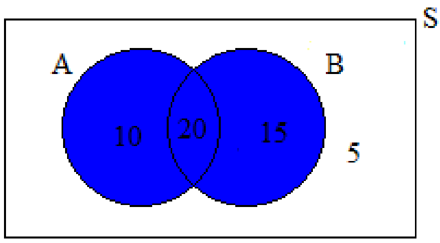
$$P(A \cup B) = \frac{5}{10} + \frac{4}{10} - \frac{1}{10} = \frac{8}{10} = \frac{4}{5}$$

3) En una clase, 10 alumnos tienen como preferencia solamente la asignatura de Matemática, 15 prefieren solamente Estadística, 20 prefieren Matemática y Estadística, y 5 no tienen preferencia por ninguna de estas asignaturas. Calcular la probabilidad que de un alumno de la clase seleccionado al azar tenga preferencia por Matemática o Estadística o ambas asignaturas.

Solución:

Realizando un diagrama de Venn-Euler se obtiene:





A U B

Simbología:
S = espacio muestral
A= Matemática
B = Estadística
a = Solamente Matemática
b = Solamente Estadística
c = Matemática y Estadística
d = Ninguna de las dos asignaturas

Datos y cálculos:
$$Número\ total\ de\ resultados\ posibles = n(S) = 10 + 20 + 15 + 5 = 50$$
$$Número\ de\ resultados\ favorables = n(E) = 10 + 20 + 15 = 45$$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(A\ o\ B) = \frac{45}{50} = \frac{9}{10}$$

O también, observando el diagrama de Venn-Euler se tiene que:

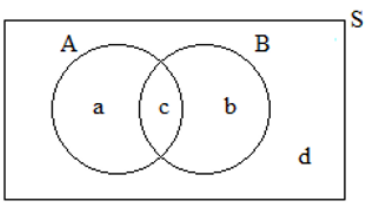
$$P(A) = \frac{30}{50}$$
$$P(B) = \frac{35}{10}$$
$$P(A \cap B) = \frac{20}{50}$$

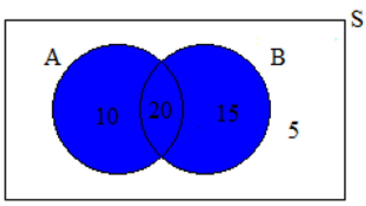
Aplicando la regla para eventos no mutuamente excluyentes se obtiene:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cup B) = \frac{30}{50} + \frac{35}{50} - \frac{20}{50} = \frac{9}{10}$$

Los cálculos en Excel se muestran en la siguiente figura:

	B	C	D	E	F
1	n (E)		P (E)		
2	10				
3	15				
4	20		2/5		
5	5				
6	50	=B2+B3+B4+B5			
7	30	=B2+B4	3/5	=B7/$B$6	
8	35	=B3+B4	7/10	=B8/$B$6	
9	45	=B2+B3+B4	**9/10**	=B9/$B$6	$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)}$
10			**9/10**	=D7+D8-D4	$P(A \cup B) = P(A) + P(B) - P(A \cap B)$



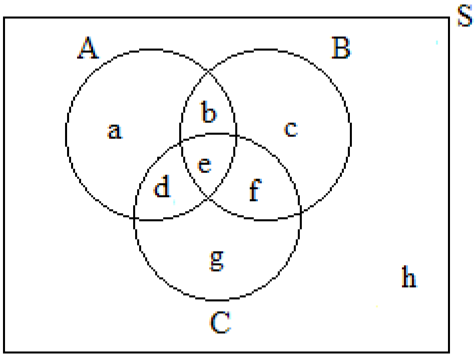


A U B

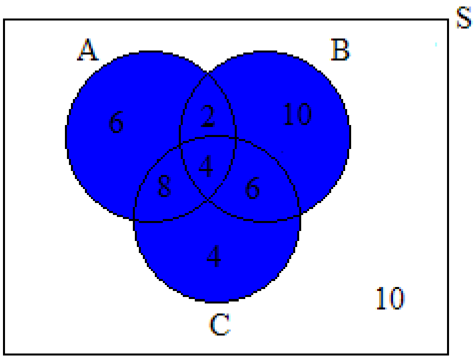
4) En un grupo de 50 personas, 6 tienen como preferencia solamente el color amarrillo, 10 prefieren solamente el color blanco, 6 prefieren el color amarrillo y blanco, 10 prefieren el color blanco y café, 12 prefieren el color amarrillo y café, 4 prefieren los 3 colores y 10 no tienen preferencia por ninguno de los tres colores.

4.1) Elaborar un diagrama de Venn-Euler
4.2) Calcular la probabilidad que de una persona del grupo seleccionada al azar tenga preferencia por lo menos uno de los tres colores.

Solución:



4.1)
Simbología:
S = espacio muestral
A= amarrillo
B= blanco
C = café
a = Solamente amarrillo
b = Amarrillo y blanco, pero no café
c = Solamente blanco
d = Amarrillo y café, pero no blanco
e = Los 3 colores
f = Blanco y café, pero no amarrillo
g = Solamente café
h = Ninguno de los tres colores



AUBUC

Datos y cálculos:
a = 6
c = 10
e = 4
h = 10
b + e = 6 → b = 6 – e = 6 – 4 = 2
e + f = 10 → f = 10 – e = 10 – 4 = 6
d + e = 12 → d = 12 – e = 12 – 4 = 8
S = 50 = a + b + c + d+ e + f+ g + h
g = 50 – a – b – c – d – e – f – h = → g = 50 – 6 – 2 – 10 – 8 – 4 – 6 – 10 = 4

4.2)

$\textit{Número total de resultados posibles} = n(S) = 50$
$\textit{Número de resultados favorables} = n(E) = 6 + 2 + 10 + 8 + 4 + 6 + 4 = 40$
Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{\textit{número de resultados favorables}}{\textit{número total de resultados posibles}} = \frac{n(E)}{n(S)} \Rightarrow P(A\ o\ B\ o\ C) = \frac{40}{50} = \frac{4}{5}$$

Nota:

Si A, B y C son tres eventos cualesquiera de modo que ocurra A o bien B o bien C o bien los tres a la vez se emplea la regla:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

Observando el diagrama de Venn-Euler se tiene que:

$$P(A) = \frac{6+2+4+8}{50} = \frac{20}{50} = \frac{2}{5}$$

$$P(B) = \frac{10+6+4+2}{50} = \frac{22}{50} = \frac{11}{25}$$

$$P(C) = \frac{4+8+4+6}{50} = \frac{22}{50} = \frac{11}{25}$$

$$P(A \cap B) = \frac{2+4}{50} = \frac{6}{50} = \frac{3}{25}$$

$$P(B \cap C) = \frac{6+4}{50} = \frac{10}{50} = \frac{1}{5}$$

$$P(A \cap C) = \frac{8+4}{50} = \frac{12}{50} = \frac{6}{25}$$

$$P(A \cap B \cap C) = \frac{4}{50} = \frac{2}{25}$$

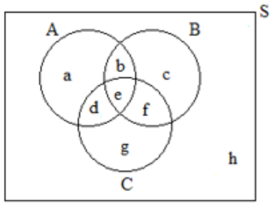
Reemplazando valores en la regla se obtiene:

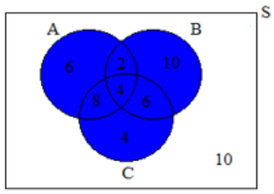
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

$$P(A \cup B \cup C) = \frac{2}{5} + \frac{11}{25} + \frac{11}{25} - \frac{3}{25} - \frac{1}{5} - \frac{6}{25} + \frac{2}{25} = \frac{4}{5}$$

Los cálculos en Excel se muestran en la siguiente figura:

	B	C	D	E	F
1	n (E)		P (E)		
2	6				
3	2	=B10-B6			
4	10				
5	8	=B12-B6			
6	4		2/25	=B6/B13	
7	6	=B11-B6			
8	4	=B13-B2-B3-B4-B5-B6-B7-B9			
9	10				
10	6		3/25	=B10/$B$13	
11	10		1/5	=B11/$B$13	
12	12		6/25	=B12/$B$13	
13	50				
14	20	=B2+B3+B5+B6	2/5	=B14/$B$13	
15	22	=B3+B4+B6+B7	11/25	=B15/$B$13	
16	22	=B5+B6+B7+B8	11/25	=B16/$B$13	
17	40	=B2+B3+B4+B5+B6+B7+B8	**4/5**	=B17/$B$13	$P(E) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{n(E)}{n(S)}$
18			**4/5**	=D14+D15+D16-D10-D11-D12+D6	
19			P(AUBUC) = P(A) + P(B) + P(C) - P(A∩B) - P(B∩C) - P(A∩C) + P(A∩B∩C)		





AUBUC

5) En una clase hay 45 estudiantes. Cada estudiante practica un solo deporte. La siguiente tabla muestra los diferentes deportes y el género de los estudiantes que lo practican.

Deporte / Estudiante	Fútbol	Básquet	Atletismo	Total
Hombre	15	3	3	21
Mujer	5	12	7	24
Total	20	15	10	**45**

Si se elige un estudiante al azar, calcular la probabilidad de que:

5.1) Sea hombre o practique fútbol
5.2) Sea mujer o practique fútbol
5.3) Sea hombre o practique básquet
5.4) Sea mujer o practique atletismo

Solución:

A partir de la tabla anterior, llamada *tabla de contingencia*, se elabora una *tabla de probabilidades*, la cual se realiza dividiendo cada una de las entradas de la tabla de contingencia por el total. Los resultados se muestran en la siguiente tabla de probabilidades:

Deporte / Estudiante	Fútbol	Básquet	Atletismo	Total
Hombre	15/45 = 1/3	3/45 = 1/15	3/45 = 1/15	21/45 = 7/15
Mujer	5/45 = 1/9	12/45 = 4/15	7/45	24/45 = 8/15
Total	20/45 = 4/9	15/45 = 1/3	10/45 = 2/9	**45/45 = 1**

Interpretación:

Los valores en las márgenes de la tabla (4/9, 1/3, 2/9, 7/15 y 8/15) se llaman *probabilidades marginales*, así por ejemplo, la probabilidad de seleccionar al azar un estudiante que sea mujer es P(M) = 8/15 y la probabilidad se seleccionar al azar un estudiante que practique atletismo es P(A) = 2/9.

Las *probabilidades conjuntas* en las celdas de la estructura principal de la tabla (1/3, 1/15, 1/15,1/9, 4/15 y 7/45) representan la probabilidad de la intersección entre dos eventos, así por ejemplo, la probabilidad de seleccionar un estudiante hombre que practique fútbol es P(H∩F) = 1/3.

Una probabilidad marginal se calcula sumando las probabilidades conjuntas correspondientes, así por ejemplo, la probabilidad marginal de seleccionar al azar un estudiante que sea mujer es
P(M) = P(M∩F) + P(M∩B) + P(M∩A), es decir, P(M) = 1/9 + 4/15 + 7/45 = 8/15

La suma de las probabilidades marginales verticales y horizontales da como resultado la unidad, así por ejemplo, P(H) + P(M) = 7/15 + 8/15 = 1 y P(F) + P(B) + P(A) = 4/9 + 1/3 + 2/9 = 1

5.1) Sea hombre o practique fútbol: $P(H\ o\ F) = P(H \cup F)$

$$P(H \cup F) = P(H) + P(F) - P(H \cap F)$$
$$P(H \cup F) = \frac{7}{15} + \frac{4}{9} - \frac{1}{3} = \frac{26}{45}$$

5.2) Sea mujer o practique fútbol: $P(M\ o\ F) = P(M \cup F)$

$$P(M \cup F) = P(M) + P(F) - P(M \cap F)$$
$$P(M \cup F) = \frac{8}{15} + \frac{4}{9} - \frac{1}{9} = \frac{13}{15}$$

5.3) Sea hombre o practique básquet: $P(H\ o\ B) = P(H \cup B)$

$$P(H \cup B) = P(H) + P(B) - P(H \cap B)$$
$$P(H \cup B) = \frac{7}{15} + \frac{1}{3} - \frac{1}{15} = \frac{11}{15}$$

5.4) Sea mujer o practique atletismo: $P(M\ o\ A) = P(M \cup A)$

$$P(M \cup A) = P(M) + P(A) - P(M \cap A)$$
$$P(M \cup A) = \frac{8}{15} + \frac{2}{9} - \frac{7}{45} = \frac{3}{5}$$

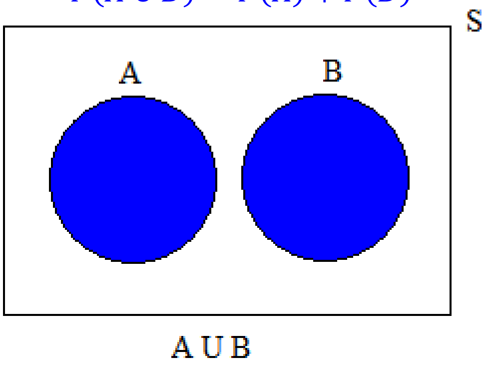
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1		Fúltbol	Básquet	Altletismo	Total
2	Hombre	15	3	3	21
3	Mujer	5	12	7	24
4	Total	20	15	10	**45**
5					
6		Fúltbol	Básquet	Altletismo	Total
7	Hombre	1/3	1/15	1/15	7/15
8	Mujer	1/9	4/15	7/45	8/15
9	Total	4/9	1/3	2/9	1
10					
11		Fúltbol	Básquet	Altletismo	Total
12	Hombre	=B2/$E$4	=C2/$E$4	=D2/$E$4	=E2/$E$4
13	Mujer	=B3/$E$4	=C3/$E$4	=D3/$E$4	=E3/$E$4
14	Total	=B4/$E$4	=C4/$E$4	=D4/$E$4	=E4/$E$4
15					
16	P(H U F)	26/45	=E7+B9-B7		
17	P(M U F)	13/15	=E8+B9-B8		
18	P(H U B)	11/15	=E7+C9-C7		
19	P(M U A)	3/5	=E8+D9-D8		

ii) REGLA PARTICULAR O ESPECIAL PARA EVENTOS MUTUAMENTE EXCLUYENTES

Si A y B son dos eventos mutuamente excluyentes (eventos no intersecantes), es decir, si la ocurrencia de cualquiera de ellos excluye la del otro, no pueden ocurrir a la vez, o cuando no tienen ningún punto muestral en común ($A \cap B = \emptyset$), entonces se aplica la siguiente regla para calcular dicha probabilidad:

$$P(AoB) = P(A) + P(B)$$
$$o$$
$$P(A \cup B) = P(A) + P(B)$$



En donde:

El conectivo lógico "*o*" corresponde a la "*unión*" en la teoría de conjuntos (o =∪)
El espacio muestral (S) corresponde al conjunto universo en la teoría de conjuntos

Ejemplos ilustrativos

1) Sea A el suceso de sacar un As de una baraja estándar de 52 cartas y B sacar un Rey de corazón rojo. Calcular la probabilidad de sacar un As o un Rey de corazón rojo en una sola extracción.

Solución:
A y B son sucesos mutuamente excluyentes porque no es posible obtener ambos a la vez.
Las probabilidades son:

$$P(A) = \frac{4}{52}$$
$$P(B) = \frac{1}{52}$$

Remplazando los anteriores valores en la regla particular de la adición de probabilidades para eventos mutuamente excluyentes se obtiene:

$$P(AoB) = P(A) + P(B)$$
$$P(AoB) = \frac{4}{52} + \frac{1}{52} = \frac{5}{52}$$

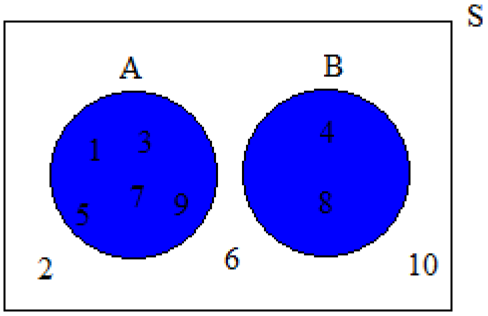
2) En una urna existe 10 bolas numeradas del 1 al 10. ¿Qué probabilidad existe de sacar en una sola extracción una bola enumerada con un número impar o con un número múltiplo de 4?

Solución:
$Espacio\ muestral = S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \Longrightarrow n(S) = 10$
A = número impar = {1, 3, 5, 7, 9}
B = número múltiplo de 4 = {4, 8}
$Resultados\ favorables = A \cup B = \{1, 3, 5, 7, 9, 4, 8\} \Rightarrow n(E) = 7$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(AoB) = \frac{7}{10}$$

O también, realizando un diagrama de Venn-Euler se obtiene:

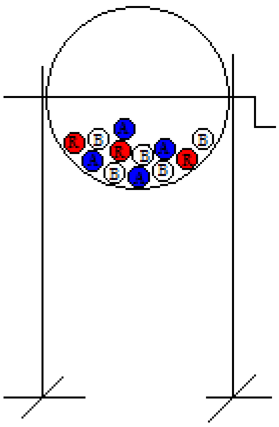


A U B

$$P(A) = \frac{5}{10}$$
$$P(B) = \frac{2}{10}$$

Entonces, aplicando la regla para eventos mutuamente excluyentes se obtiene:
$$P(A \cup B) = P(A) + P(B)$$
$$P(A \cup B) = \frac{5}{10} + \frac{2}{10} = \frac{7}{10}$$
3) De una tómbola que contiene 3 bolas rojas, 5 blancas y 4 azules, Mathías extrae una bola, calcular la probabilidad de que la bola extraída sea:

3.1) Roja o Blanca
3.2) Roja o Azul
3.3) Blanca o Azul



Solución:
R = Roja
B = Blanca
A = Azul

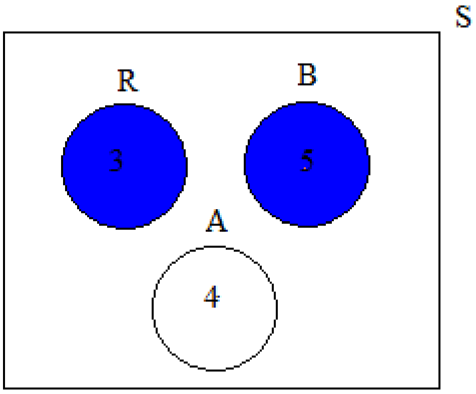
$$Número\ total\ de\ resultados\ posibles = n(S) = 3 + 5 + 4 = 12$$

3.1) Roja o Blanca (R o B)
$Número\ de\ resultados\ favorables = R \cup B = n(E) = 3 + 5 = 8$

Aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(R\ o\ B) = \frac{8}{12} = \frac{2}{3}$$

O también, realizando un diagrama de Venn-Euler se obtiene:



R U B

$$P(R) = \frac{3}{12} = \frac{1}{4}$$
$$P(B) = \frac{5}{12}$$

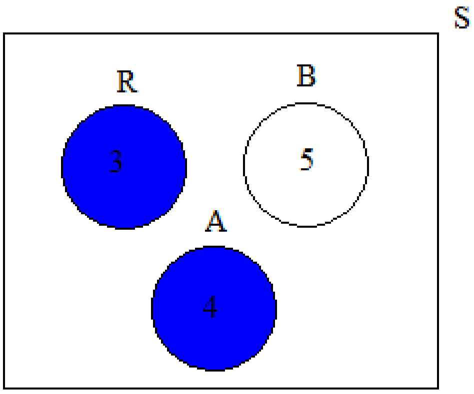
Entonces, aplicando la regla para eventos mutuamente excluyentes se obtiene:

$$P(R \cup B) = P(R) + P(B)$$
$$P(R \cup B) = \frac{1}{4} + \frac{5}{12} = \frac{2}{3}$$

3.2) Roja o Azul (R o A)

$$Número\ de\ resultados\ favorables = R \cup A = n(E) = 3 + 4 = 7$$
Aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(R\ o\ A) = \frac{7}{12}$$
O también, realizando un diagrama de Venn-Euler se obtiene:



R U A

$$P(R) = \frac{3}{12} = \frac{1}{4}$$
$$P(A) = \frac{4}{12} = \frac{1}{3}$$

Entonces, aplicando la regla para eventos mutuamente excluyentes se obtiene:
$$P(R \cup A) = P(R) + P(A)$$
$$P(R \cup A) = \frac{1}{4} + \frac{1}{3} = \frac{7}{12}$$

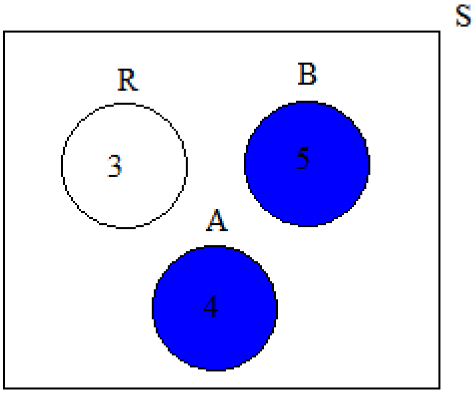
3.3) Blanca o Azul (B o A)

$Número\ de\ resultados\ favorables = B \cup A = n(E) = 5 + 4 = 9$

Aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(B\ o\ A) = \frac{9}{12} = \frac{3}{4}$$

O también, realizando un diagrama de Venn-Euler se obtiene:



B U A

$$P(B) = \frac{5}{12}$$
$$P(A) = \frac{4}{12} = \frac{1}{3}$$

Entonces, aplicando la regla para eventos mutuamente excluyentes se obtiene:

$$P(B \cup A) = P(B) + P(A)$$
$$P(B \cup A) = \frac{5}{12} + \frac{1}{3} = \frac{3}{4}$$

Los cálculos realizados en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	E	n(E)		P (E)	
2	R	3		1/4	=B2/$B$5
3	B	5		5/12	=B3/$B$5
4	A	4		1/3	=B4/$B$5
5	S	12	=SUMA(B2:B4)		
6		P (E)		P (E)	
7	R o B	2/3	=(B2+B3)/B5	2/3	=D2+D3
8	R o A	7/12	=(B2+B4)/B5	7/12	=D2+D4
9	B o A	3/4	=(B3+B4)/B5	3/4	=D3+D4

TAREA DE INTERAPRENDIZAJE N° 3

1) Consulte sobre la biografía de Euler y realice un organizador gráfico de la misma.

2) Sea A el suceso de sacar un Rey de una baraja estándar de 52 cartas y B sacar una carta con corazón rojo. Calcular la probabilidad de sacar un Rey o un corazón rojo o ambos en una sola extracción.

4/13

3) Sea A el suceso de sacar un Rey de una baraja estándar de 52 cartas y B sacar una Reina de corazón rojo. Calcular la probabilidad de sacar un Rey o Reina de corazón rojo en una sola extracción.

5/52

4) Sea A el suceso de sacar una Reina de una baraja estándar de 52 cartas y B sacar una carta con corazón negro. Calcular la probabilidad de sacar una Reina o un corazón negro o ambas en una sola extracción.

4/13

5) Sea A el suceso de sacar una Reina de una baraja estándar de 52 cartas y B sacar un As. Calcular la probabilidad de sacar una Reina o un As en una sola extracción.

2/13

6) Plantee y resuelva dos problemas de aplicación similares a los anteriores.

7) En una urna existe 10 bolas numeradas del 1 al 10. Elabore un diagrama de Venn-Euler y calcule la probabilidad de obtener en una sola extracción una bola enumerada con un número impar o con un número múltiplo de 5.

3/5

8) Plantee y resuelva un problema similar al anterior.

9) En una urna existe 10 bolas numeradas con los números dígitos. Elabore un diagrama de Venn-Euler y calcule la probabilidad de obtener en una sola extracción una bola enumerada con un número par o con un número divisor de 9.

7/10

10) De una tómbola que contiene 5 bolas rojas, 3 blancas y 2 azules, se extrae una bola. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel de que la bola extraída sea:

10.1) Roja o Blanca

4/5

10.2) Roja o Azul

7/10

10.3) Blanca o Azul

1/2

11) Plantee y resuelva un problema similar al anterior.

12) Un dado tiene tres caras blancas numeradas 4, 5 y 6, y tres caras rojas numeradas 1, 2 y 3. Si se lanza una vez este dado. Elabore un diagrama de Venn-Euler y calcule la probabilidad de obtener un número par o una cara blanca.

2/3

13) En un grupo de jóvenes, 22 estudian, 7 solamente estudian, 8 solamente trabajan y 10 no estudian ni trabajan. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de un joven seleccionado al azar estudie o trabaje o ambas actividades a la vez.

3/4

14) A la empresa D & M, 15 trabajadores se trasladan solamente en vehículo particular, 18 en transporte público, 10 solamente en transporte público y 7 se trasladan mediante otros medios. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de un trabajador seleccionado al azar se transporte a la empresa D & M en vehículo particular o transporte público o ambos transportes a la vez.

33/40

15) Plantee y resuelva un problema similar al anterior de manera manual y empleando Excel.

16) En un grupo de deportistas, 18 practican el básquet, 8 el atletismo y fútbol, 4 solamente el fútbol, 4 solamente el atletismo, 6 fútbol y básquet, pero no atletismo, 2 atletismo y básquet, pero no fútbol, 2 practican los tres deportes, y 8 practican otros deportes. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de un deportista seleccionado al azar practique por lo menos uno de estos tres deportes.

4/5

17) En un grupo de fábricas, 18 confeccionan ropa deportiva, 12 ropa deportiva y formal, 10 ropa casual y formal, 9 ropa deportiva y casual, 2 solamente ropa formal, 4 solamente ropa casual, 5 ropa deportiva y casual, pero no formal, y 10 confeccionan otros artículos. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de una fábrica seleccionada al azar confeccione por lo menos una de estas tres tipos de ropas.

3/4

18) En un grupo de 70 personas, 15 solamnete prefieren el color amarrillo, 10 solamente prefieren el color rojo, 20 solamente prefieren el color verde, 13 prefieren el color amarillo y rojo, 11 prefieren el color rojo y verde, 10 prefieren el color amarillo y verde, y 7 prefieren otros colores. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de una persona seleccionada al azar tenga por preferencia por lo menos uno de los tres colores.

63/70

19) La siguiente tabla muestra el nombre y la edad de los integrantes de una familia ecuatoriana

Nombre	Emily	Mathías	Dyanita	Mario	Bertha	Segundo	Victoria	Alberto	Carmen
Edad	11 meses	5 años	37 años	35 años	66 años	64 años	64 años	65 años	78 años

19.1) Llene la siguiente tabla de contingencia

Edad Género	Mayor de edad	Menor de edad	Total
Hombre			
Mujer			
Total			

19.2) A partir de la tabla anterior elabore de manera manual y empleando Excel una tabla de probabilidades

19.3) Si se elige una persona al azar, calcule la probabilidad de manera manual y empleando Excel de que sea hombre o menor de edad

5/9

19.4) Si se elige una persona al azar, calcule la probabilidad de manera manual y empleando Excel de que sea mujer o mayor de edad

8/9

20) El personal que labora en una institución educativa es de 50. Cada persona desempeña solo un cargo. La siguiente tabla muestra los diferentes cargos y el género del personal.

Cargo / Género	Docente	Administrativo	Auxiliar	Total
Hombre	23	1	2	26
Mujer	16	6	2	24
Total	39	7	4	**50**

Si se elige una persona al azar, calcule la probabilidad de manera manual y empleando Excel de que:

20.1) Sea hombre o docente

21/25

20.2) Sea mujer o docente

47/50

20.3) Sea hombre o auxiliar

14/25

20.4) Sea mujer o administrativo

1/2

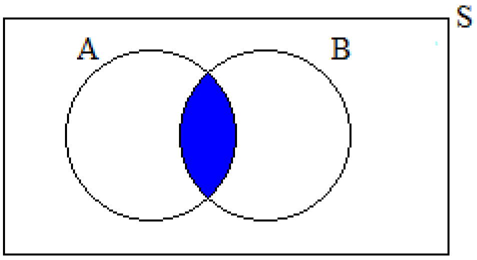
21) Plantee y resuelva de manera manual y empleando Excel un problema similar al anterior con datos de la institución educativa en la cual usted estudia.

B) REGLA DE LA MULTIPLICACIÓN DE PROBABILIDADES

i) REGLA GENERAL PARA EVENTOS DEPENDIENTES

Si A y B son dos eventos dependientes, es decir, si la ocurrencia de A afecta la probabilidad de ocurrencia de B, entonces, dicha probabilidad de calcula empleando la siguiente regla:

$$P(AyB) = P(A) \cdot P(B/A)$$
$$o$$
$$P(A \cap B) = P(A) \cdot P(B/A)$$



A ∩ B

En donde:
El conectivo "*y*" corresponde a la "*intersección*" en la teoría de conjuntos (y =∩)
El espacio muestral (S) corresponde al conjunto universo en la teoría de conjuntos
$P(B/A) = Probabilidad\ condicional\ de\ B, dado\ A$

Nota:
La probabilidad del evento B, calculada bajo la suposición de que el evento A ha ocurrido, se denomina *probabilidad condicional de B, dado A*, y se denota por P (B/A).
Si se desea obtener una fórmula para calcular la probabilidad condicional se despeja de la fórmula general de la multiplicación$P(A \cap B) = P(A) \cdot P(B/A)$, obteniéndose:

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

La probabilidad condicional de A dado B se denota por P(A/B) y se calcula empleando la siguiente fórmula:

$$P(A/B) = \frac{P(B \cap A)}{P(B)}$$

Ejemplos ilustrativos

1) De una baraja estándar de 52 cartas sea A el suceso de sacar un As en la primera extracción y B sacar un As en la segunda extracción. Calcular la probabilidad de sacar dos Ases en dos extracciones sin devolver la carta extraída.

Solución:

A y B son sucesos dependientes porque la ocurrencia de A afecta la probabilidad de ocurrencia de B.

La probabilidad de que la primera carta sea un As es:
$P(A) = \dfrac{4}{52}$

La probabilidad de que la segunda carta sea un As, dado que ya se sacó un As en la primera extracción, es:

$$P(B/A) = \frac{3}{51}$$

Remplazando los anteriores valores en la regla general de la multiplicación de probabilidades para eventos dependientes se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$
$$P(AyB) = \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}$$

2) Sea A el suceso de sacar un As de una baraja estándar de 52 cartas y B sacar un Rey de corazón rojo. Calcular la probabilidad de sacar un As y un Rey de corazón rojo en dos extracciones sin devolver la carta extraída.

Solución:

A y B son sucesos dependientes porque la ocurrencia de A afecta la probabilidad de ocurrencia de B.

La probabilidad de que la primera carta sea un As es:

$$P(A) = \frac{4}{52}$$

La probabilidad de que la segunda carta sea un Rey, dado que ya se sacó un As en la primera extracción es:

$$P(B/A) = \frac{1}{51}$$

Remplazando los anteriores valores en la regla general de la multiplicación de probabilidades para eventos dependientes se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$
$$P(AyB) = \frac{4}{52} \cdot \frac{1}{51} = \frac{1}{663}$$

3) En una urna existe 10 bolas numeradas del 1 al 10. ¿Qué probabilidad existe de sacar en una sola extracción una bola enumerada con un número par y primo?

Solución:

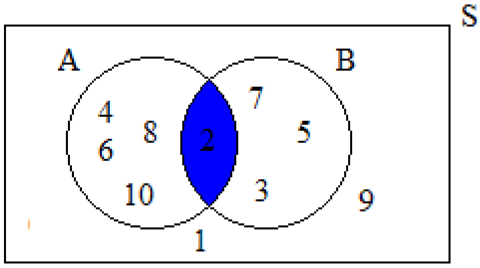
$$Espacio\ muestral = S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \Longrightarrow n(S) = 10$$

A = número par = {2, 4, 6, 8, 10}
B = número primo = {2, 3, 5, 7}
$Resultados\ favorables = A \cap B = \{2\} \Rightarrow n(E) = 1$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(A\ y\ B) = \frac{1}{10}$$

O también, realizando un diagrama de Venn-Euler se obtiene directamente la probabilidad solicitada



A ∩ B

$$P(A \cap B) = \frac{1}{10}$$

Para aplicar la regla de la multiplicación para eventos dependientes, observando el diagrama de Venn-Euler se tiene:

$$P(A) = \frac{5}{10}$$

La suposición de que la bola seleccionada esté numerada con un número par significa que sólo consideremos el conjunto A = número par = {2, 4, 6, 8, 10}, de estos 5 elementos, sólo uno, el número 2 es primo. Por lo tanto la probabilidad condicional P(B/A) = 1/5

Remplazando valores en la regla de la multiplicación para eventos dependientes se obtiene:

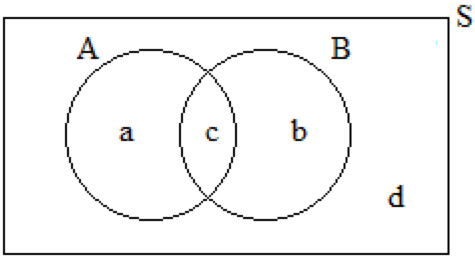
$$P(A \cap B) = P(A) \cdot P(B/A)$$

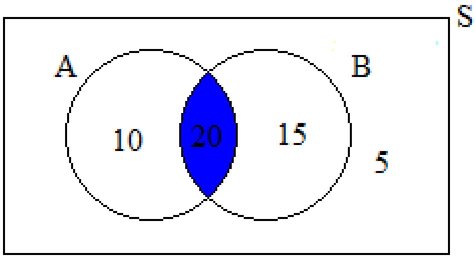
$$P(A \cap B) = \frac{5}{10} \cdot \frac{1}{5} = \frac{1}{10}$$

4) En una clase de 50 alumnos, 10 alumnos tienen como preferencia solamente la asignatura de Matemática, 15 prefieren solamente Estadística y 5 no tienen preferencia por ninguna de estas asignaturas. Calcular la probabilidad que de un alumno de la clase seleccionado al azar tenga preferencia por

4.1) Matemática y Estadística.
4.2) Estadística y Matemática

Solución:
Realizando un diagrama de Venn-Euler se obtiene:





A ∩ B

Simbología:
S = espacio muestral
A= Matemática
B = Estadística

a = Solamente Matemática
b = Solamente Estadística
c = Matemática y Estadística
d = Ninguna de las dos asignaturas

Datos y cálculos:
a = 10
b = 15
c = S – a – b – d = 50 – 10 – 15 – 5 = 20
d = 5
S = 50

4.1) Matemática y Estadística.

$$Número\ de\ resultados\ favorables = A \cap B = n(E) = 20$$

$$Número\ total\ de\ resultados\ posibles = n(S) = 50$$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(A \cap B) = \frac{20}{50} = \frac{2}{5}$$

O también, observando el diagrama de Venn-Euler se tiene directamente la probabilidad solicitada:

$$P(A \cap B) = \frac{20}{50} = \frac{2}{5}$$

Para aplicar la regla de la multiplicación para eventos dependientes, observando el diagrama de Venn-Euler se tiene:

$$P(A) = \frac{30}{50} = \frac{3}{5}$$

La suposición de que el alumno seleccionado tenga preferencia por Matemática significa que sólo consideremos el conjunto A, de los 30 elementos de A, sólo 20 tienen preferencia por Estadística. Por lo tanto la probabilidad condicional P(B/A) = 20/30 = 2/3

O también, observando el diagrama de Venn-Euler y aplicando la fórmula de la probabilidad condicional se tiene:

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{2}{5}}{\frac{3}{5}} = \frac{2}{3}$$

Remplazando valores en la regla de la multiplicación para eventos dependientes se obtiene:

$$P(A \cap B) = P(A) \cdot P(B/A)$$

$$P(A \cap B) = \frac{3}{5} \cdot \frac{2}{3} = \frac{2}{5}$$

4.2) Estadística y Matemática.

$Número\ de\ resultados\ favorables = B \cap A = n(E) = 20$

$Número\ total\ de\ resultados\ posibles = n(S) = 50$

Entonces, aplicando la fórmula de la probabilidad teórica se obtiene:

$$P(E) = \frac{número\ de\ resultados\ favorables}{número\ total\ de\ resultados\ posibles} = \frac{n(E)}{n(S)} \Rightarrow P(B \cap A) = \frac{20}{50} = \frac{2}{5}$$

O también observando el diagrama de Venn-Euler se tiene directamente la probabilidad solicitada:

$$P(B \cap A) = \frac{20}{50} = \frac{2}{5}$$

Para aplicar la regla de la multiplicación para eventos dependientes, observando el diagrama de Venn-Euler se tiene:

$$P(B) = \frac{35}{50} = \frac{7}{10}$$

La suposición de que el alumno seleccionado tenga preferencia por Estadística significa que sólo consideremos el conjunto B, de los 35 elementos de B, sólo 20 tienen preferencia por Matemática. Por lo tanto la probabilidad condicional P(A/B) = 20/35 = 4/7

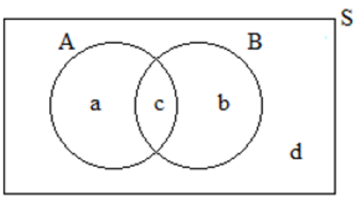
Remplazando valores en la regla general de la multiplicación:

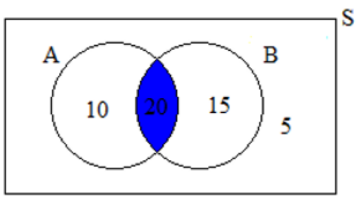
$$P(A \cap B) = P(A) \cdot P(B/A) \Rightarrow P(B \cap A) = P(B) \cdot P(A/B)$$

$$P(B \cap A) = \frac{7}{10} \cdot \frac{4}{7} = \frac{2}{5}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	E	n (E)		P (E)		
2	a	10				
3	b	15				
4	c	20	=B6-B2-B3-B5			
5	d	5				
6	S	50				
7	A	30	=B2+B4	3/5	=B7/B6	
8	B	35	=B3+B4	7/10	=B8/B6	
9						
10	B/A	20	=B4	2/3	=B10/B7	
11	A ∩ B	20	=B4	**2/5**	=B11/B6	
12				**2/5**	=D7*D10	$P(A \cap B) = P(A) \cdot P(B/A)$
13						
14	A/B	20	=B4	4/7	=B14/B8	
15	B ∩ A	20	=B4	**2/5**	=B15/B6	
16				**2/5**	=D8*D14	$P(B \cap A) = P(B) \cdot P(A/B)$





A ∩ B

Notas:

En los eventos dependientes se cumple:

$P(B/A) \neq P(A/B)$

El resultado de $P(A \cap B) = P(B \cap A)$, sin embargo, el proceso de cálculo es diferente.

5) En la tabla de contingencia que aparece a continuación se ha registrado el color de ojos de 40 estudiantes.

Ojos / Género	Café	Verde	Negro	Total
Hombre	12	1	2	15
Mujer	20	3	2	25
Total	32	4	4	40

5.1) Se selecciona un estudiante al azar, calcule la probabilidad de que el estudiante seleccionado:

a) Sea hombre y tenga los ojos negros
b) Tenga los ojos verdes, sabiendo que se trata de una mujer

5.2) Se selecciona dos estudiantes al azar, calcule la probabilidad de que los estudiantes seleccionados:

a) Los dos sean hombres
b) Un hombre y una mujer
c) Los dos tengan los ojos de color café
d) Ojos no verdes el primer estudiante y ojos no negros el segundo estudiante
e) Los dos no tengan los ojos de color café
f) Hombre de ojos café el primer estudiante y mujer de ojos verdes el segundo estudiante

Solución:

Elaborando una tabla de probabilidades se tiene:

Ojos / Género	Café	Verde	Negro	Total
Hombre	12/40 = 3/10	1/40	2/40 = 1/20	15/40 = 3/8
Mujer	20/40 = 1/2	3/40	2/40 = 1/20	25/40 = 5/8
Total	32/40 = 4/5	4/40 = 1/10	4/40 = 1/10	40/40 = 1

5.1) Se selecciona un estudiante al azar, calcule la probabilidad de que el estudiante seleccionado:

a) Sea hombre y tenga los ojos negros
La probabilidad de que un estudiante sea hombre es:

$$P(H) = \frac{15}{40} = \frac{3}{8}$$

De 15 hombres existen 2 estudiantes que tienen los ojos de color negro, entonces, la probabilidad de que tenga los ojos de color negro, siempre que sea hombre es:

$$P(N/H) = \frac{2}{15}$$

O también, observando la tabla de probabilidades y aplicando la fórmula de la probabilidad condicional se tiene:

$$P(N/H) = \frac{P(N \cap H)}{P(H)} = \frac{\frac{1}{20}}{\frac{3}{8}} = \frac{2}{15}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$P(AyB) = P(A) \cdot P(B/A)$

$P(HyN) = P(H) \cdot P(N/H)$

$P(HyN) = \frac{3}{8} \cdot \frac{2}{15} = \frac{1}{20}$

Nota: Si observamos directamente en la tabla de contingencia, se tiene que 2 de los 40 estudiantes son “hombres y tienen los ojos negros”, que en la tabla de probabilidades representa 1/20. Esto confirma que con la regla general para la multiplicación de probabilidades sí se obtiene la respuesta correcta.

b) Tenga los ojos verdes, sabiendo que se trata de una mujer

La probabilidad de que un estudiante tenga los ojos verdes es:

$P(V) = \frac{4}{40} = \frac{1}{10}$

Existen 3 mujeres de 4 estudiantes que tienen los ojos de color verde, entonces, la probabilidad de que sea mujer, siempre que tenga los ojos de color verde es:

$P(M/V) = \frac{3}{4}$

Remplazando valores en la regla general de la multiplicación se obtiene:

$P(AyB) = P(A) \cdot P(B/A)$

$P(VyM) = P(V) \cdot P(M/V)$

$P(VyM) = \frac{1}{10} \cdot \frac{3}{4} = \frac{3}{40}$

O también, observando directamente en la tabla de probabilidades se obtiene:

$P(VyM) = \frac{3}{40}$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G
1		Café	Verde	Negro	Total		
2	Hombre	12	1	2	15		
3	Mujer	20	3	2	25		
4	Total	32	4	4	40		
5							
6		Café	Verde	Negro	Total		
7	Hombre	3/10	1/40	1/20	3/8		
8	Mujer	1/2	3/40	1/20	5/8		
9	Total	4/5	1/10	1/10	1		
10							
11	P(H)	3/8	=E7				
12	P(N/H)	2/15	=D2/E2				
13	P(HyN)	1/20	=B11*B12		$P(HyN) = P(H) \cdot P(N/H)$		
14		1/20	=D7				
15							
16	P(V)	1/10	=C9				
17	P (M/V)	3/4	=C3/C4				
18	P(VyM)	3/40	=B16*B17		$P(VyM) = P(V) \cdot P(M/V)$		
19		3/40	=C8				

5.2) Se selecciona dos estudiantes al azar, calcule la probabilidad de que los estudiantes seleccionados:

Ojos / Género	Café	Verde	Negro	Total
Hombre	12	1	2	15
Mujer	20	3	2	25
Total	32	4	4	40

a) Los dos sean hombres

La probabilidad de que el primer estudiante seleccionado sea hombre es:

$$P(H) = \frac{15}{40} = \frac{3}{8}$$

La probabilidad de que el segundo estudiante seleccionado sea hombre, dado que ya se seleccionó un hombre en la primera extracción es:

$$P(H/H) = \frac{14}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$

$$P(HyH) = P(H) \cdot P(H/H)$$

$$P(HyH) = \frac{3}{8} \cdot \frac{14}{39} = \frac{7}{52}$$

b) Un hombre y una mujer

La probabilidad de que el primer estudiante seleccionado sea hombre es:

$$P(H) = \frac{15}{40} = \frac{3}{8}$$

La probabilidad de que el segundo estudiante seleccionado sea mujer, dado que ya se seleccionó un estudiante hombre en la primera extracción es:

$$P(M/H) = \frac{25}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$

$$P(HyM) = P(H) \cdot P(M/H)$$

$$P(HyM) = \frac{3}{8} \cdot \frac{25}{39} = \frac{25}{104}$$

c) Los dos tengan los ojos de color café

La probabilidad de que el primer estudiante seleccionado tenga los ojos de color café es:

$$P(C) = \frac{32}{40} = \frac{4}{5}$$

La probabilidad de que el segundo estudiante seleccionado tenga los ojos de color café, dado que ya se seleccionó un estudiante que tiene los ojos de color café en la primera extracción es:

$$P(C/C) = \frac{31}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$

$$P(CyC) = P(C) \cdot P(C/C)$$

$$P(CyC) = \frac{4}{5} \cdot \frac{31}{39} = \frac{124}{195}$$

d) Ojos no verdes el primer estudiante y ojos no negros el segundo estudiante

La probabilidad de que el primer estudiante seleccionado no tenga los ojos de color verde es:

$$P(\bar{V}) = 1 - P(V) = 1 - \frac{4}{40} = \frac{9}{10}$$

O también como el número de estudiantes que no tienen los ojos de color verde son 32 + 4 = 36, entonces,

$$P(\bar{V}) = \frac{36}{40} = \frac{9}{10}$$

La probabilidad de que el segundo estudiante seleccionado no tenga los ojos de color negro, dado que ya se seleccionó un estudiante que no tiene los ojos de color verde en la primera extracción es:

$$P(\bar{N}/\bar{V}) = 1 - P(N/\bar{V}) = 1 - \frac{4}{39} = \frac{35}{39}$$

O también observando la tabla de contingencia se obtiene:

$$P(\bar{N}/\bar{V}) = \frac{36-1}{40-1} = \frac{35}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$

$$P(\bar{V}y\bar{N}) = P(\bar{V}) \cdot P(\bar{N}/\bar{V})$$

$$P(\bar{V}y\bar{N}) = \frac{9}{10} \cdot \frac{35}{39} = \frac{21}{26}$$

e) Los dos no tengan los ojos de color café

La probabilidad de que el primer estudiante seleccionado no tenga los ojos de color café es:

$$P(\bar{C}) = 1 - P(C) = 1 - \frac{32}{40} = \frac{1}{5}$$

O también como el número de estudiantes que no tienen los ojos de color café son 4 + 4 = 8, entonces,

$$P(\bar{C}) = \frac{8}{40} = \frac{1}{5}$$

La probabilidad de que el segundo estudiante seleccionado no tenga los ojos de color café, dado que ya se seleccionó un estudiante que no tiene los ojos de color café en la primera extracción es:

$$P(\bar{C}/\bar{C}) = 1 - P(C/\bar{C}) = 1 - \frac{32}{39} = \frac{7}{39}$$

O también observando la tabla de contingencia se obtiene:

$$P(\bar{C}/\bar{C}) = \frac{8-1}{40-1} = \frac{7}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyB) = P(A) \cdot P(B/A)$$

$$P(\bar{C}y\bar{C}) = P(\bar{C}) \cdot P(\bar{C}/\bar{C})$$

$$P(\bar{C}y\bar{C}) = \frac{1}{5} \cdot \frac{7}{39} = \frac{7}{195}$$

f) Hombre de ojos café el primer estudiante y mujer de ojos verdes el segundo estudiante

Si observamos directamente en la tabla de contingencia, se tiene que 12 de los 40 estudiantes son "hombres y tienen los ojos de color café", entonces, la probabilidad de que el primer estudiante seleccionado sea hombre y tenga los ojos de color café es:

$$P(HyC) = \frac{12}{40} = \frac{3}{10}$$

Si observamos directamente en la tabla de contingencia, se tiene que 12 de los 40 estudiantes son "hombres y tienen los ojos de color café", entonces, la probabilidad de que el primer estudiante seleccionado sea hombre y tenga los ojos de color café es:

Si observamos directamente en la tabla de contingencia, se tiene que 3 de los 40 estudiantes son "mujeres y tienen los ojos de color verde", entonces, la probabilidad de que el segundo estudiante seleccionado sea mujer y tenga los ojos de color verde, dado que ya se seleccionó un estudiante hombre que tiene los ojos de color café en la primera extracción es:

$$P((MyV)/(HyC)) = \frac{3}{39}$$

Remplazando valores en la regla general de la multiplicación se obtiene:

$P(AyB) = P(A) \cdot P(B/A)$

$P((HyC)y(MyV)) = P(HyC) \cdot P((MyV)/(HyC))$

$$P\big((HyC)y(MyV)\big) = \frac{3}{10} \cdot \frac{3}{39} = \frac{3}{130}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I	J
1		Café	Verde	Negro	Total					
2	Hombre	12	1	2	15					
3	Mujer	20	3	2	25					
4	Total	32	4	4	40					
5										
6	P(HyH)	7/52	=E2/E4*(E2-1)/(E4-1)				$P(HyH) = P(H) \cdot P(H/H)$			
7										
8	P(HyM)	25/104	=E2/E4*E3/(E4-1)				$P(HyM) = P(H) \cdot P(M/H)$			
9										
10	P(CyC)	124/195	=B4/E4*(B4-1)/(E4-1)				$P(CyC) = P(C) \cdot P(C/C)$			
11										
12	$P(\bar{V}y\bar{N})$	21/26	=(B4+D4)/E4*(B4+C4-1)/(E4-1)				$P(\bar{V}y\bar{N}) = P(\bar{V}) \cdot P(\bar{N}/\bar{V})$			
13										
14	$P(\bar{C}y\bar{C})$	7/195	=(C4+D4)/E4*(C4+D4-1)/(E4-1)				$P(\bar{C}y\bar{C}) = P(\bar{C}) \cdot P(\bar{C}/\bar{C})$			
15										
16	$P\big((HyC)y(MyV)\big)$	3/130	=B2/E4*C3/(E4-1)				$P((HyC)y(MyV)) = P(HyC) \cdot P((MyV)/(HyC))$			

6) De una tómbola que contiene 3 bolas rojas y 5 blancas, Mathías extrae tres bolas, sin volver a la tómbola la bola extraída, calcular la probabilidad de que las 3 bolas extraídas sean:

6.1) Rojas
6.2) 2 rojas y una blanca
6.3) Una roja y 2 blancas
6.4) 3 blancas

Solución:

6.1) Rojas

En 3 sucesos la fórmula de la regla general de probabilidades es:

$P(AyByC) = P(A) \cdot P(B/A) \cdot P(C/AyB)$

La probabilidad de seleccionar una bola roja la primera extracción es:

$$P(R) = \frac{3}{8}$$

La probabilidad de seleccionar una bola roja la segunda extracción, dado que ya se seleccionó una bola roja en la primera extracción es:

$$P(R/R) = \frac{2}{7}$$

La probabilidad de seleccionar una bola roja la tercera extracción, dado que ya se seleccionó una bola roja en la primera y en la segunda extracción es:

$$P(R/RyR) = \frac{1}{6}$$

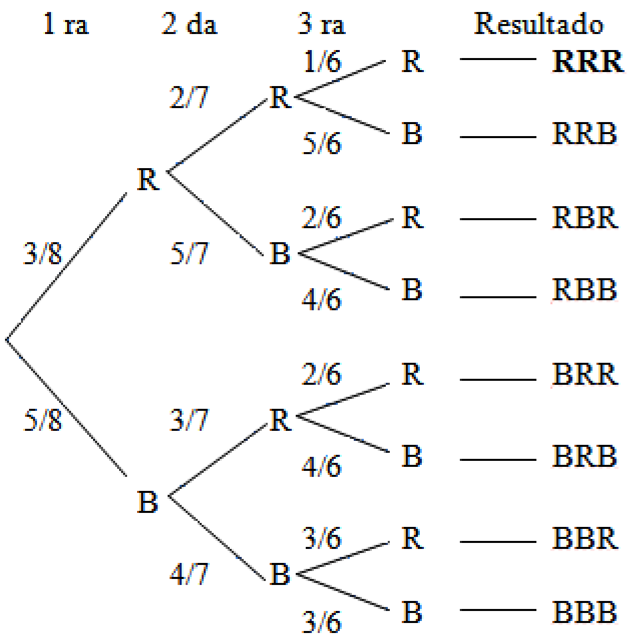
Remplazando valores en la regla general de la multiplicación se obtiene:

$$P(AyByC) = P(A) \cdot P(B/A) \cdot P(C/AyB)$$

$$P(RyRyR) = P(R) \cdot P(R/R) \cdot P(R/RyR)$$

$$P(RyRyR) = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} = \frac{1}{56}$$

O también, elaborando un diagrama de árbol se tiene todas las probabilidades:



En el diagrama de árbol, la probabilidad correspondiente a cada rama del árbol corresponde a la probabilidad condicional de que ocurra el evento específico, dado que han ocurrido los eventos de las ramas precedentes. Al describir un evento mediante una trayectoria a través del diagrama de árbol, la probabilidad de que ocurra dicho evento es igual a producto de las probabilidades de las ramas que forman la trayectoria que representa al mencionado evento.

La solución empleando el diagrama de árbol para $P(RyRyR)$ es multiplicando las ramas RRR, es decir,

$$P(RRR) = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{1}{6} = \frac{1}{56}$$

6.2) 2 rojas y una blanca

La solución empleando el diagrama de árbol se obtiene multiplicando las ramas RRB

$$P(RRB) = \frac{3}{8} \cdot \frac{2}{7} \cdot \frac{5}{6} = \frac{5}{56}$$

6.3) Una roja y 2 blancas

La solución empleando el diagrama de árbol se obtiene multiplicando las ramas RBB

$$P(RBB) = \frac{3}{8} \cdot \frac{5}{7} \cdot \frac{4}{6} = \frac{5}{28}$$

6.4) 3 blancas

La solución empleando el diagrama de árbol se obtiene multiplicando las ramas BBB

$P(BBB) = \frac{5}{8} \cdot \frac{4}{7} \cdot \frac{3}{6} = \frac{5}{28}$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	R	3			
2	B	5			
3	Total	8			
4					
5	P(RRR)	1/56	=B1/B3*(B1-1)/(B3-1)*(B1-2)/(B3-2)		
6	P(RRB)	5/56	=B1/B3*(B1-1)/(B3-1)*B2/(B3-2)		
7	P(RBB)	5/28	=B1/B3*B2/(B3-1)*(B2-1)/(B3-2)		
8	P(BBB)	5/28	=B2/B3*(B2-1)/(B3-1)*(B2-2)/(B3-2)		

ii) REGLA PARTICULAR O ESPECIAL PARA EVENTOS INDEPENDIENTES

Si A y B son dos eventos independientes, es decir, si el conocimiento de la incidencia de uno de ellos no tiene efecto en la probabilidad de ocurrencia del otro, entonces, para calcular la probabilidad de dichos eventos se aplica la siguiente regla:

$$P(AyB) = P(A) \cdot P(B)$$
$$o$$
$$P(A \cap B) = P(A) \cdot P(B)$$

Nota: Dos eventos A y B son independientes si la ocurrencia de uno de ellos no afecta la probabilidad de ocurrencia del otro, esto es, si $P(B/A) = P(B)$

Ejemplos ilustrativos

1) De una baraja estándar de 52 cartas sea A el suceso de sacar un As en la primera extracción y B sacar un Rey en la segunda extracción. Calcular la probabilidad de sacar un As y un Rey en dos extracciones devolviendo la carta extraída.

Solución:

A y B son sucesos independientes porque la ocurrencia de A afecta la probabilidad de ocurrencia de B.

La probabilidad de que la primera carta sea un As es:

$P(A) = \frac{4}{52}$

La probabilidad de que la segunda carta sea un Rey, dado que se devolvió el As de la primera extracción, es:

$P(B/A) = P(B) = \frac{4}{52}$

Remplazando los anteriores valores en la regla particular de la multiplicación se obtiene:

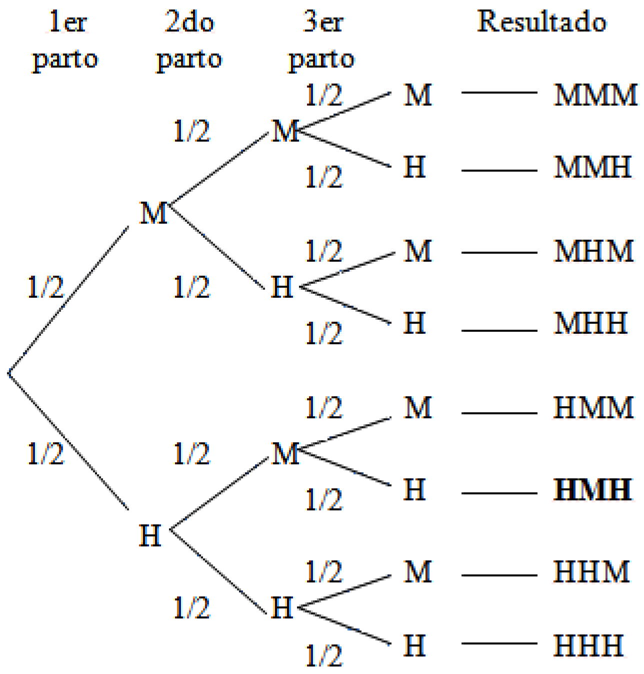
$$P(AyB) = P(A) \cdot P(B)$$
$$P(AyB) = \frac{4}{52} \cdot \frac{4}{52} = \frac{1}{169}$$

2) Una pareja de esposos desean tener 3 hijos. Suponiendo que las probabilidades de tener un niño o una niña son iguales, calcular la probabilidad de éxito en tener hombre en el primer nacimiento, mujer en el segundo nacimiento y hombre en el tercer nacimiento.

Solución:

M = mujer
H = hombre

Elaborando un diagrama de árbol se tiene todas las probabilidades:



Entonces,

$$P(HyMyH) = P(HMH) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

TAREA DE INTERAPRENDIZAJE N° 4

1) De una baraja estándar de 52 cartas sea A el suceso de sacar un Rey en la primera extracción y B sacar un Rey en la segunda extracción. Calcular la probabilidad de sacar Rey y Rey, sin devolver la carta extraída.

1/221

2) Se tienen tres ruletas con 9 símbolos, se gana el premio mayor cuando se obtienen tres 3. ¿Cuál es la probabilidad de ganar el premio mayor?

1/729

3) De una baraja estándar de 52 cartas sea A el suceso de sacar un Rey en la primera extracción y B sacar una Reina en la segunda extracción. Calcular la probabilidad de sacar Rey y Reina, devolviendo la carta extraída.

1/169

4) En un sobre hay tres cartulinas: una blanca, una amarilla y una roja. ¿Qué probabilidad existe de obtener tres veces consecutivas la cartulina roja si se vuelve a guardar lo que se sacó anteriormente?

1/27

5) En un cesto hay 5 manzanas y 4 peras; en un segundo cesto, 3 manzanas y 2 naranjas. Si los cestos se encuentran totalmente cubiertos, ¿qué probabilidad existe de que al sacar una fruta de cada cesto se obtengan manzanas?

1/3

6) Se colocan 9 frascos de esencias en una caja, 5 son de vainilla y 4 de menta. Calcular la probabilidad de sacar dos frascos de vainilla en dos extracciones consecutivas, si se utiliza el extracto y se vuelve a guardar en la caja

25/81

7) Se tienen 15 fichas en una caja: 5 verdes, 5 amarillas y 5 azules. ¿Qué probabilidad existe de obtener en tres ocasiones consecutivas una ficha verde si se vuelve a guardar la que se sacó anteriormente?.

1/27

8) Dyanita y Mario, una pareja de esposos, desean tener 2 hijos. Suponiendo que las probabilidades de tener un niño o una niña son iguales, calcular la probabilidad de éxito en tener hombre en el primer nacimiento y mujer en el segundo nacimiento. Elabore un diagrama de árbol.

1/4

9) En una urna existe 10 bolas numeradas del 1 al 10. ¿Qué probabilidad existe de sacar en una sola extracción una bola enumerada con un número impar y múltiplo de 3?

1/5

10) En una bandeja se tienen 15 sánduches, 7 de pernil y 8 de jamón. Si se brinda a dos amigos, calcule la probabilidad de manera manual y empleando Excel de que los dos amigos escojan sánduches de jamón cada uno. También elabore un diagrama de árbol.

4/15

11) En una caja hay 18 chocolates, 10 con relleno y 8 sin relleno. Si se realiza tres extracciones sin volver a guardar en la caja el chocolate extraído, calcule la probabilidad de manera manual y empleando Excel de que se obtengan tres chocolates rellenos. También elabore un diagrama de árbol.

5/34

12) En una urna existe 10 bolas numeradas del 1 al 10. Elabore un diagrama de Venn-Euler y calcule la probabilidad de obtener en una sola extracción una bola enumerada con un número impar y múltiplo de 5

1/10

13) Un dado tiene tres caras blancas numeradas 4, 5 y 6, y tres caras rojas numeradas 1, 2 y 3. Si se lanza una vez este dado. Elabore un diagrama de Venn-Euler y calcule la probabilidad de obtener una cara blanca numerada con un número par.

1/3

14) En una clase de 40 alumnos, 10 alumnos tienen como preferencia solamente la asignatura de Matemática, 15 prefieren solamente Estadística y 5 no tienen preferencia por ninguna de estas asignaturas. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel de que un alumno seleccionado al azar tenga preferencia por Matemática y Estadística

1/4

15) En un grupo de 50 personas, 6 tienen como preferencia solamente el color amarrillo, 10 prefieren solamente el color blanco, 4 prefieren solamente el color café, 6 prefieren el color amarrillo y blanco, 10 prefieren el color blanco y café, 12 prefieren el color amarrillo y café, y 10 no tienen preferencia por ninguno de los tres colores. Elabore un diagrama de Venn-Euler. Aplique la fórmula de la probabilidad de 3 sucesos para calcular la probabilidad de que un alumno seleccionado al azar tenga preferencia por los 3 colores.

2/25

16) En un grupo de 60 fábricas, 4 confeccionan solamente ropa deportiva, 10 solamente ropa casual, 9 solamente ropa formal, 11 ropa casual y formal, 12 ropa deportiva y casual, 15 ropa deportiva y formal, y 11 confeccionan otros artículos. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel que de una fábrica seleccionada al azar confeccione los tres tipos de ropas.

7/60

17) En una clase, la probabilidad de que los alumnos tengan como preferencia la asignatura de Matemática es 3/5, la probabilidad de que los alumnos tengan como preferencia la asignatura de Estadística es 7/10 y la probabilidad de que los estudiantes no prefieran ninguna de estas asignaturas es 1/10. Elabore un diagrama de Venn-Euler y calcule la probabilidad de manera manual y empleando Excel de que un alumno seleccionado al azar

17.1) Tenga preferencia por Matemática y Estadística

2/5

17.2) Sabiendo que un alumno seleccionado al azar tiene preferencia por Matemática, calcule la probabilidad de que tenga preferencia por Estadística.

P(E/M) =2/3

18) En la tabla de contingencia que aparece a continuación se ha registrado el deporte y el género de 45 estudiantes que lo practican.

Deporte / Género	Fútbol	Atletismo	Básquet	Total
Hombre	15	3	2	20
Mujer	2	3	20	25
Total	17	6	22	45

18.1) Se selecciona un estudiante al azar, calcule la probabilidad de manera manual y empleando Excel de que el estudiante seleccionado:

a) Sea hombre y practique básquet

2/45

b) Practique atletismo, sabiendo que se trata de una mujer

1/15

18.2) Se selecciona dos estudiantes al azar, calcule la probabilidad de manera manual y empleando Excel de que los estudiantes seleccionados:

a) Los dos sean hombres

19/99

b) Un hombre y una mujer

25/99

c) Los dos practique fútbol

68/495

d) Que no practique atletismo el primer estudiante y que no practique básquet el segundo estudiante

13/30

19) Plantee y resuelva un problema similar al anterior de manera manual y empleando Excel.

20) Un recipiente contiene 3 bolas negras, 2 rojas y 5 blancas. ¿Cuál es la probabilidad de que, al hacer dos extracciones sucesivas, se obtenga una bola roja en la primera extracción y en la segunda extracción una bola negra, sin devolver a al recipiente la primera bola extraída?. Elabore un diagrama de árbol

1/15

21) De una urna que contiene 5 fichas de color rojo y 3 blancas, se extrae tres bolas, sin devolver a la urna la ficha extraída. Elabore un diagrama de árbol y calcule la probabilidad de manera manual y empleando Excel de que las 3 fichas extraídas sean 2 rojas y una blanca.

5/28

22) Dyanita va a una papelería a comprar 3 marcadores. En la papelería únicamente disponen de 6 marcadores de tinta rojos, 5 de tinta negra y 9 de tinta azul. Elabore un diagrama de árbol y calcule la probabilidad de manera manual y empleando Excel de que si Dyanita compra al lazar, compre un marcador con tinta azul y dos marcadores con tinta negra.

1/38

23) Plantee y resuelva un problema similar al anterior

24) Para integrar una comisión, se debe elegir a tres alumnos de entre 28 mujeres y 12 hombres. Se preparan papeles con los nombres de los integrantes y se introducen en una ánfora para elegir al azar. Elabore un diagrama de árbol y calcule la probabilidad de manera manual y empleando Excel de que la comisión esté formada por dos mujeres y un hombre.

15,304%

25) De una baraja estándar de 52 cartas, se extraen 3 cartas sin reposición. Elabore un diagrama de árbol y calcule la probabilidad de que las tres cartas sean Ases.

1/5525

1.4) PROBABILIDAD TOTAL Y TEOREMA DE BAYES

A) PROBABILIDAD TOTAL

Sea $A_1, A_2, \ldots\ldots A_n$ un sistema completo de eventos tales que la probabilidad de cada uno de ellos es distinto de cero, y sea B un evento cualquiera del que se conocen las probabilidades condicionales $P(B/A_i)$, entonces, la probabilidad del evento B, llamada probabilidad total, se calcula empleando la siguiente fórmula:

$$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \cdots P(A_n) \cdot P(B/A_n)$$

B) TEOREMA DE BAYES

El teorema de Bayes se utiliza para revisar probabilidades previamente calculadas cuando se posee nueva información. Desarrollado por el reverendo Thomas Bayes en el siglo XVII, el teorema de Bayes es una extensión de lo que ha aprendido hasta ahora acerca de la probabilidad condicional.

Comúnmente se inicia un análisis de probabilidades con una asignación inicial, probabilidad a priori. Cuando se tiene alguna información adicional se procede a calcular las probabilidades revisadas o a posteriori. El teorema de Bayes permite calcular las probabilidades a posteriori y es:

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{P(B)}$$

Donde:
$P(A_i)$ = Probabilidad a priori
$P(B/A_i)$ = Probabilidad condicional
$P(B)$ = Probabilidad Total
$P(A_i/B)$ = Probabilidad aposteriori

Ejemplo ilustrativo

Una compañía de transporte público tiene tres líneas en una ciudad, de forma que el 45% de los autobuses cubre el servicio de la línea 1, el 25% cubre la línea 2 y el 30% cubre el servicio de la línea 3. Se sabe que la probabilidad de que, diariamente, un autobús se averíe es del 2%, 3% y 1% respectivamente, para cada línea.

1) Calcular la probabilidad de que, en un día, un autobús sufra una avería
2) Calcular la probabilidad de que, en un día, un autobús no sufra una avería
3) ¿De qué línea de transporte es más probable que un autobús sufra una avería?

Solución:

Simbología:
A_1 = Cubre el servicio de la línea 1
A_2 = Cubre el servicio de la línea 2
A_3= Cubre el servicio de la línea 3
B_1= Sufre una avería
B_2= No sufre una avería

Datos:
$P(A_1) = 45\% = 0{,}45$; $P(A_2) = 25\% = 0{,}25$; $P(A_3) = 30\% = 0{,}3$
$P(B_1/A_1) = 2\% = 0{,}02$; $P(B_1/A_2) = 3\% = 0{,}03$; $P(B_1/A_3) = 1\% = 0{,}01$

Las probabilidades de no sufrir una avería para cada línea son:
$P(B_2/A_1) = 1 - P(B_1/A_1) = 1 - 0{,}02 = 0{,}98$
$P(B_2/A_2) = 1 - P(B_1/A_2) = 1 - 0{,}03 = 0{,}97$
$P(B_2/A_3) = 1 - P(B_1/A_3) = 1 - 0{,}01 = 0{,}99$

1) Calcular la probabilidad de que, en un día, un autobús sufra una avería

Empleando la fórmula de probabilidad total se obtiene:

$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \cdots P(A_n) \cdot P(B/A_n)$
$P(B_1) = P(A_1) \cdot P(B_1/A_1) + P(A_2) \cdot P(B_1/A_2) + P(A_3) \cdot P(B_1/A_3)$
$P(B_1) = 0{,}45 \cdot 0{,}02 + 0{,}25 \cdot 0{,}03 + 0{,}3 \cdot 0{,}01 = 0{,}0195$

2) Calcular la probabilidad de que, en un día, un autobús no sufra una avería

Empleando la fórmula de probabilidad total se obtiene:

$P(B) = P(A_1) \cdot P(B/A_1) + P(A_2) \cdot P(B/A_2) + \cdots P(A_n) \cdot P(B/A_n)$
$P(B_2) = P(A_1) \cdot P(B_2/A_1) + P(A_2) \cdot P(B_2/A_2) + P(A_3) \cdot P(B_2/A_3)$
$P(B_2) = 0{,}45 \cdot 0{,}98 + 0{,}25 \cdot 0{,}97 + 0{,}3 \cdot 0{,}99 = 0{,}9805$

O también, sabiendo que $P(\bar{E}) = 1 - P(E)$, entonces
$P(B_2) = 1 - P(B_1) = 1 - 0{,}0195 = 0{,}9805$

3) ¿De qué línea de transporte es más probable que un autobús sufra una avería?

Se debe calcular las tres probabilidades aposteriori empleando el Teorema de Bayes

$$P(A_1/B) = \frac{P(A_1) \cdot P(B/A_1)}{P(B)}$$

La probabilidad de que sea de la línea 1, sabiendo que sufre una avería es:

$$P(A_1/B_1) = \frac{P(A_1) \cdot P(B/A_1)}{P(B_1)} = \frac{0{,}45 \cdot 0{,}02}{0{,}0195} = 0{,}4615$$

La probabilidad de que sea de la línea 2, sabiendo que sufre una avería es:

$$P(A_2/B_1) = \frac{P(A_2) \cdot P(B/A_2)}{P(B_1)} = \frac{0{,}25 \cdot 0{,}03}{0{,}0195} = 0{,}3846$$

La probabilidad de que sea de la línea 3, sabiendo que sufre una avería es:

$$P(A_3/B_1) = \frac{P(A_3) \cdot P(B_1/A_3)}{B_1} = \frac{0{,}3 \cdot 0{,}01}{0{,}0195} = 0{,}1538$$

Entonces, sabiendo que el autobús sufre una avería, lo más probable es que sea de la línea 1, ya que esta probabilidad $P(A_1/B_1) = 0{,}4615$, es la mayor.

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I	J
1	$P(A_1)$	0,45								
2	$P(A_2)$	0,25								
3	$P(A_3)$	0,3								
4	$P(B_1/A_1)$	0,02								
5	$P(B_1/A_2)$	0,03								
6	$P(B_1/A_3)$	0,01								
7	$P(B_2/A_1)$	0,98	=1-B4		$P(B_2/A_1) = 1 - P(B_1/A_1)$					
8	$P(B_2/A_2)$	0,97	=1-B5		$P(B_2/A_2) = 1 - P(B_1/A_2)$					
9	$P(B_2/A_3)$	0,99	=1-B6		$P(B_2/A_3) = 1 - P(B_1/A_3)$					
10	$P(B_1)$	0,0195	=B1*B4+B2*B5+B3*B6		$P(B_1) = P(A_1)\cdot P(B_1/A_1) + P(A_2)\cdot P(B_1/A_2) + P(A_3)\cdot P(B_1/A_3)$					
11	$P(B_2)$	0,9805	=1-B10		$P(B_2) = 1 - P(B_1)$					
12	$P(A_1/B_1)$	0,4615	=B1*B4/B10		$P(A_i/B_1) = \frac{P(A_i)\cdot P(B/A_1)}{P(B_1)}$					
13	$P(A_2/B_1)$	0,3846	=B2*B5/B10							
14	$P(A3/B_1)$	0,1538	=B3*B6/B10							
15	$P\ Max(A/B_1)$	0,4615	=MAX(B12:B14)							

TAREA DE INTERAPRENDIZAJE N° 5

1) Consulte en la biblioteca o en el internet sobre la biografía de Bayes, y realice un organizador gráfico de la misma.

2) Conteste las siguientes preguntas:
2.1) ¿Qué entiende por probabilidad total?
2.2) ¿Qué entiende por probabilidad a priori?
2.3) ¿Qué entiende por probabilidad a posteriori?
2.4) ¿Qué entiende por teorema de Bayes?

3) Una empresa dedicada a la comercialización de televisores está considerando comercializar un nuevo televisor. En el pasado el 90% de los televisores que comercializó tuvieron éxito y el 10% no fueron exitosos. Se sabe que la probabilidad que habría recibido un reporte favorable de investigación fue del 85% y 35%, respectivamente.

3.1) Escribir la simbología del problema
$A_1 =$
$A_2 =$
$B_1 =$
$B_2 =$

3.2) Llenar los datos del problema
$P(A_1)$ $\quad$ $P(A_2) =$ $\quad$ $P(B_1/A_1) =$ $\quad$ $P(B_1/A_2) =$

3.3) Calcule la probabilidad que los televisores exitosos reciban un reporte desfavorable de investigación

$$P(B_2/A_1) = 0{,}15$$

3.4) Calcule la probabilidad que los televisores no exitosos reciban un reporte desfavorable de investigación

$$P(B_2/A_2) = 0{,}65$$

3.5) Calcule la probabilidad de que un televisor reciba un reporte favorable de investigación.

$P(B_1) = 0{,}8$

3.6) Calcule la probabilidad de que un televisor reciba un reporte desfavorable de investigación.

$P(B_2) = 0{,}2$

3.7) ¿Cuál es la probabilidad de que el equipo de televisor tenga éxito en el mercado?

$P(A_1/B_1) = 0{,}9563$

3.8) Resuelva los ítems 3.3) al 3.7) empleando Excel

4) La probabilidad de que una persona tenga una determinada enfermedad es de 0,02. Existen pruebas de diagnóstico médico disponibles para determinar si una persona tiene realmente la enfermedad. Si la enfermedad realmente está presente, la probabilidad de que la prueba de diagnóstico indique la presencia de la enfermedad es de 0,95.

4.1) ¿Cuál es la probabilidad de tener la enfermedad, si la prueba de diagnóstico indica la presencia de la misma?. Calcule de manera manual y empleando Excel

$P(A_1/B_1) = 0{,}2794$

4.2) ¿Cuál es la probabilidad de no tener la enfermedad, si la prueba de diagnóstico no indica la presencia de la misma?. Calcule de manera manual y empleando Excel

$P(A_2/B_2) = 0{,}9989$

5) Una fábrica de sacos tiene 3 máquinas independientes que producen el mismo tipo de sacos. La máquina 1 produce el 15% de los sacos con un 1% de sacos defectuosos. La máquina 2 produce el 45% de los sacos con un 3% de sacos defectuosos. La máquina 3 produce el 40% de los sacos con un 2% de sacos defectuosos.

5.1) Si se selecciona al azar un saco. ¿Cuál es la probabilidad de que sea defectuoso?

0,023

5.2) ¿De qué máquina es más probable que provenga, si el saco no resulta defectuoso?

Máquina 2 con 0,4468

6) El primer año de bachillerato de un colegio está integrado por 35 estudiantes en la especialidad de físico matemático, 47 en químico biólogo, 40 en sociales y 38 en bachillerato general. Se sabe que la probabilidad de que un estudiante pierda el año es del 5% 4%, 3% y 4%, respectivamente. ¿De qué especialidad es más probable que sea el estudiante, si se sabe que un estudiante ha perdido el año?

Químico biólogo con 0,296

7) Plantee y resuelva un problema de aplicación del teorema de Bayes en forma manual y empleando Excel.

CAPÍTULO II
DISTRIBUCIONES DE PROBABILIDAD

RESULTADOS DE APRENDIZAJE DEL CAPÍTULO

Al finalizar el presente capítulo el lector podrá evidenciar que:

- ✓ Interpreta las definiciones, características, propiedades y aplicaciones de las distribuciones de probabilidad: binomial, de Poisson, hipergeométrica, exponencial, uniforme y normal.
- ✓ Aplica algoritmos relativos a las distribuciones de probabilidad en la resolución de ejercicios y problemas prácticos de manera manual, empleando Excel, Winstats y GeoGebra.
- ✓ Plantea y resuelve ejercicios y problemas de aplicación sobre distribuciones de probabilidad de manera manual, utilizando Excel, Winstats y GeoGebra.

CONTENIDOS

- ✓ Distribuciones Discretas: Distribución Binomial, Distribución de Poisson y Distribución Hipergeométrica.
- ✓ Distribuciones Continuas: Distribución Exponencial, Distribución Uniforme y Distribución Normal

2.1) DISTRIBUCIONES DISCRETAS

A) INTRODUCCIÓN

Una distribución de probabilidad es una representación de todos los resultados posibles de algún experimento y de la probabilidad relacionada con cada uno.
Una distribución de probabilidad es discreta cuando los resultados posibles del experimento son obtenidos de variables aleatorias discretas, es decir, de variables que sólo puede tomar ciertos valores, con frecuencia números enteros, y que resultan principalmente del proceso de conteo.

Ejemplos de variables aleatorias discretas son:
Número de caras al lanzar una moneda
El resultado del lanzamiento de un dado
Número de hijos de una familia
Número de estudiantes de una universidad

Ejemplo ilustrativo
Sea el experimento aleatorio de lanzar 2 monedas al aire. Determinar la distribución de probabilidades del número de caras.

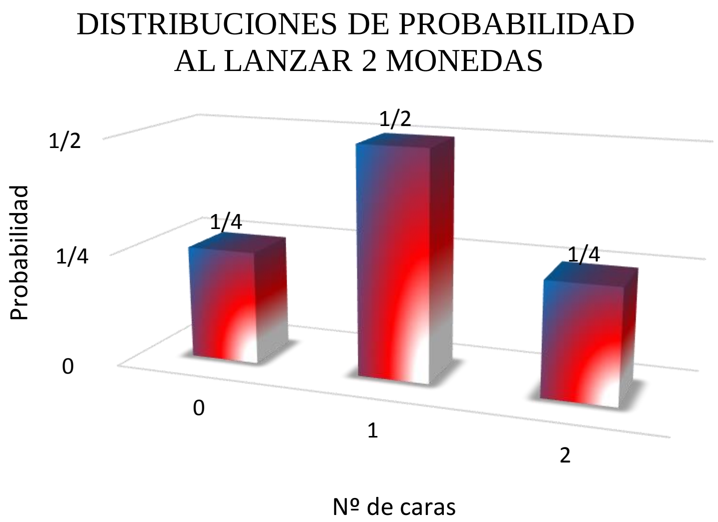
Solución:

El espacio muestral es S = {CC, CS, SC, SS}
La probabilidad de cada punto muestral es de 1/4, es decir, P(CC) = P(CS) = P(SC) = P(SS) = 1/4

La distribución de probabilidades del número de caras se presenta en la siguiente tabla:

Resultados (N° de Caras)	Probabilidad
0	1/4 = 0,25 = 25%
1	2/4 = 0,50 = 50%
2	1/4 = 0,25 = 25%

El gráfico de distribuciones de probabilidad en 3D elaborado en Excel se muestra en la siguiente figura:



Interpretación:
La probabilidad de obtener 0 caras al lanzar 2 monedas al aire es de 1/4 = 0,25 = 25%
La probabilidad de obtener una cara al lanzar 2 monedas al aire es de 2/4 = 0,5 = 50%
La probabilidad de obtener 2 caras al lanzar 2 monedas al aire es de 1/4 = 0,25 = 25%

B) LA MEDIA Y LA VARIANZA DE LAS DISTRIBUCIONES DISCRETAS

i) Media

La media llamada también valor esperado, esperanza matemática o simplemente esperanza de una distribución de probabilidad discreta es la media aritmética ponderada de todos los resultados posibles en los cuales los pesos son las probabilidades respectivas de tales resultados. Se halla multiplicando cada resultado posible por su probabilidad y sumando los resultados. Se expresa mediante la siguiente fórmula:

$$\mu = \mathrm{E(X)} = \Sigma(x_i \cdot P(x_i))$$

Donde:
$\mu = \mathrm{E(X)}$ = Media, Valor Esperado, Esperanza Matemática o simplemente Esperanza
x_i = Posible resultado
$P(x_i)$ = Probabilidad del posible resultado

ii) Varianza

La varianza es el promedio de las desviaciones al cuadrado con respecto a la media. La varianza mide la dispersión de los resultados alrededor de la media y se halla calculando las diferencias entre cada uno de los resultados y su media, luego tales diferencias se elevan al cuadrado y se multiplican por sus respectivas probabilidades, y finalmente se suman los resultados. Se expresa mediante la siguiente fórmula:

$$\sigma^2 = \Sigma[(x_i - \mu)^2 \cdot P(x_i)]$$

Nota: La varianza se expresa en unidades al cuadrado, por lo que es necesario calcular la desviación estándar que se expresa en las mismas unidades que la variable aleatoria y que por lo tanto tiene una interpretación más lógica de la dispersión de los resultados alrededor de la media. La desviación estándar se calcula así: $\sigma = \sqrt{\sigma^2}$

Ejemplo ilustrativo:
Hallar la esperanza matemática, la varianza y la desviación estándar del número de caras al lanzar tres monedas al aire.

Solución:
El espacio muestral es S = {CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS}
La probabilidad de cada punto muestral es de 1/8

Se elabora las distribuciones de probabilidad y se realiza los cálculos respectivos. Estos resultados se presentan en la siguiente tabla:

x_i	$P(x_i)$	$x_i \cdot P(x_i)$	$(x_i - \mu)^2 \cdot P(x_i)$
0	1/8	0·1/8 = 0	$(0\text{-}1{,}5)^2 \cdot 1/8 = 0{,}281$
1	3/8	1·3/8 = 3/8	$(1\text{-}1{,}5)^2 \cdot 3/8 = 0{,}094$
2	3/8	2·3/8 = 3/4	$(2\text{-}1{,}5)^2 \cdot 3/8 = 0{,}094$
3	1/8	3·1/8 = 3/8	$(3\text{-}1{,}5)^2 \cdot 1/8 = 0{,}281$
Total	1	1,5	0,750

Observando la tabla se tiene:
$\mu = \mathrm{E(X)} = 1{,}5$; $\sigma^2 = 0{,}75$

Y calculando la desviación estándar se obtiene:
$\sigma = \sqrt{\sigma^2} = \sqrt{0{,}75} = 0{,}866$

Los cálculos en Excel de la esperanza matemática, la varianza y la desviación estándar se muestran en la siguiente figura:

	A	B	C	D	E
1	x_i	$P(x_i)$	$(x_i-\mu)^2$		
2	0	1/8	2,25	=(A2-$B$7)^2	
3	1	3/8	0,25	=(A3-$B$7)^2	
4	2	3/8	0,25	=(A4-$B$7)^2	
5	3	1/8	2,25	=(A5-$B$7)^2	
6					
7	μ	1,5	=SUMAPRODUCTO(A2:A5;B2:B5)		
8	σ^2	0,75	=SUMAPRODUCTO(C2:C5;B2:B5)		
9	σ	0,866	=RAIZ(B8)		

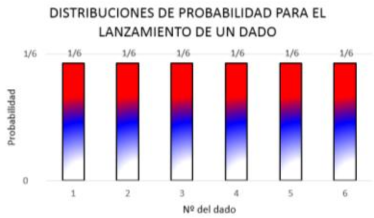
Interpretación:
El valor de $\mu = E(X) = 1{,}5$ significa que si se promedian los resultados del lanzamiento de las tres monedas (teóricamente, un número infinito de lanzamientos), se obtendrá 1,5.
Los valores de $\sigma^2 = 0{,}75$ y $\sigma = 0{,}866$ miden la dispersión de los resultados de lanzar las tres monedas alrededor de su media.

TAREA DE INTERAPRENDIZAJE N° 6

1) Elabore un organizador gráfico sobre las distribuciones discretas

2) Al ser la esperanza matemática una media aritmética ponderada, explique el por qué en su fórmula no aparece la división por la suma de los pesos como en cualquier fórmula de la media aritmética ponderada.

3) Sea el experimento aleatorio de lanzar un dado al aire.
3.1) Elabore un gráfico de distribuciones de probabilidad en 2D de manera manual y empleando Excel



3.2) Elabore un gráfico de distribuciones de probabilidad en 3D empleando Excel

3.3) Calcule la esperanza matemática, la varianza y desviación estándar de manera manual y empleando Excel.

$E(X) = 3{,}5$; $\sigma^2 = 2{,}917$; $\sigma = 1{,}71$

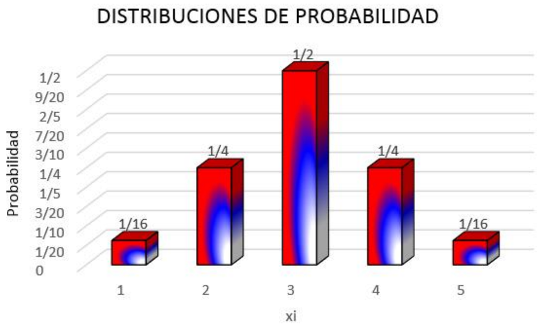
4) Dada las distribuciones de probabilidad

x_i	$P(x_i)$
0	x
1	1/4
2	6x
3	4x
4	1/16

4.1) Calcular el valor de x

1/16

4.2) Elabore un gráfico de distribuciones de probabilidad en 3D empleando Excel



4.3) Calcule la esperanza matemática, la varianza y desviación estándar de manera manual y empleando Excel.

E(X) = 2 ; $\sigma^2 = 1; \sigma = 1$

5) El número de automóviles que la empresa D & M vendió mensualmente varió de 4 a 12 junto con la frecuencia de ventas que se muestra en la siguiente tabla:

N° meses	Automóviles (x_i)
6	4
8	8
12	10
10	12
8	14
4	12

En meses anteriores el número promedio de ventas mensuales fue de 8 con una variabilidad de 4,2. Empleando las cifras presentadas, determine que ha pasado el promedio mensual de ventas y su variabilidad de la empresa D & M en comparación con los meses anteriores. Realice los cálculos empleando Excel.

Como E(X) = 10,167 y $\sigma = 2{,}995$ se evidencia que la empresa ha incrementado su promedio mensual de ventas y ha reducido su variabilidad en comparación con los meses anteriores.

6) Plantee y resuelva 3 ejercicios de aplicación sobre distribuciones discretas. En cada ejercicio elabore gráficos de distribuciones de probabilidad en 3D empleando Excel. Realice los cálculos empleando Excel.

C) DISTRIBUCIÓN BINOMIAL

i) Definición:
Cuando se dispone de una expresión matemática, es factible calcular la probabilidad de ocurrencia exacta correspondiente a cualquier resultado específico para la variable aleatoria.

La *distribución de probabilidad binomial* es uno de los modelos matemáticos (expresión matemática para representar una variable) que se utiliza cuando la variable aleatoria discreta es el número de éxitos en una muestra compuesta por n observaciones.

ii) Propiedades:

- La muestra se compone de un número fijo de observaciones n

- Cada observación se clasifica en una de dos categorías, *mutuamente excluyentes* (los eventos no pueden ocurrir de manera simultánea. Ejemplo: Una persona no puede ser de ambos sexos) y *colectivamente exhaustivos* (uno de los eventos debe ocurrir. Ejemplo: Al lanzar una moneda, si no ocurre cruz, entonces ocurre cara). A estas categorías se las denomina éxito y fracaso.

- La probabilidad de que una observación se clasifique como *éxito, p,* es constante de una observación o otra. De la misma forma, la probabilidad de que una observación se clasifique como *fracaso, 1-p,* es constante en todas las observaciones.

- La variable aleatoria binomial tiene un rango de 0 a n

iii) Ecuación:

$$P(X) = \frac{n!}{X!\,(n-X)!} \cdot p^X \cdot (1-p)^{n-X}$$

Donde:
$P(X)$ =Probabilidad de X éxitos, dadas n y p
n = Número de observaciones
p = Probabilidad de éxitos
$1-p$ = Probabilidad de fracasos
X = Número de éxitos en la muestra (X= 0, 1, 2, 3, 4,………n)

iv) Media de la distribución binomial
La media μ de la distribución binomial es igual a la multiplicación del tamaño n de la muestra por la probabilidad de éxito p

$$\mu = np$$

v) Desviación estándar de la distribución binomial

$$\sigma = \sqrt{\sigma^2} = \sqrt{np(1-p)}$$

Ejemplos ilustrativos

1) Determine P(X = 8) para n = 10 y p = 0,5

Solución:
Aplicando la ecuación se obtiene:

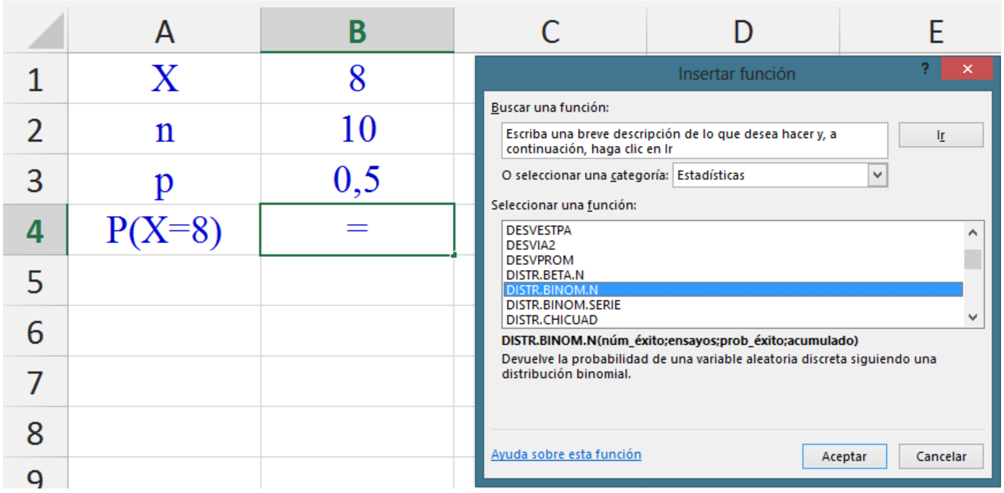
$$P(X) = \frac{n!}{X!\,(n-X)!} \cdot p^X \cdot (1-p)^{n-X}$$

$$P(X=8) = \frac{10!}{8!\,(10-8)!} \cdot 0{,}5^8 \cdot (1-0{,}5)^{10-8}$$

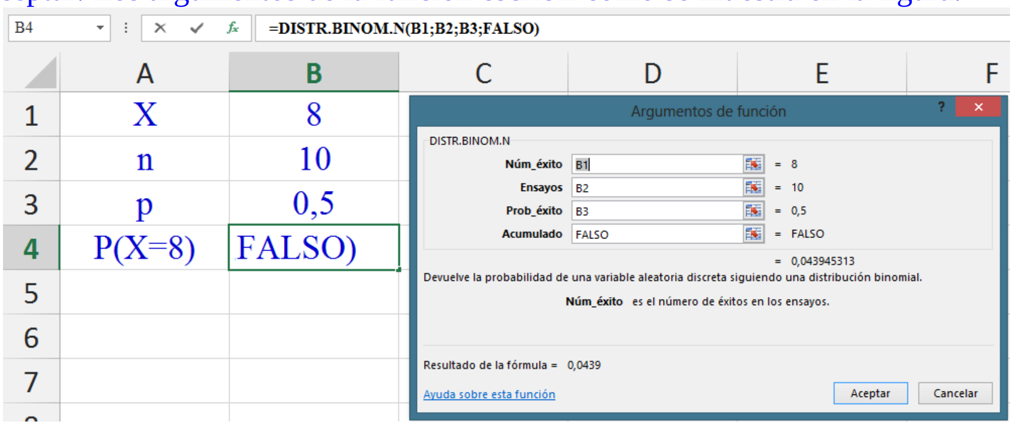
$$P(X=8) = 45 \cdot 0{,}003906 \cdot 0{,}25 = 0{,}0439$$

En Excel se calcula de la siguiente manera:

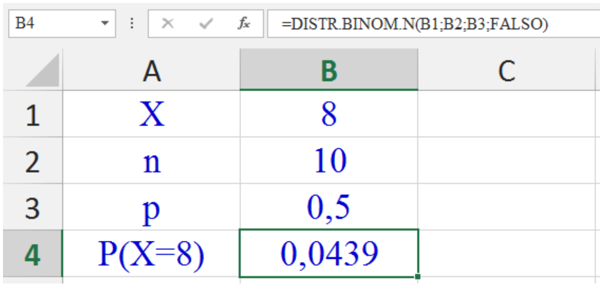
a) Se escribe los datos y se inserta la función DISTR.BINOM.N como se muestra en la siguiente figura:



b) Clic en Aceptar. Los argumentos de la función escribir como se muestra en la figura:



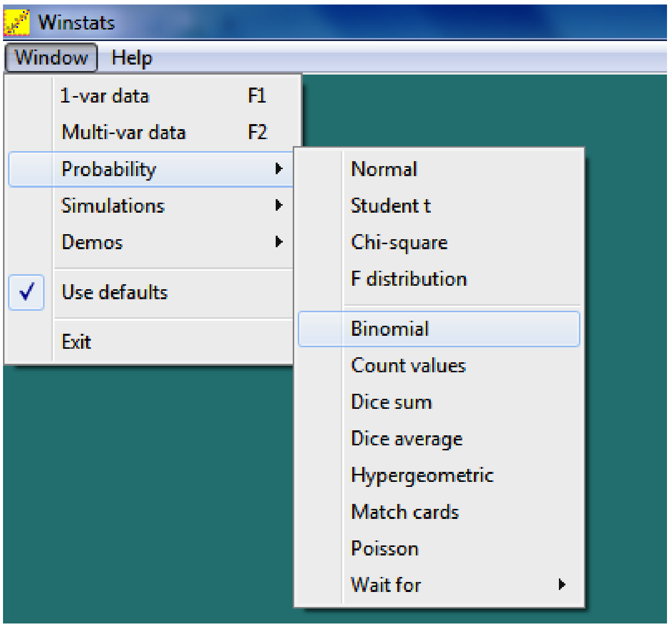
c) Clic en Aceptar



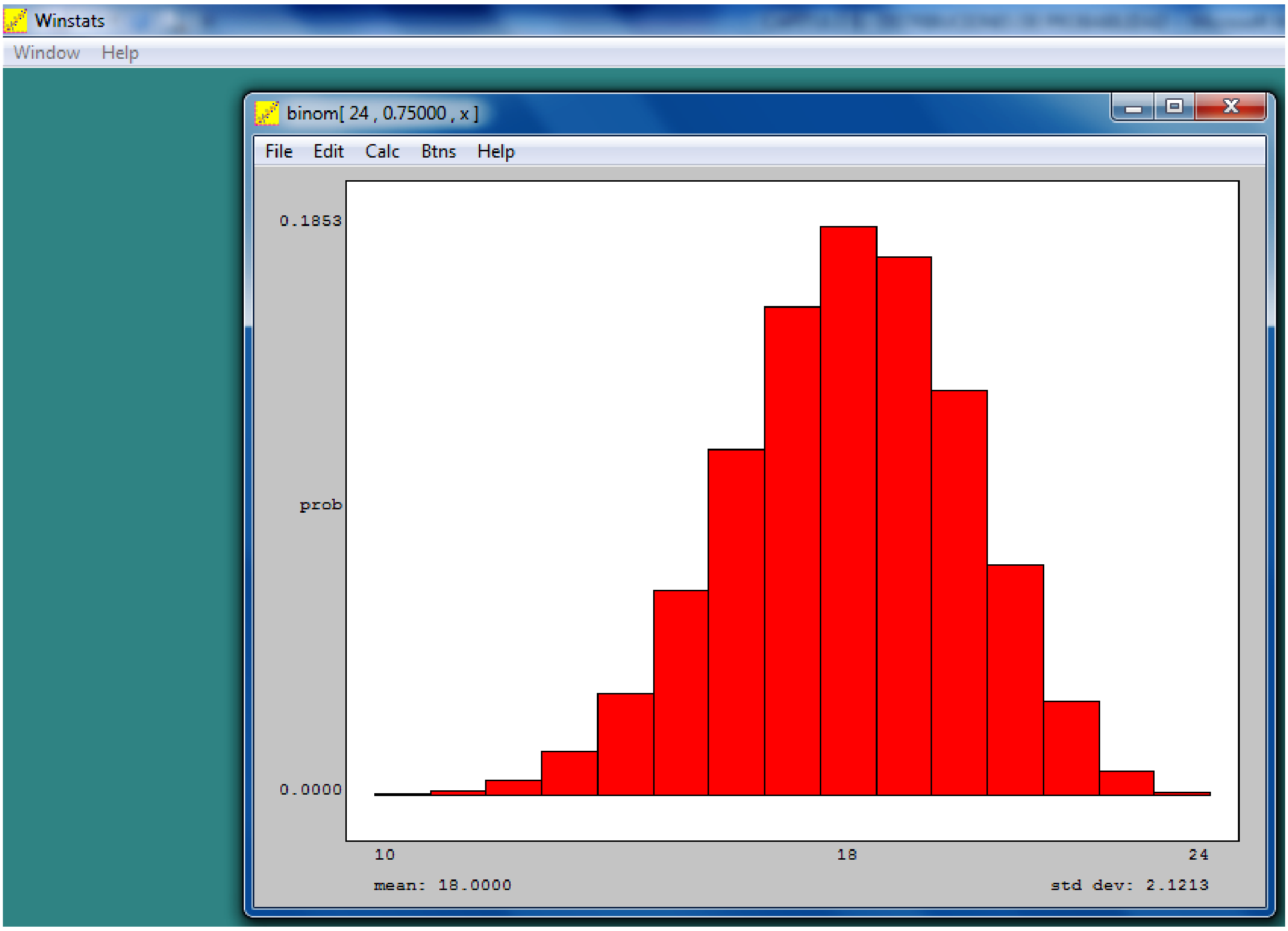
En Winstats se procede de la siguiente manera

a) Se ingresa al programa Winstats

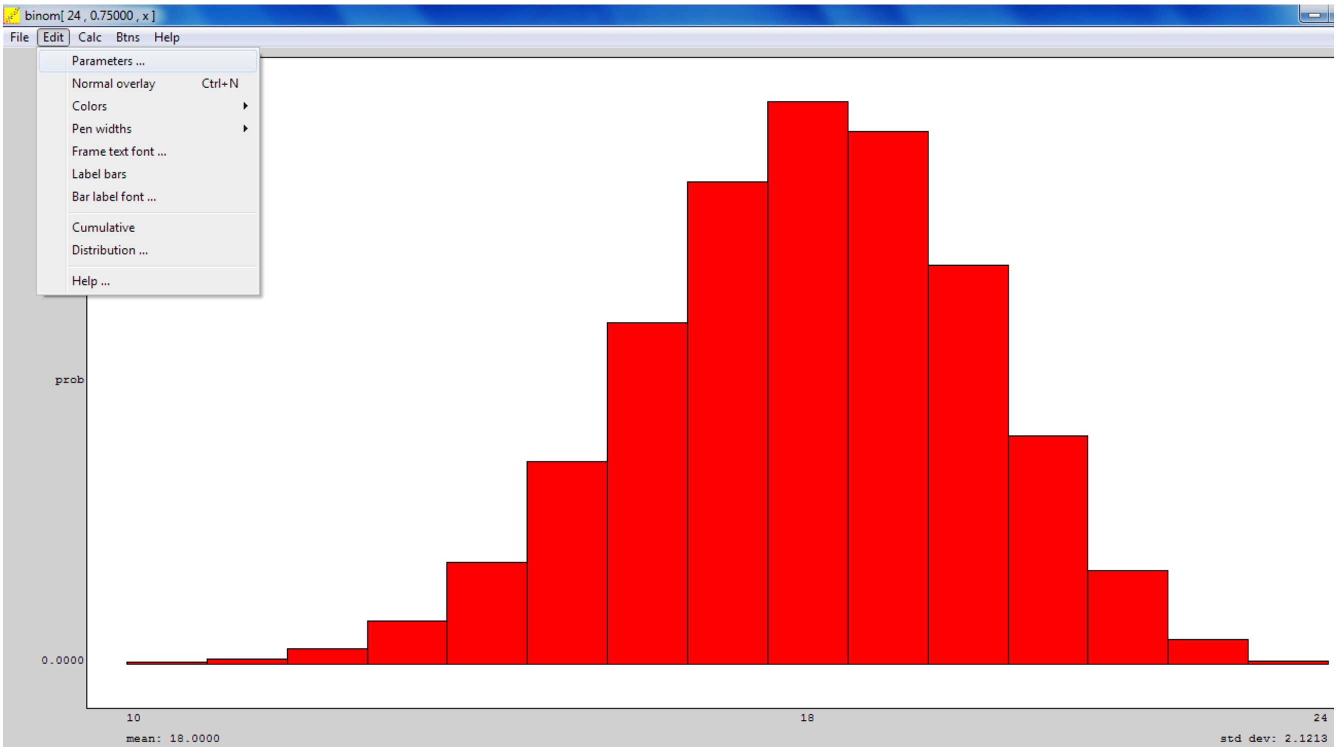
b) Clic en Window y luego en Probability



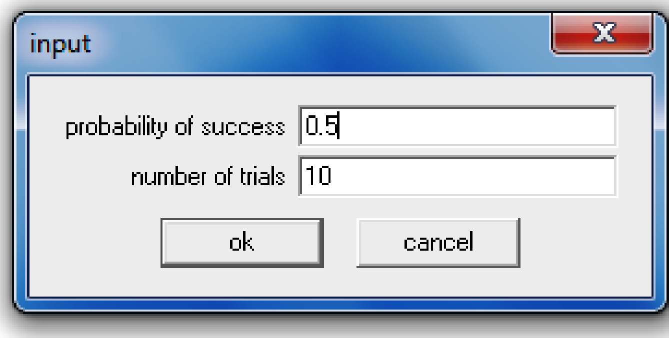
c) En Probability escoger Binomial



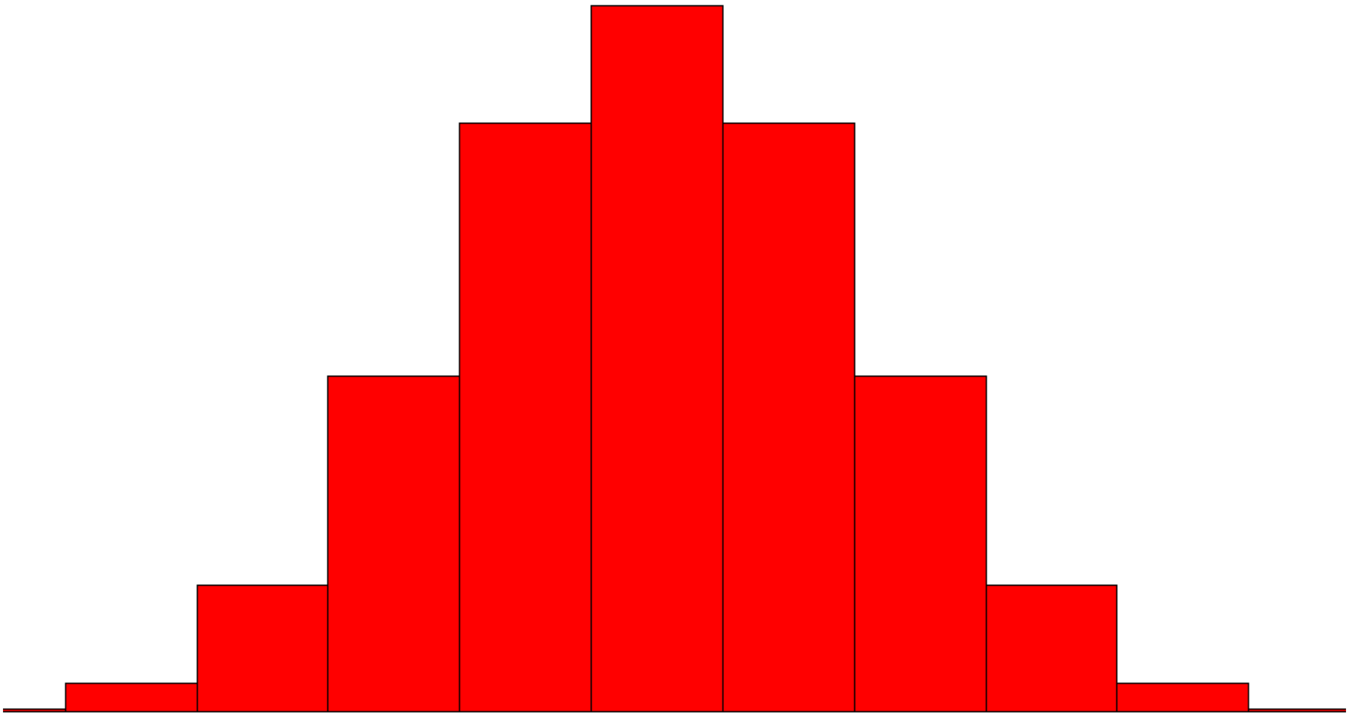
d) Clic en Edit.



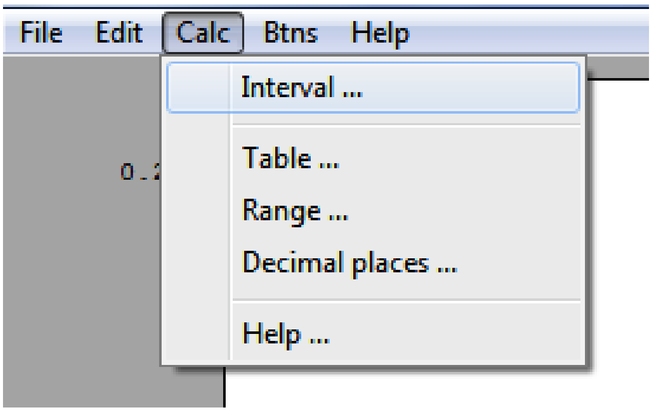
e) Clic en Parameters. En la casilla en probability of success escribir 0,5 y en number of trials escribir 10



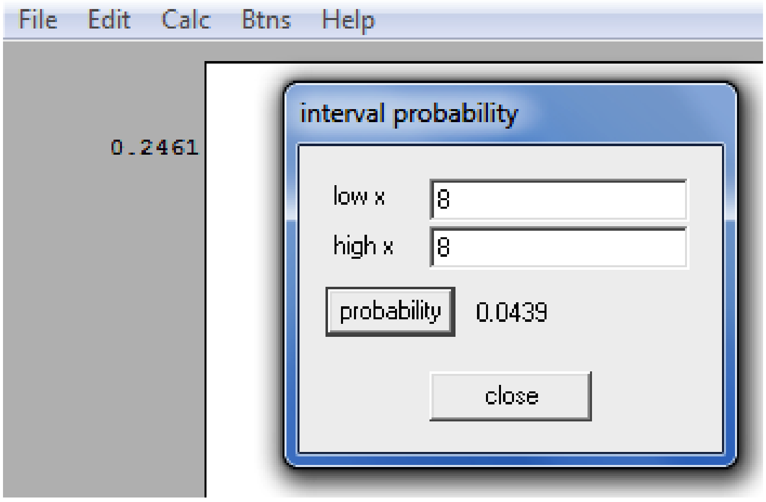
f) Clic en ok

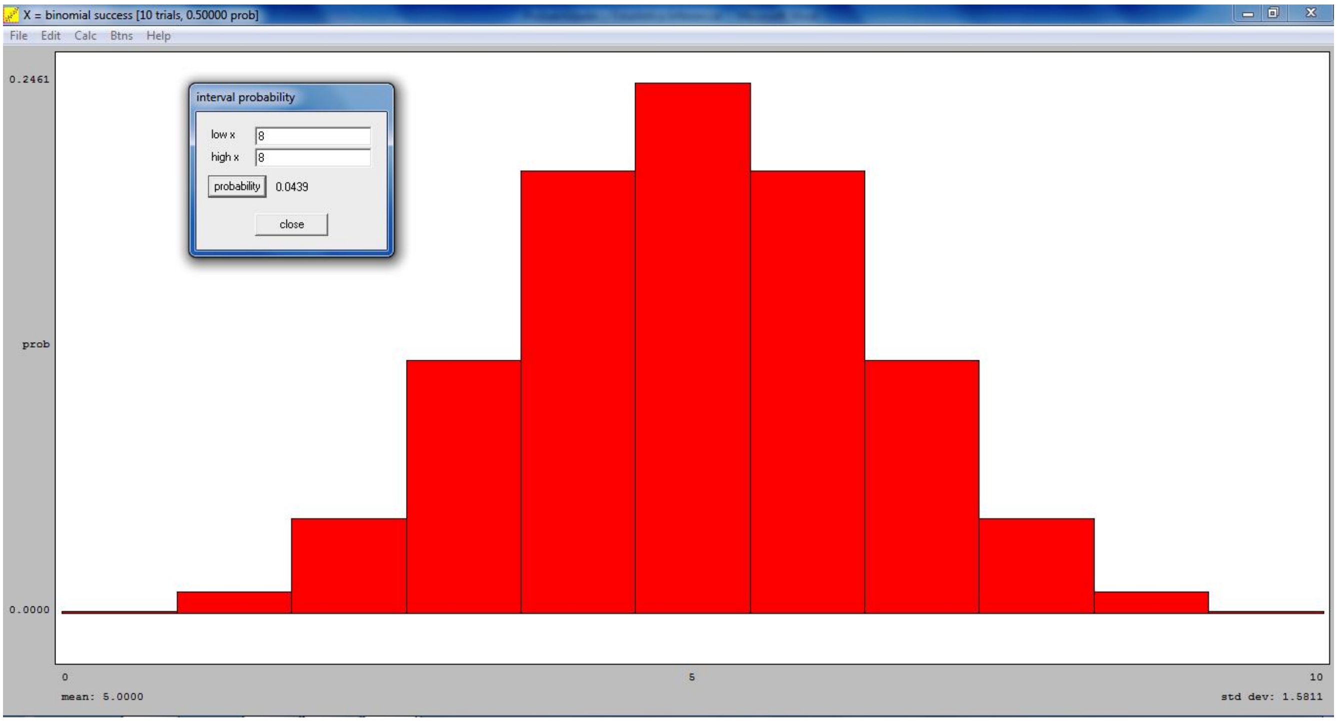


g) Clic en Calc



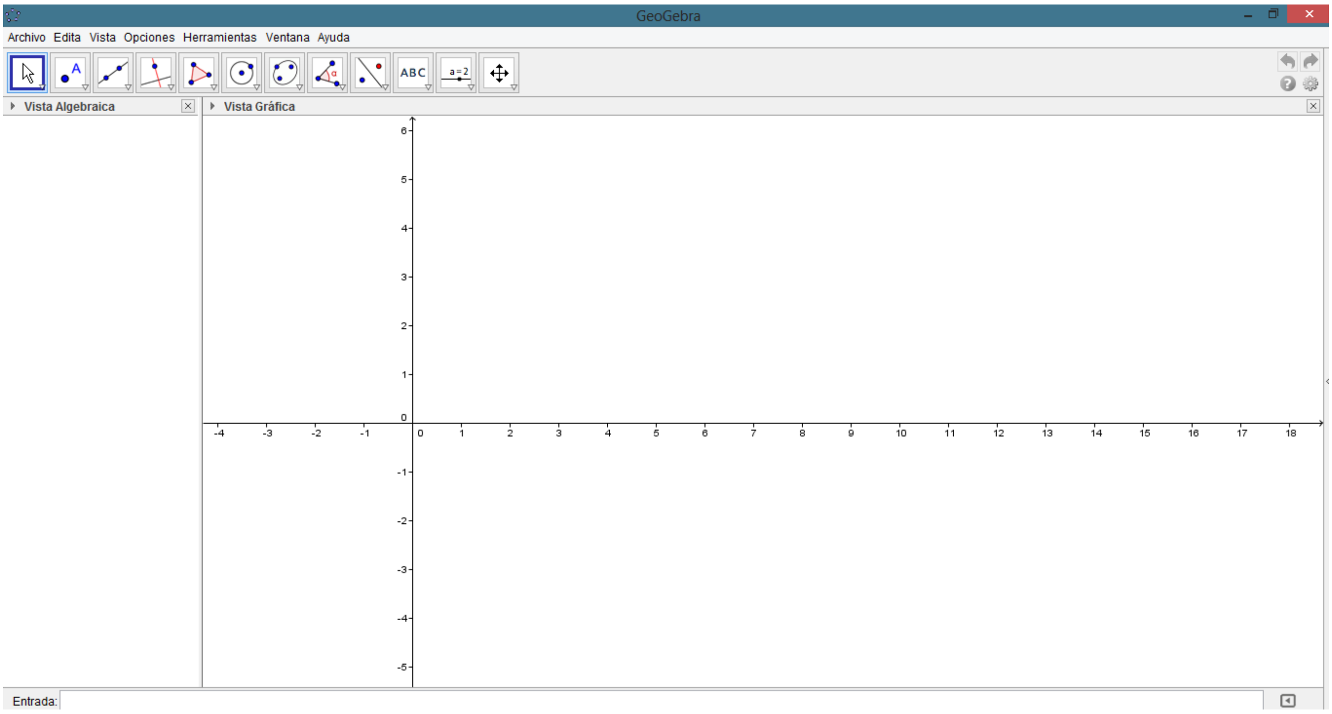
h) Clic en Intervalo. En la casilla low x escribir 8 y en la casilla high x escribir 8. Clic en probability



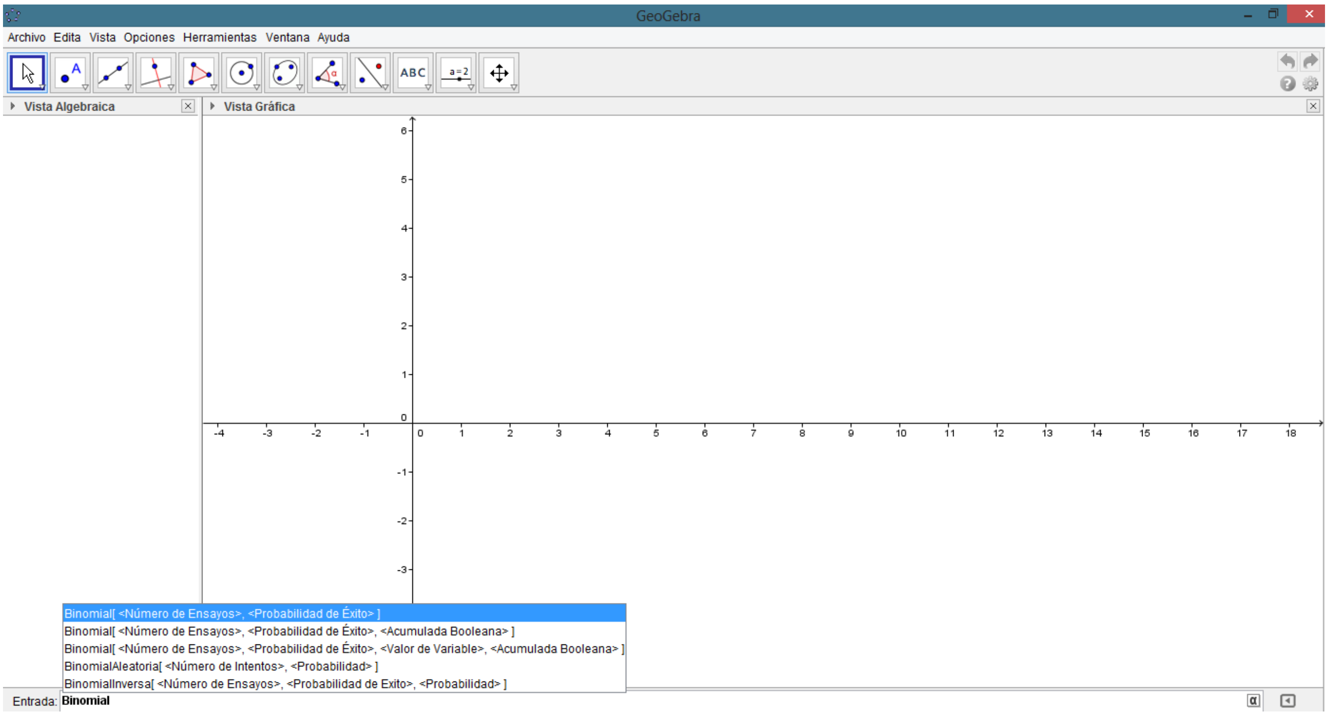


En GeoGebra se procede de la siguiente manera:

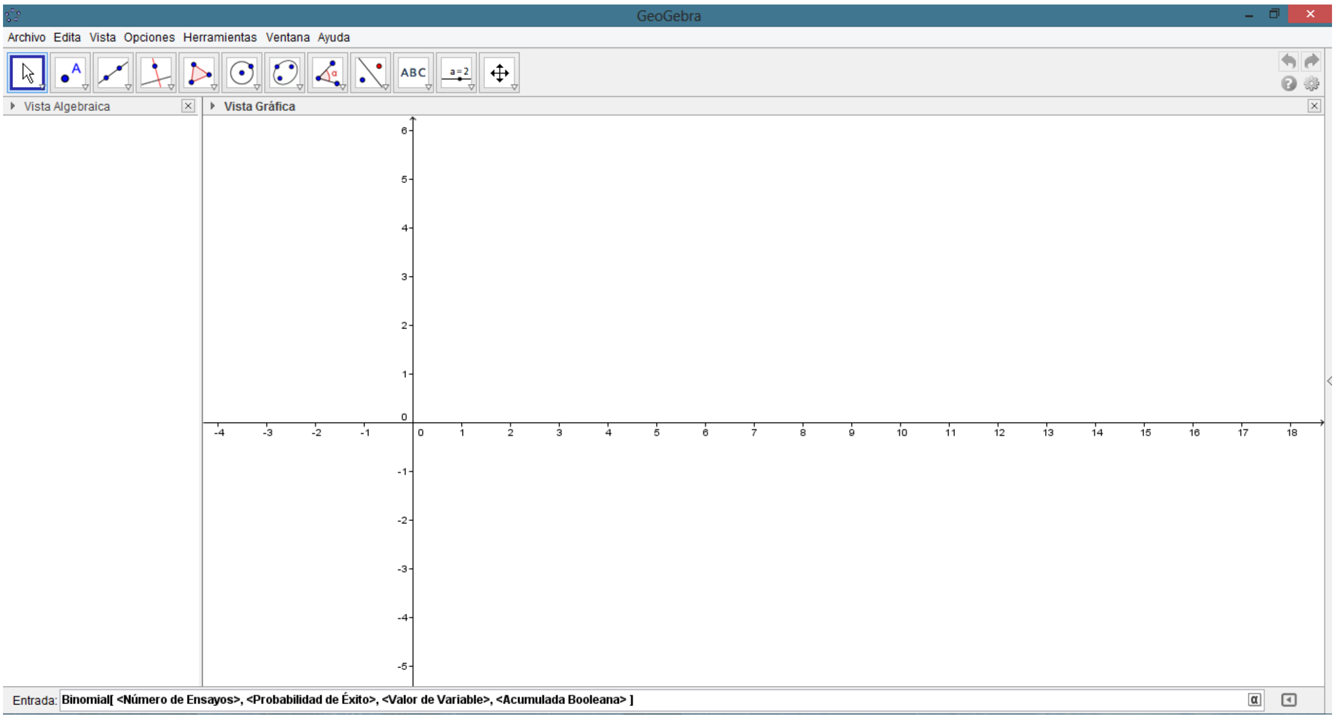
a) Se ingresa al programa



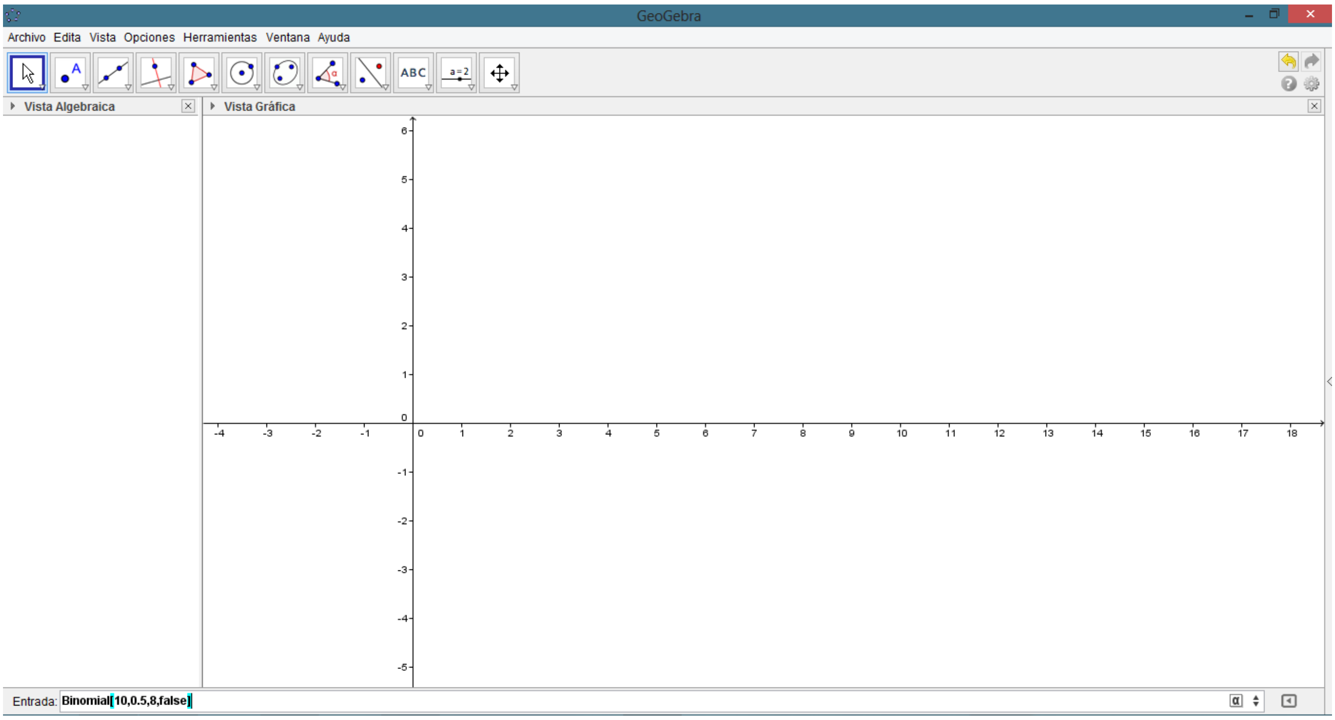
b) En la casilla Entrada escribir Binomial para que se desplieguen algunas opciones.



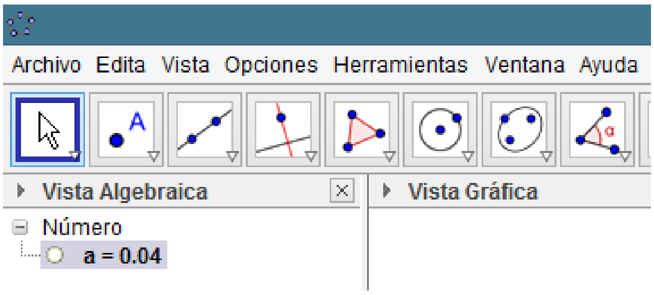
c) Seleccionar la opción Binomial[<Número de Ensayos>, <Probabilidad de Éxito>, <Valor de Variable>, <Acumulada Booleana>]

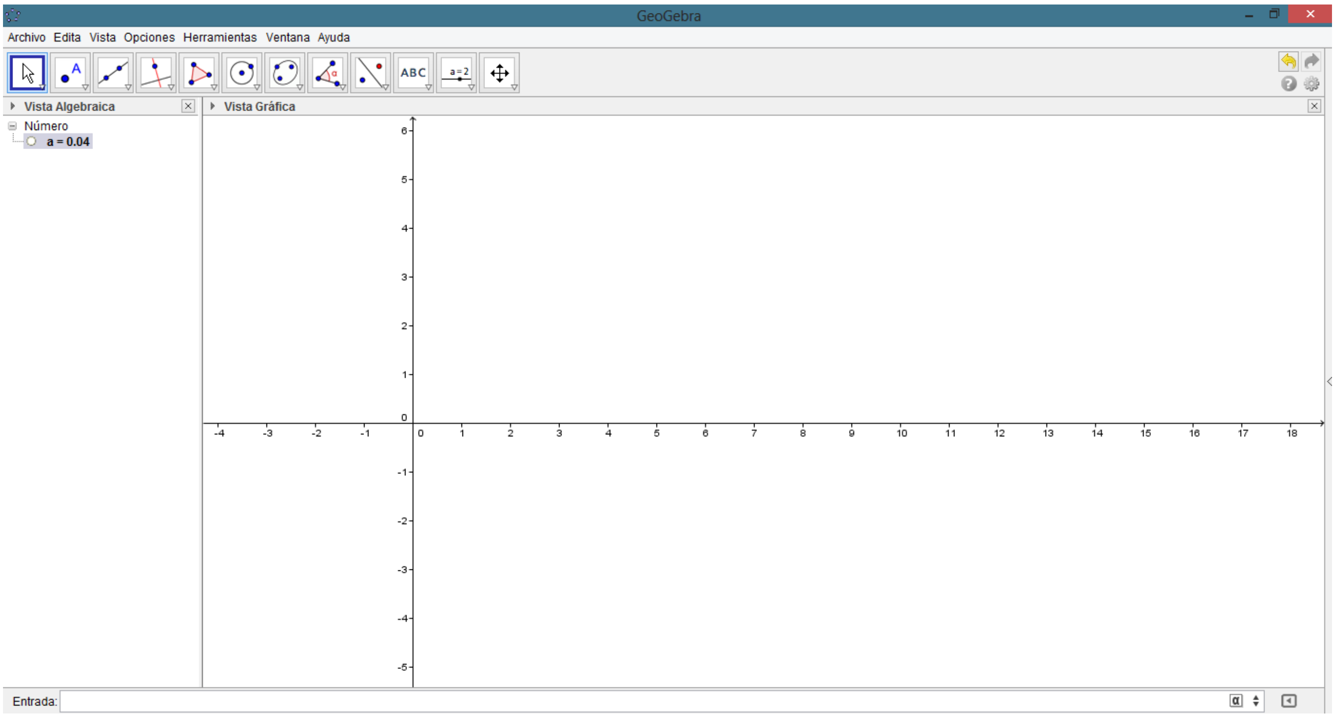


d) Escribir 10 en <Número de Ensayos>, 0.5 en <Probabilidad de Éxito>, 8 en <Valor de Variable> y false en <Acumulada Booleana>

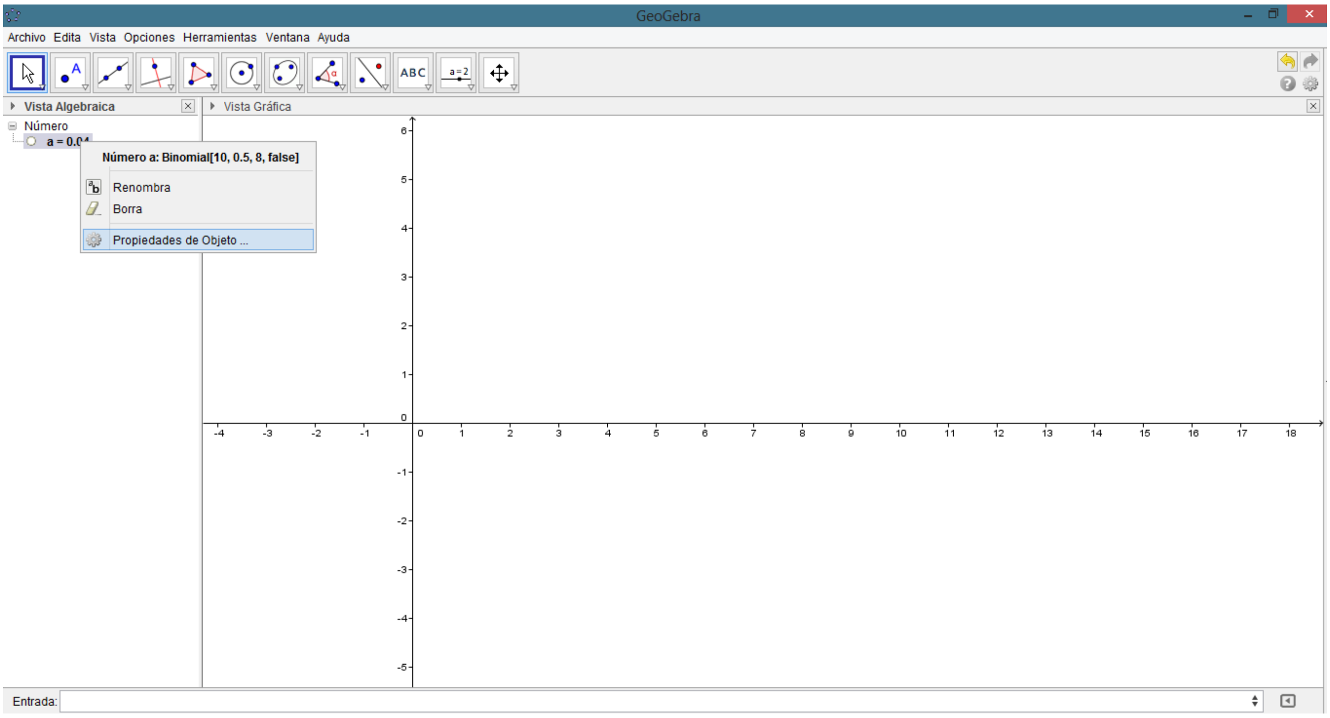


e) Enter

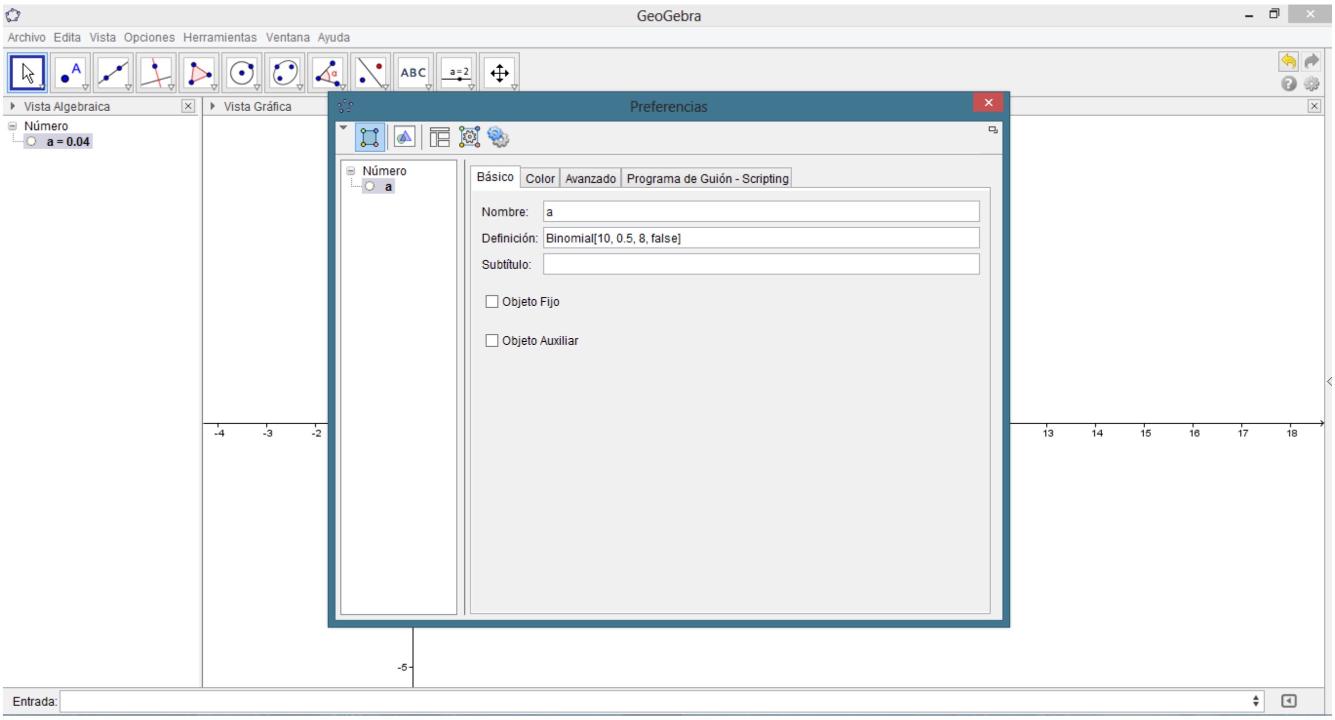




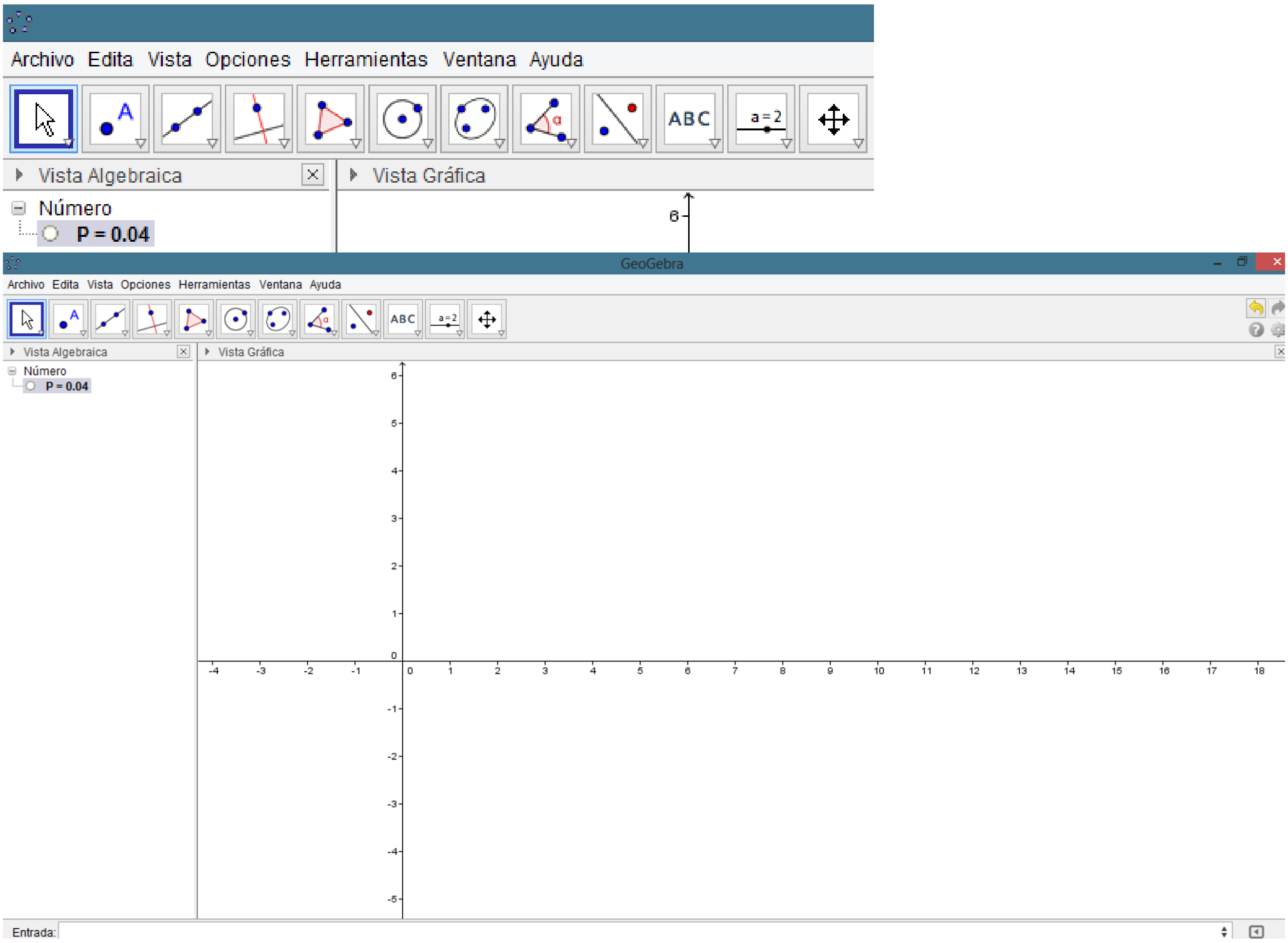
f) Para editar. Clic derecho en a = 0.04



g) Escoger la opción Propiedades de Objeto

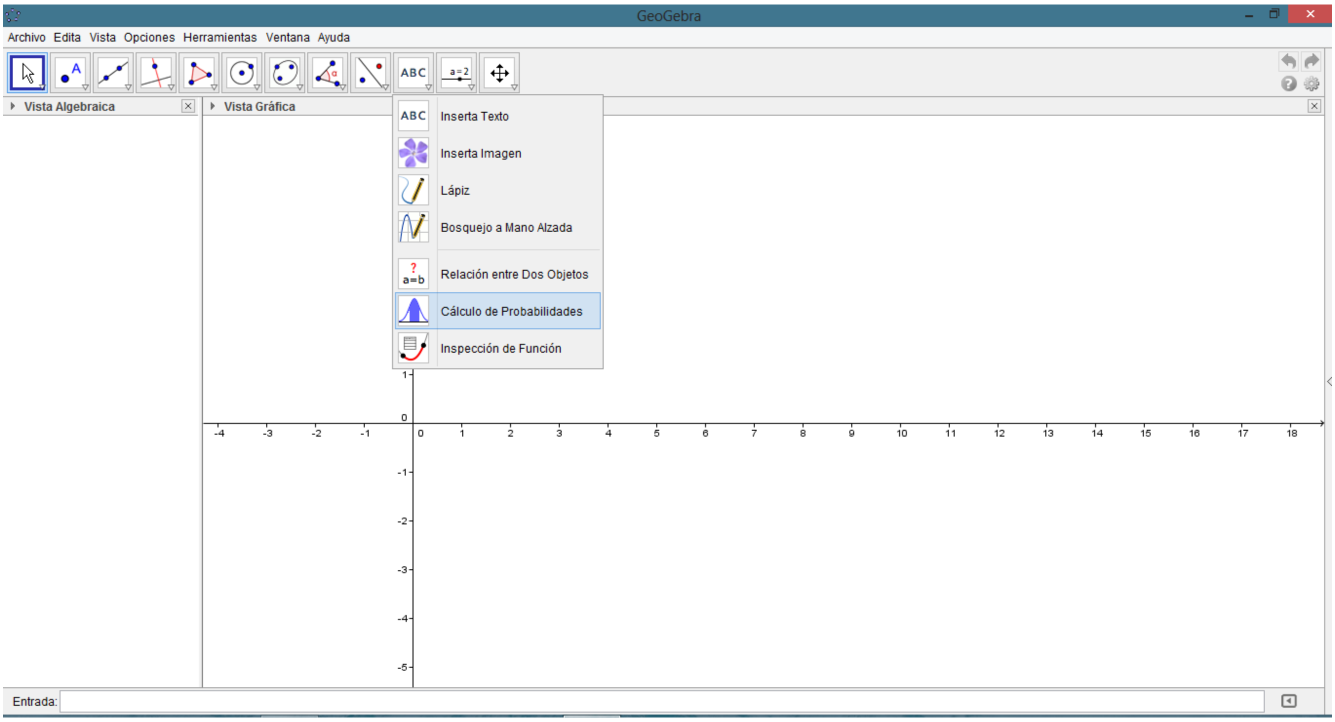


h) En la ventana Referencias, en la casilla Nombre, borrar la letra a y escribir P. Cerrar la ventana Refreencias

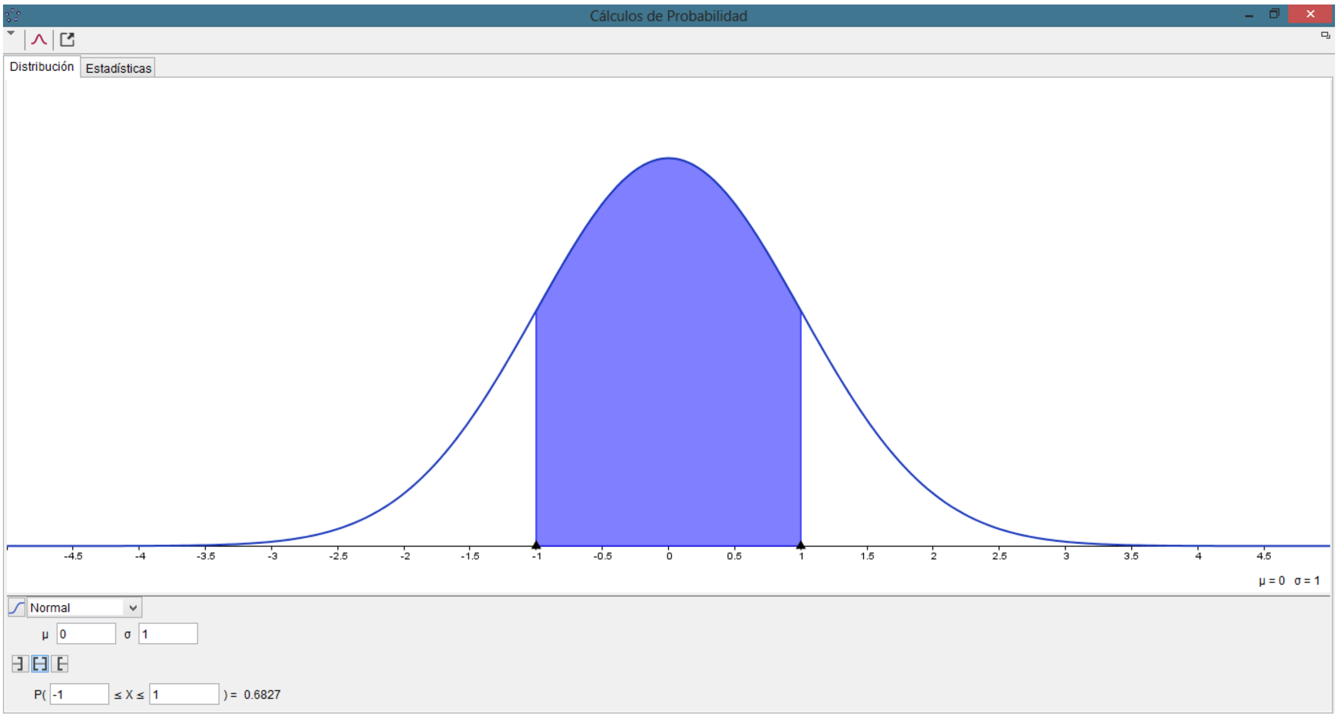


Para calcular con el gráfico en GeoGebra:

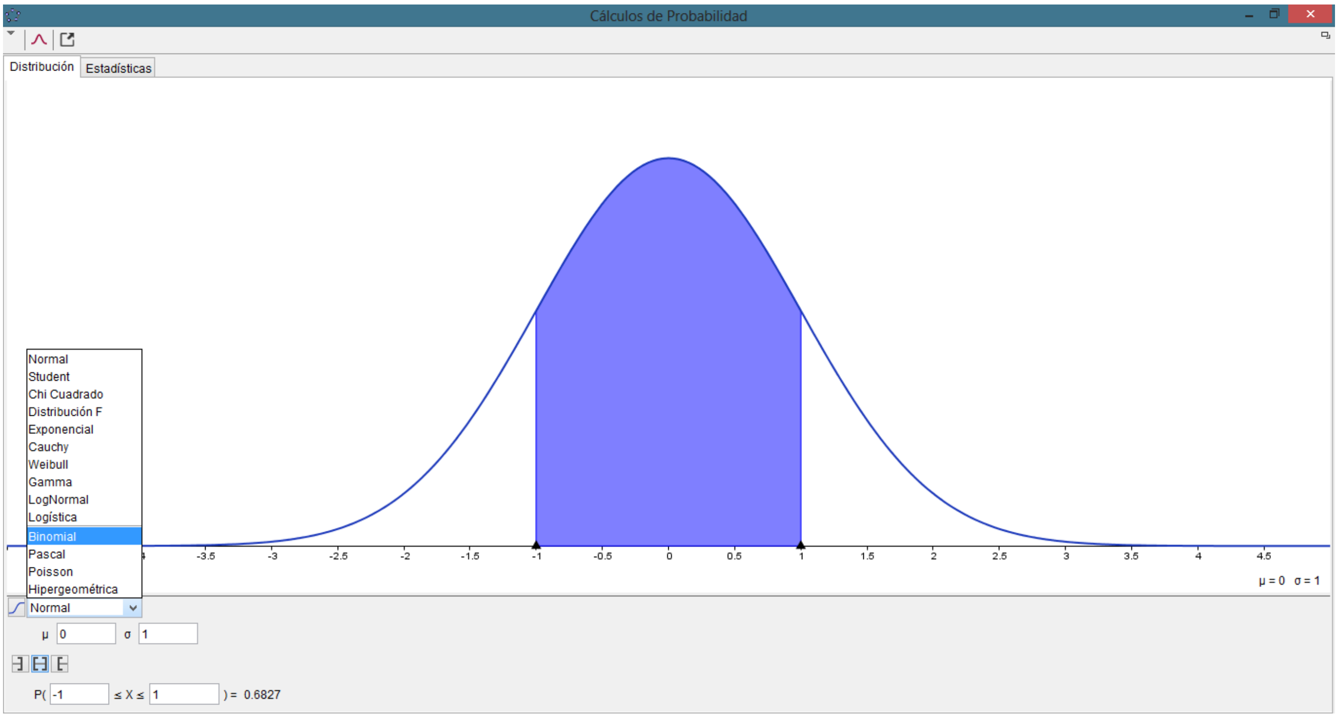
a) Ingresar al programa. En insertar texto, clic en punto de posición del texto



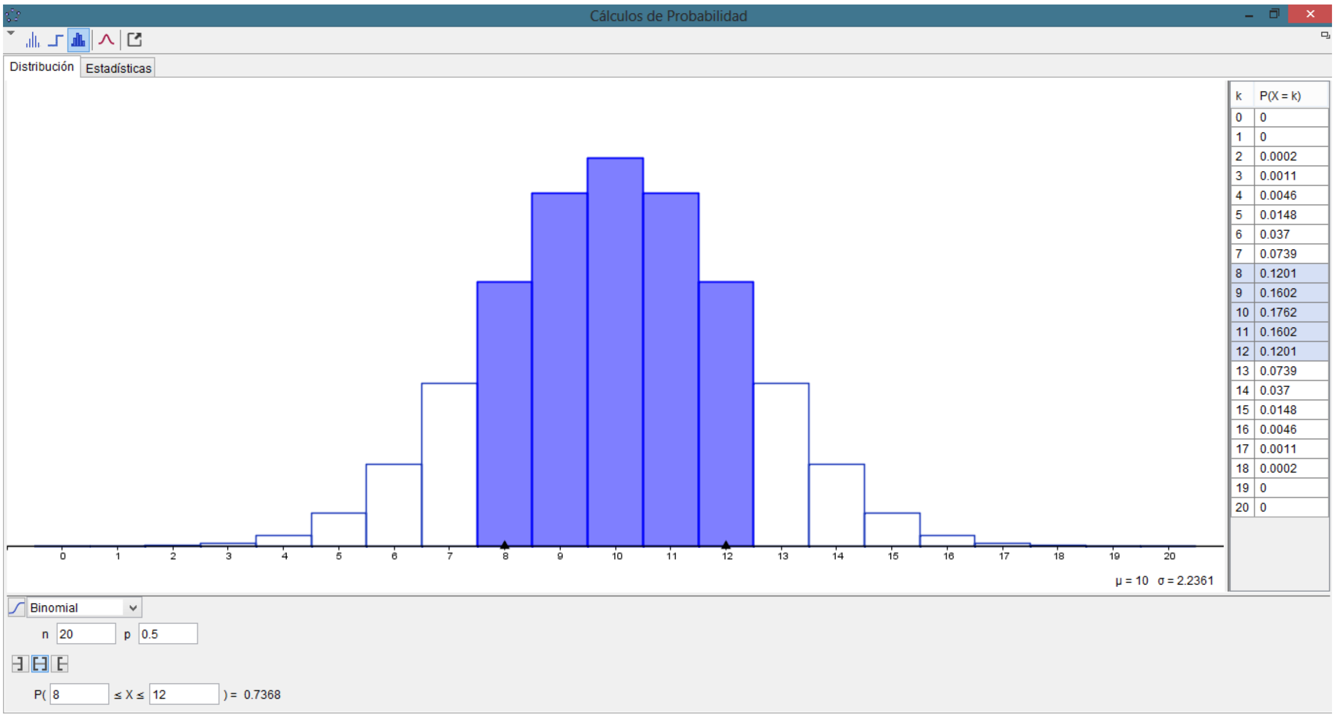
b) Seleccionar Cálculo de Probabilidades



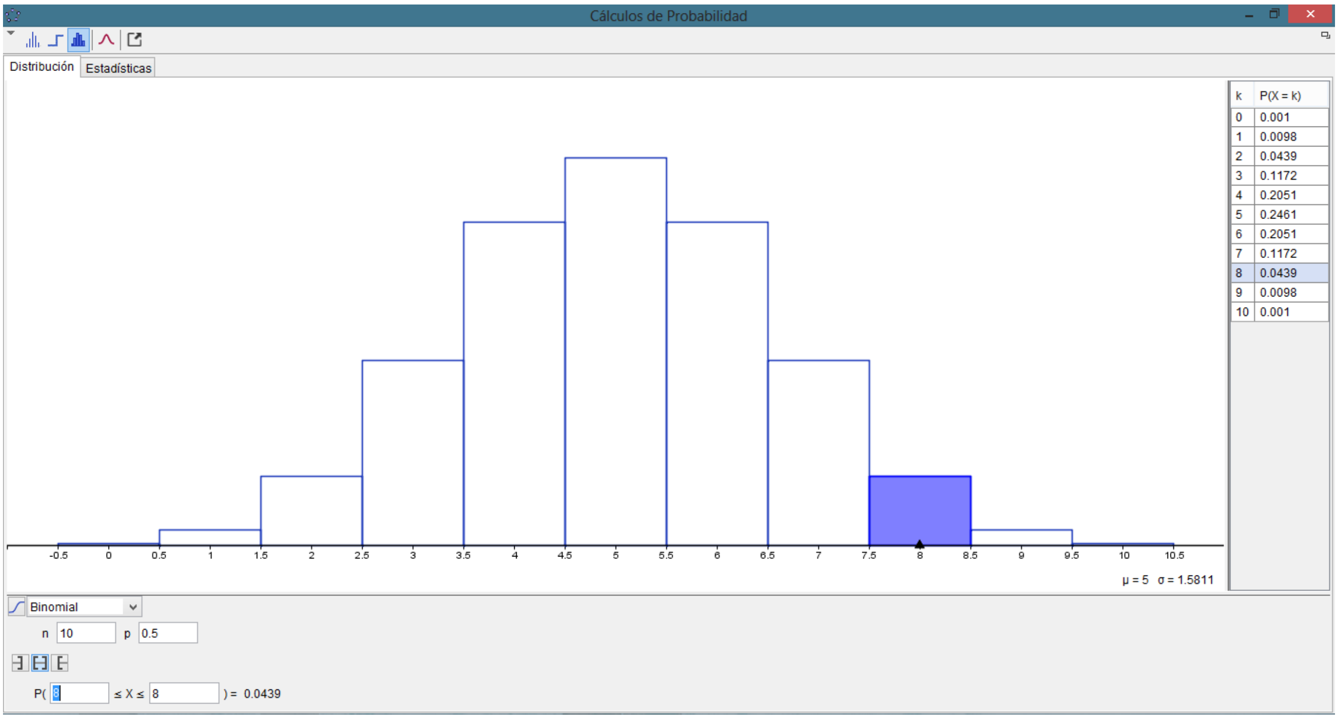
c) Clic en la pestaña de la casilla Normal para que se despliegue otras opciones.

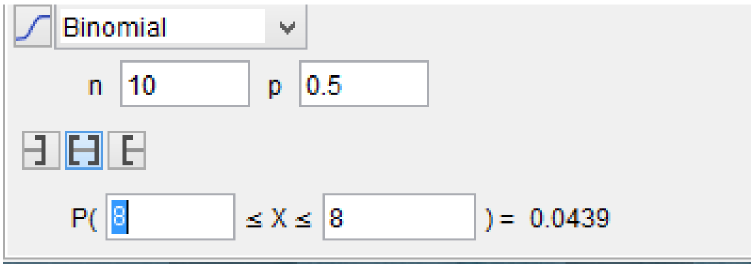


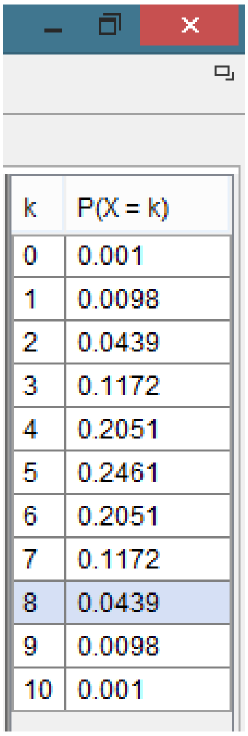
d) Clic en Binomial



e) En la casilla n escribir 10. En la casilla p escribir 0.5. En la casilla P escribir 8. En la casilla X≤ escribir 8. Enter







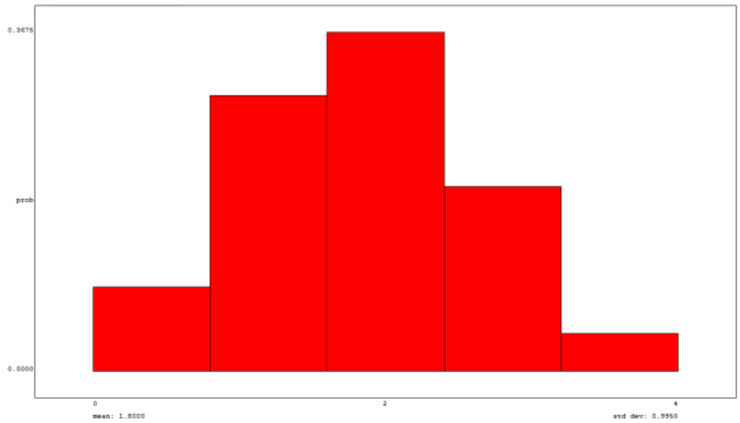
2) Determinar P (X ≤ 3) para n =4 y p = 0,45

Solución:

$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$

Se puede aplicar la ecuación para cada probabilidad, pero para ahorrar tiempo se recomienda encontrar las probabilidades con lectura en la tabla de probabilidades binomiales.

TABLA Nº 1
PROBABILIDADES BINOMIALES

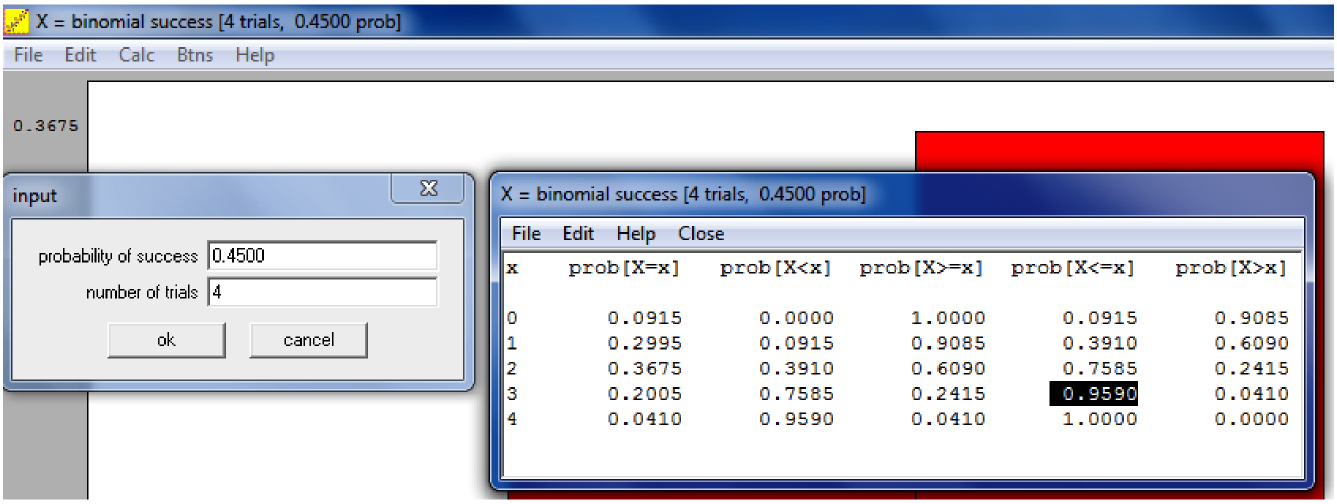


$\boldsymbol{Ejemplo: Para\ n = 4\ y\ p = 0,45\ \Rightarrow P(X = 2) = 0,3675}$

n	X	P									
		0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	0	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000	0,6500	0,6000	0,5500	0,5000
1	1	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000	0,3500	0,4000	0,4500	0,5000
n	X	P									
2	0	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900	0,4225	0,3600	0,3025	0,2500
2	1	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200	0,4550	0,4800	0,4950	0,5000
2	2	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900	0,1225	0,1600	0,2025	0,2500
n	X	P									
3	0	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430	0,2746	0,2160	0,1664	0,1250
3	1	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410	0,4436	0,4320	0,4084	0,3750
3	2	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890	0,2389	0,2880	0,3341	0,3750
3	3	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270	0,0429	0,0640	0,0911	0,1250
n	X	P									
4	0	0,8145	0,6561	0,5220	0,4096	0,3164	0,2401	0,1785	0,1296	0,0915	0,0625
4	1	0,1715	0,2916	0,3685	0,4096	0,4219	0,4116	0,3845	0,3456	0,2995	0,2500
4	2	0,0135	0,0486	0,0975	0,1536	0,2109	0,2646	0,3105	0,3456	0,3675	0,3750
4	3	0,0005	0,0036	0,0115	0,0256	0,0469	0,0756	0,1115	0,1536	0,2005	0,2500
4	4	0,0000	0,0001	0,0005	0,0016	0,0039	0,0081	0,0150	0,0256	0,0410	0,0625

Realizando la lectura en la tabla de P(X=0) con n=4 y p = 0,45 se obtiene 0,0915. Continuando con la respectivas lecturas en la tabla se obtiene: 0,2995 para P(X=1), 0,3675 para P(X=2) y 0,2005 para P(X=3).
Por lo tanto $P(X \leq 3) = 0{,}0915 + 0{,}2995 + 0{,}3675 + 0{,}2005 = 0{,}9590$

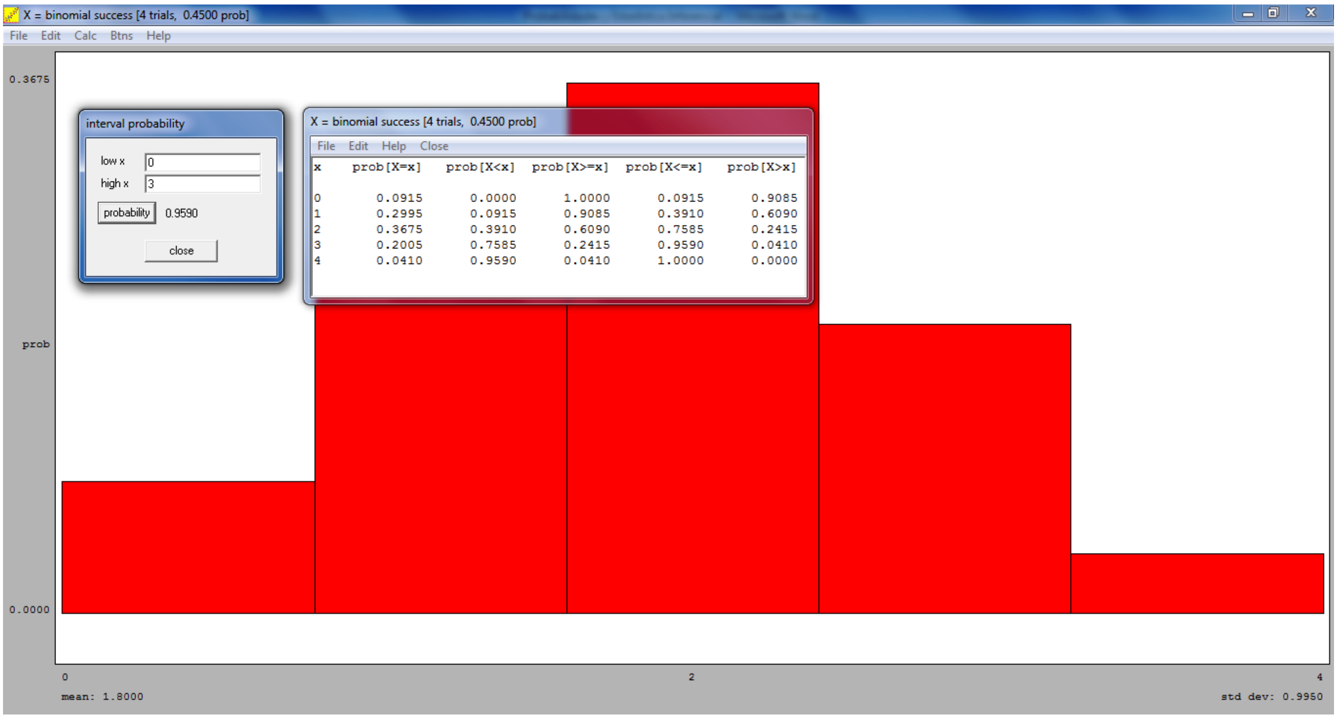
Para que aparezca la tabla en Winstats se hace clic en Edit y luego en parámetros. En la ventana de parámetros, en la casilla trials, escribir 4 y en success prob escribir 0,45. Finalmente clic Calc y luego en table



Los cálculos realizados en Excel se muestran en la siguiente figura:

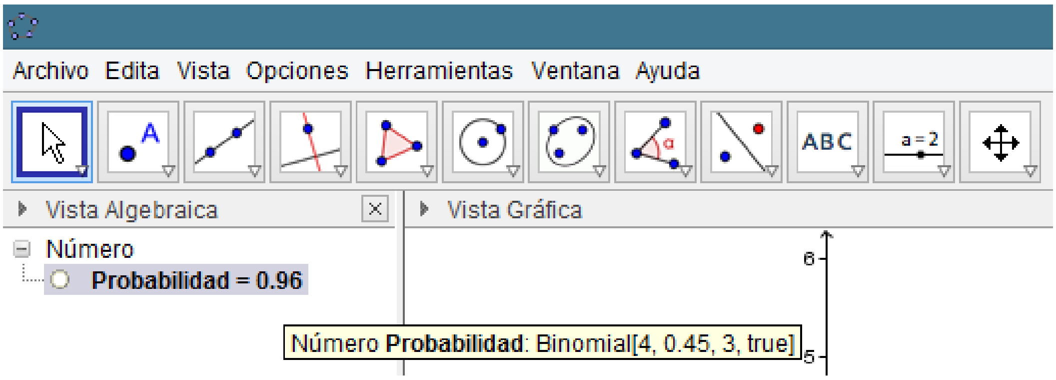
	A	B	C	D	E	F
1	X	3				
2	n	4				
3	p	4/9				
4	$P(X \leq 3)$	0,9590	=DISTR.BINOM.N(B1;B2;B3;VERDADERO)			

Los cálculos realizados en Winstats se muestran en la siguiente figura:



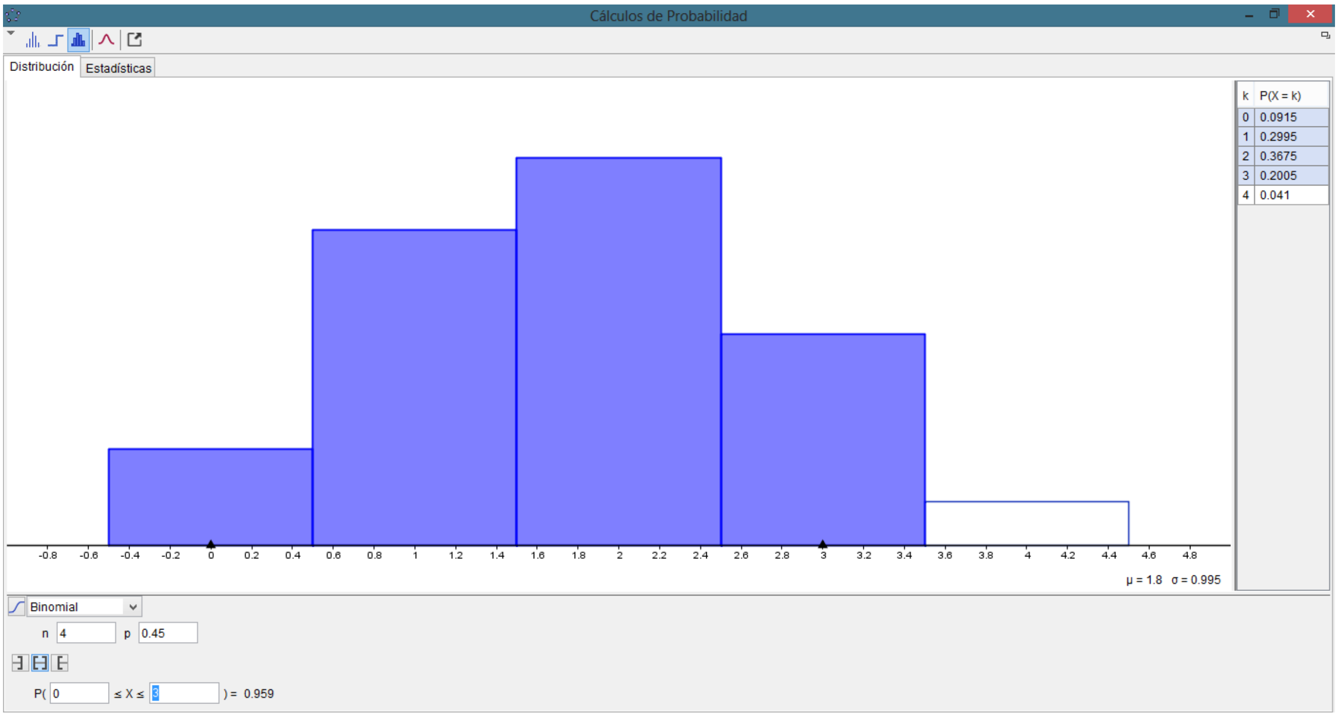
En GeoGebra

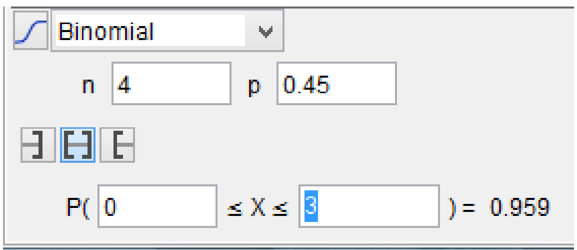
Escribir 4 en <Número de Ensayos>, 0.45 en <Probabilidad de Éxito>, 3 en <Valor de Variable> y true en <Acumulada Booleana>



Nota:
Para P(X= 3), siendo 3 el número de éxitos, en <Acumulada Booleana> se escribe false
Para P(X≤ 3), siendo 3 el número de éxitos, en <Acumulada Booleana> se escribe true

O También





3) Se lanza ocho dados.
3.1) Calcular la probabilidad de obtener 2 seis
3.2) Calcular la probabilidad de obtener máximo 2 seis
3.3) Calcular la probabilidad de obtener al menos 2 seis

Solución:

3.1)

$P(X = 2) = ?\ ; n = 8;\ p = \frac{1}{6}$

Aplicando la fórmula se obtiene:

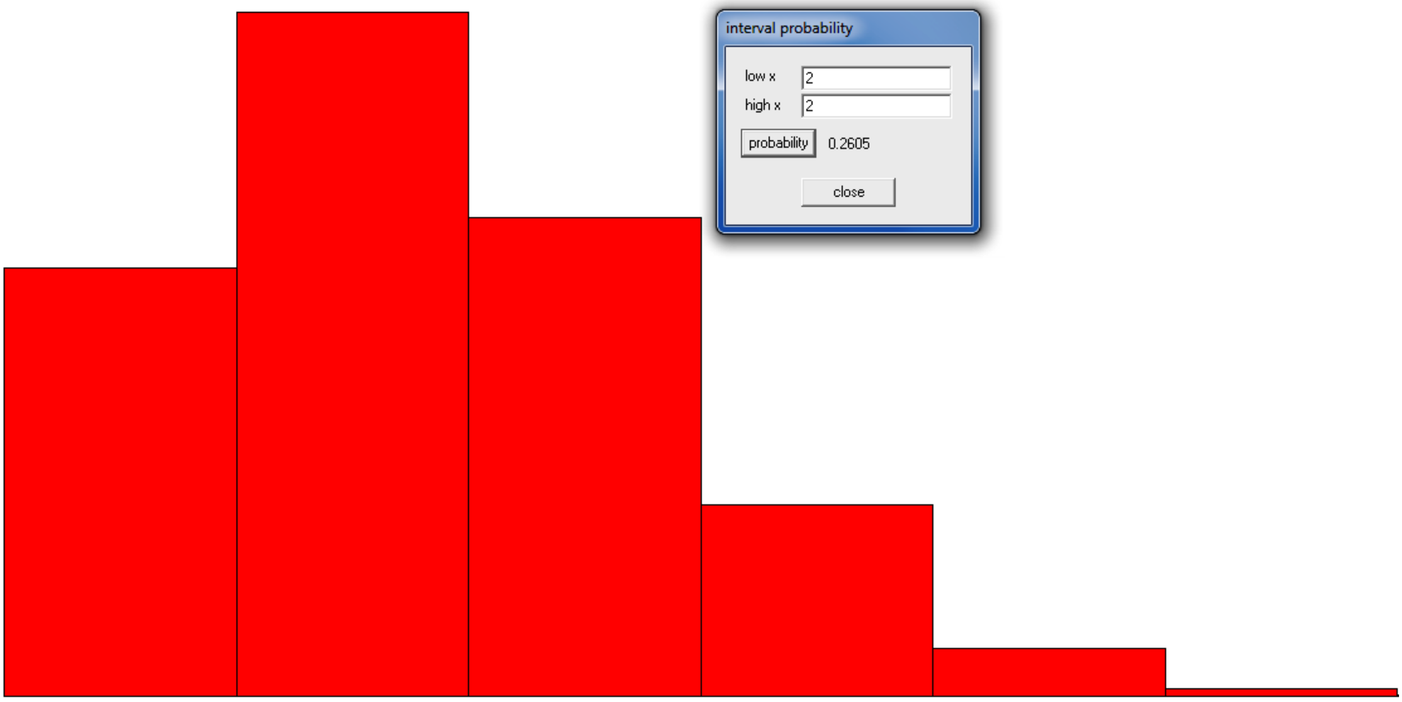
$$P(X) = \frac{n!}{X!\,(n-X)!} \cdot p^X \cdot (1-p)^{n-X}$$

$$P(X = 2) = \frac{8!}{2!\,(8-2)!} \cdot \frac{1}{6}^2 \cdot \left(1 - \frac{1}{6}\right)^{8-2} = 0{,}2605$$

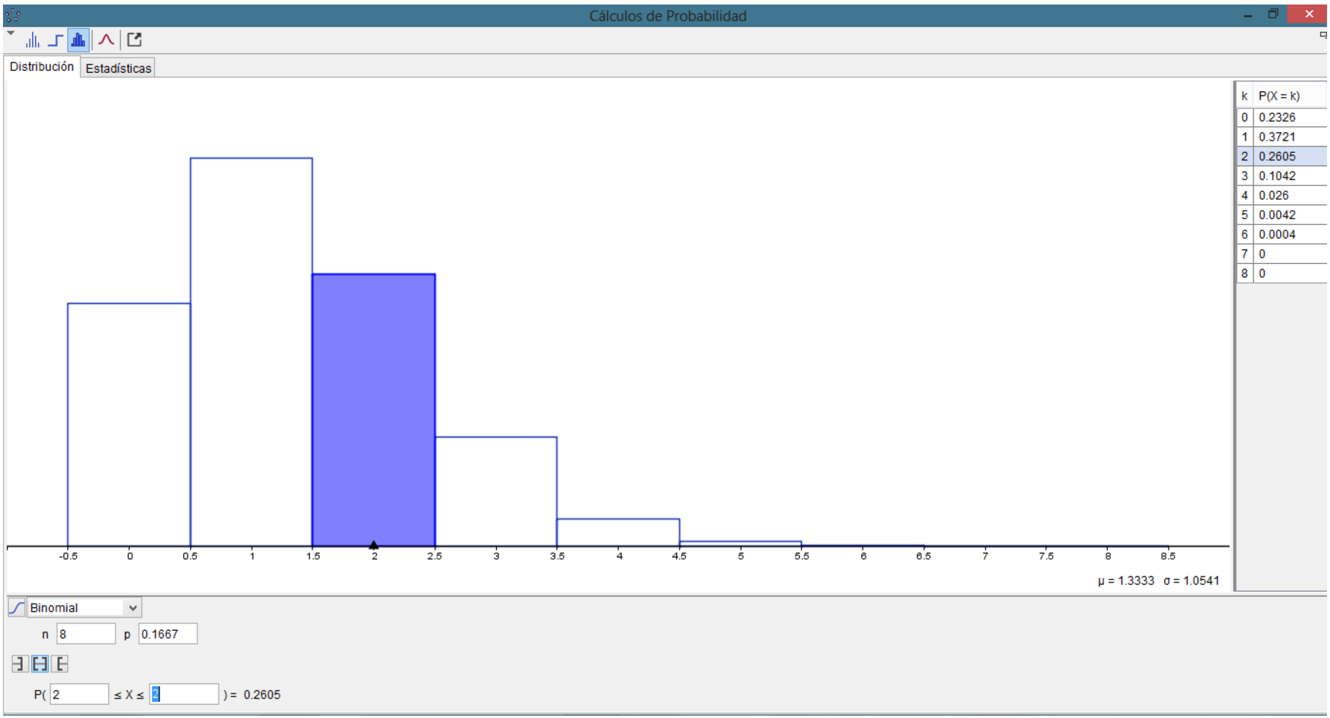
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	2			
2	n	8			
3	p	1/6			
4	P(X=2)	0,2605	=DISTR.BINOM.N(B1;B2;B3;FALSO)		

Los cálculos en Winstats se muestran en la siguiente figura:



Los cálculos en GeoGebra se muestran en la siguiente figura:



3.2)

$$P(X \leq 2) =?\ ; n = 8;\ p = \frac{1}{6}$$

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	X	2				
2	n	8				
3	p	1/6				
4	$P(X \leq 2)$	0,8652	=DISTR.BINOM.N(B1;B2;B3;VERDADERO)			

3.3)

$$P(X \geq 2) =?\ ; n = 8;\ p = \frac{1}{6}$$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - P(X \leq 1)$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	X	1				
2	n	8				
3	p	1/6				
4	$P(X \leq 1)$	0,6047	=DISTR.BINOM.N(B1;B2;B3;VERDADERO)			
5	$P(X \geq 2)$	0,3953	=1-B4			

4) Se lanzan simultáneamente tres monedas, calcular la probabilidad de que se obtengan:
4.1) Tres caras.
4.2) Dos caras y un sello
4.3) Una cara y dos sellos
4.4) Tres sellos
4.5) Al menos una cara

Solución:
Designando por C = cara y por S = sello se tiene:
Espacio muestral = S = {CCC, CCS, CSC, SCC, CSS, SCS, SSC, SSS}, entonces, n(S) = 8
Cada una de estos puntos muestrales son igualmente probables, con probabilidad de 1/8

Todas las probabilidades individuales se representan en la siguiente tabla:

Monedas			n(E)	P(E)
1ra	2da	3ra		
C	C	C	1	1/8
C	C	S	3	3/8
C	S	S	3	3/8
S	S	S	1	1/8
Total			8	1

4.1) Tres caras.
Observando la tabla se obtiene que P(CCC) = 1/8

Aplicando la fórmula se obtiene:

$$P(X=3)=P(CCC); n=3;\ p=\frac{1}{2}$$

$$P(X)=\frac{n!}{X!\,(n-X)!}\cdot p^X\cdot(1-p)^{n-X}$$

$$P(CCC)=\frac{3!}{3!\,(3-3)!}\cdot\frac{1^3}{2}\cdot\left(1-\frac{1}{2}\right)^{3-3}=1\cdot\frac{1}{8}\cdot 1=\frac{1}{8}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	3			
2	n	3			
3	p	1/2			
4	P(X=3)	1/8	=DISTR.BINOM.N(B1;B2;B3;FALSO)		

4.2) Dos caras y un sello
Observando la tabla se obtiene que P(CCS) = 3/8
Aplicando la fórmula se obtiene:

$$P(X=2)=P(CCS); n=2;\ p=\frac{1}{2}$$

$$P(X)=\frac{n!}{X!\,(n-X)!}\cdot p^X\cdot(1-p)^{n-X}$$

$$P(CCS)=\frac{3!}{2!\,(3-2)!}\cdot\frac{1^2}{2}\cdot\left(1-\frac{1}{2}\right)^{3-2}=3\cdot\frac{1}{4}\cdot\frac{1}{2}=\frac{3}{8}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	2			
2	n	3			
3	p	1/2			
4	P(X=2)	3/8	=DISTR.BINOM.N(B1;B2;B3;FALSO)		

4.3) Una cara y dos sellos
Observando la tabla se obtiene que P(CSS) = 3/8

Aplicando la fórmula se obtiene:

$$P(X=1)=P(CSS);n=1;\ p=\frac{1}{2}$$

$$P(X)=\frac{n!}{X!\,(n-X)!}\cdot p^{X}\cdot(1-p)^{n-X}$$

$$P(CSS)=\frac{3!}{1!\,(3-1)!}\cdot\frac{1}{2}^{1}\cdot\left(1-\frac{1}{2}\right)^{3-1}=3\cdot\frac{1}{2}\cdot\frac{1}{4}=\frac{3}{8}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	1			
2	n	3			
3	p	1/2			
4	P(X=1)	3/8	=DISTR.BINOM.N(B1;B2;B3;FALSO)		

4.4) Tres sellos
Observando la tabla se obtiene que P(SSS) = 1/8
Aplicando la fórmula se obtiene:

$$P(X=0)=P(SSS);n=3;\ p=\frac{1}{2}$$

$$P(X)=\frac{n!}{X!\,(n-X)!}\cdot p^{X}\cdot(1-p)^{n-X}$$

$$P(SSS)=\frac{3!}{0!\,(3-0)!}\cdot\frac{1}{2}^{0}\cdot\left(1-\frac{1}{2}\right)^{3-0}=1\cdot 1\cdot\frac{1}{8}=\frac{1}{8}$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	0			
2	n	3			
3	p	1/2			
4	P(X=0)	1/8	=DISTR.BINOM.N(B1;B2;B3;FALSO)		

4.5) Al menos una cara

Observando la tabla se obtiene que:

P(Al menos C) = P(CCC) + P(CCS) + P(CSS) = 1/8 +3/8 + 3/8 = 7/8

O también P(Al menos C) = 1 – P(SSS) = 1 – 1/8 = 7/8

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	X	0			
2	n	3			
3	p	1/2			
4	P(X=0)	1/8	=DISTR.BINOM.N(B1;B2;B3;FALSO)		
5	$P(X \geq 1)$	7/8	=1-B4		

TAREA DE INTERAPRENDIZAJE N° 7

1) Realice un organizador gráfico sobre la distribución binomial

2) Calcule de manera manual, empleando Excel y GeoGebra. Realice los gráficos en Winstats y GeoGebra

2.1) Para n = 4 y p = 0,12, ¿cuánto es P(X = 0)?

R: 0,5997

2.2) Para n = 10 y p = 0,40, ¿cuánto es P(X = 9)?

R: 0,0016

2.3) Para n =10 y p = 0,50, ¿cuánto es P(X = 8)?

R: 0,0439

3) En una muestra de 4 pedidos, se observa el siguiente resultado:

1er pedido	2do pedido	3er pedido	4to pedido
Marcado	Marcado	Sin marcar	Marcado

3.1) Llenar la tabla de manera manual y empleando Excel

n	p	X	$\frac{n!}{X!\,(n-X)!}$	p^X	$(1-p)^{n-X}$	$P(X)$
4	0,1	0				0,6561
		1				
		2				0,0486
		3				
		4				0,0001

Empleando la anterior tabla, resuelva los siguientes ejercicios de manera manual, empleando Excel, y GeoGebra.

3.2) ¿Qué probabilidad existe de que haya tres formatos marcados?

$P(X = 3) = 0{,}0036$

3.3) ¿Qué probabilidad existe de que haya menos de tres formatos marcados?

$P(X < 3) = 0{,}9963$

3.4) ¿Qué probabilidad existe de que haya más de tres formatos marcados?

$$P(X > 3) = 0{,}0001$$

3.5) ¿Qué probabilidad existe de que haya tres o más formatos marcados (es decir, al menos tres, por lo menos tres, o mínimo tres)?

$$P(X \geq 3) = 0{,}0037$$

3.6) ¿Qué probabilidad existe de que haya tres o menos formatos marcados? (es decir, a lo más tres)?

$$P(X \leq 3) = 0{,}9999$$

3.7) Calcular la desviación estándar

$$\sigma = 0{,}6$$

4) Crear y resolver de forma manual , empleando Excel y Winstats un problema similar al anterior.

5) El 60% de profesionales leen su contrato de trabajo, incluyendo las letras pequeñas. Suponga que el número de empleados que leen cada una de las palabras de su contrato se puede modelar utilizando la distribución binomial. Considerando un grupo de cinco empleados:

5.1) Llenar la tabla manera manual y empleando Excel

n	p	X	$\frac{n!}{X!\,(n-X)!}$	p^X	$(1-p)^{n-X}$	$P(X)$
		0				0,0102
		1				
		2				
		3				0,3456
		4				
		5				

5.2) Resuelva de manera manual y con GeoGebra la probabilidad de que:

a) Los cinco lean cada una de las palabras de su contrato

0,0778

b) Al menos tres lean cada una de las palabras de su contrato

0,6826

c) Menos de dos lean cada una de las palabras de su contrato

0,0870

6) ¿Cuáles serían los resultados para los incisos de la pegunta anterior, si la probabilidad de que un empleado lea cada una de las palabras de su contrato es de 0,80?. Resolver los siguientes ejercicios de de manera manual, empleando Excel y GeoGebra.

0,3277; 0,9421; 0,0067

7) Un examen de estadística de elección múltiple contenía 20 preguntas y cada una de ellas 5 respuestas. Si un estudiante desconocía todas las respuestas y contestó al azar, calcular de manera manual, empleando GeoGebra la probabilidad de que:

a) Conteste correctamente a 5 preguntas

0,1746

b) Conteste correctamente a lo más 5 preguntas

0,8042

8) Crear y resolver de manera manual, empleando Excel y Winstats un problema similar al anterior.

9) Se lanza simultáneamente 10 dados, calcular la probabilidad de manera manual, empleando Excel y Winstats de que se obtengan:

9.1) Exactamente 7 dos

0,00025

9.2) Exactamente 0 tres

0,16151

9.3) Menos de 7 cincos

0,99973

9.4) Más de 7 tres

0,00002

9.5) Por lo menos 7 cuatros

0,00027

10) Se lanzan simultáneamente cinco monedas, calcular la probabilidad de manera manual, empleando Excel y GeoGebra de que se obtengan:

10.1) Cinco caras

1/32

10.2) Tres caras y dos sellos

5/16

10.3) El mismo evento

0

10.4) Al menos una cara

31/32

11) Plantee y resuelva de manera manual, empleando Excel, GeoGebra y Winstats un ejercicio sobre dados y otro sobre monedas empleando la distribución binomial.

D) DISTRIBUCIÓN DE POISSON

i) Introducción.- Muchos estudios se basan en el conteo de las veces que se presenta un evento dentro de *un* área de oportunidad dada. El *área de oportunidad* es una unidad continua o intervalo de tiempo o espacio (volumen o área) en donde se puede presentar más de un evento. Algunos ejemplos serían los defectos en la superficie de un refrigerador, el número fallas de la red en un día, o el número de pulgas que tiene un perro. Cuando se tiene un área de oportunidad como éstas, se utiliza la *distribución de Poisson* para calcular las probabilidades si:

- Le interesa contar las veces que se presenta un evento en particular dentro de un área de oportunidad determinada. El área de oportunidad se define por tiempo, extensión, área, volumen, etc.
- La probabilidad de que un evento se presente en un área de oportunidad dada es igual para todas las áreas de oportunidad.
- El número de eventos que ocurren en un área de oportunidad es independiente del número de eventos que se presentan en cualquier otra área de oportunidad.
- La probabilidad de que dos o más eventos se presenten en un área de oportunidad tiende a cero conforme esa área se vuelve menor.

ii) Fórmula.- La distribución de Poisson tiene un parámetro, llamado λ (letra griega lambda minúscula), que es la media o el número esperado de eventos por unidad. La varianza de la distribución de Poisson también es igual a λ, y su desviación estándar es igual a $\sqrt{\lambda}$. El número de eventos X de la variable aleatoria de Poisson fluctúa desde 0 hasta infinito.

$$P(X) = \frac{e^{-\lambda} \cdot \lambda^{X}}{X!}$$

Donde:
$P(X) =$ Probabilidad de X eventos en un área de oportunidad
$\lambda =$ Número de eventos esperados
$X =$ Número de eventos
$e =$ Constante matemática base de los logaritmos naturales aproximadamente igual a 2718281828…. Este número es de gran importancia, tan sólo comparable a la del número π (pi), por su gran variedad de aplicaciones. El número e suele definirse como el límite de la expresión:

$$(1 + 1/\mathrm{n})^{\mathrm{n}}$$

Cuando *n* tiende hacia el infinito. Algunos valores de esta expresión para determinados valores de la *n* se muestran en la tabla siguiente:

VALOR NUMÉRICO DE $(1 + 1/\mathrm{n})^{\mathrm{n}}$ PARA VALORES CRECIENTES DE n		
n	$(1 + 1/n)^{n}$	Valor numérico
1	$(1 + 1/1)^{1}$	2
3	$(1 + 1/3)^{3}$	2,369
5	$(1 + 1/5)^{5}$	2,489
20	$(1 + 1/20)^{20}$	2,653
40	$(1 + 1/40)^{40}$	2,684
50	$(1 + 1/50)^{50}$	2,691
100	$(1 + 1/100)^{100}$	2,705
1000	$(1 + 1/1000)^{1000}$	2,717
10000	$(1 + 1/10000)^{10000}$	2,718
∞	……………………..	2,71828….

Observando la columna de la derecha de la tabla anterior, se puede ver que a medida que *n* crece el valor de la expresión se aproxima, cada vez más, a un valor límite. Este límite es 2,7182818285….

Ejemplos ilustrativos

1) Suponga una distribución de Poisson. Si $\lambda = 1$, calcular P(X= 0)

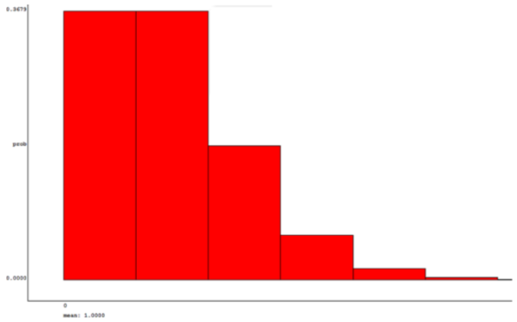
Solución:

Aplicando la fórmula se obtiene:

$$P(X) = \frac{e^{-\lambda} \cdot \lambda^{X}}{X!} = \frac{2{,}71828^{-1} \cdot 1^{0}}{0!} = 0{,}3679$$

También se puede obtener con lectura de la tabla de probabilidades de Poisson

TABLA Nº 2
PROBABILIDADES DE POISSON

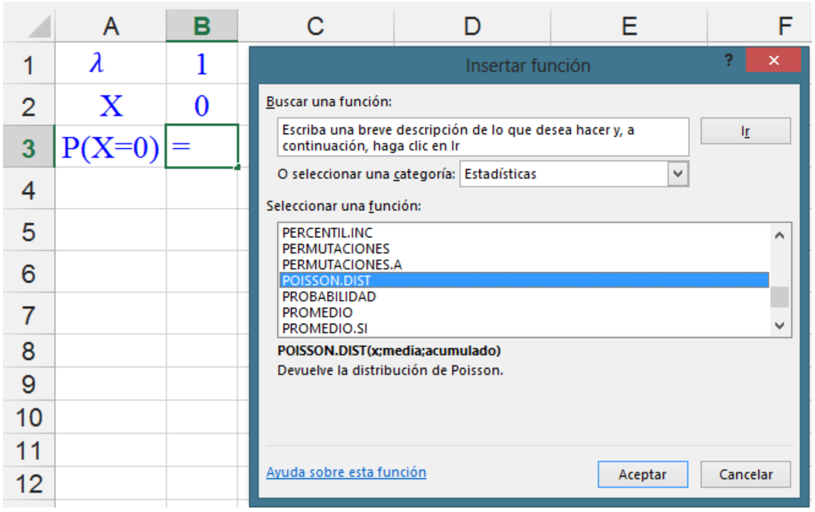


$\boldsymbol{Ejemplo: Para\ \lambda = 1\ y\ X = 0\ \Rightarrow P(X = 0) = 0{,}3679}$

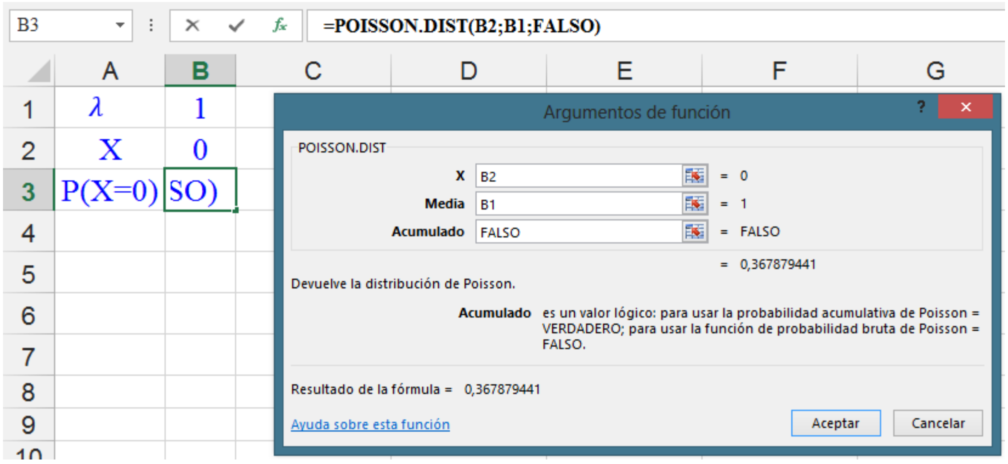
	λ									
X	**0,005**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0	**0,9950**	**0,9900**	**0,9802**	**0,9704**	**0,9608**	**0,9512**	**0,9418**	**0,9324**	**0,9231**	**0,9139**
1	**0,0050**	**0,0099**	**0,0196**	**0,0291**	**0,0384**	**0,0476**	**0,0565**	**0,0653**	**0,0738**	**0,0823**
2	**0,0000**	**0,0000**	**0,0002**	**0,0004**	**0,0008**	**0,0012**	**0,0017**	**0,0023**	**0,0030**	**0,0037**
3	**0,0000**	**0,0000**	**0,0000**	**0,0000**	**0,0000**	**0,0000**	**0,0000**	**0,0001**	**0,0001**	**0,0001**
	λ									
X	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**	**1**
0	**0,9048**	**0,8187**	**0,7408**	**0,6703**	**0,6065**	**0,5488**	**0,4966**	**0,4493**	**0,4066**	**0,3679**

El cálculo de $P(X = 0)$ con $\lambda = 1$ en Excel se realizan de la siguiente manera:

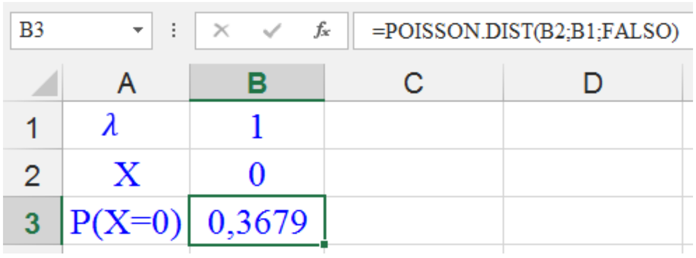
a) Se inserta la función POISSON.DIST



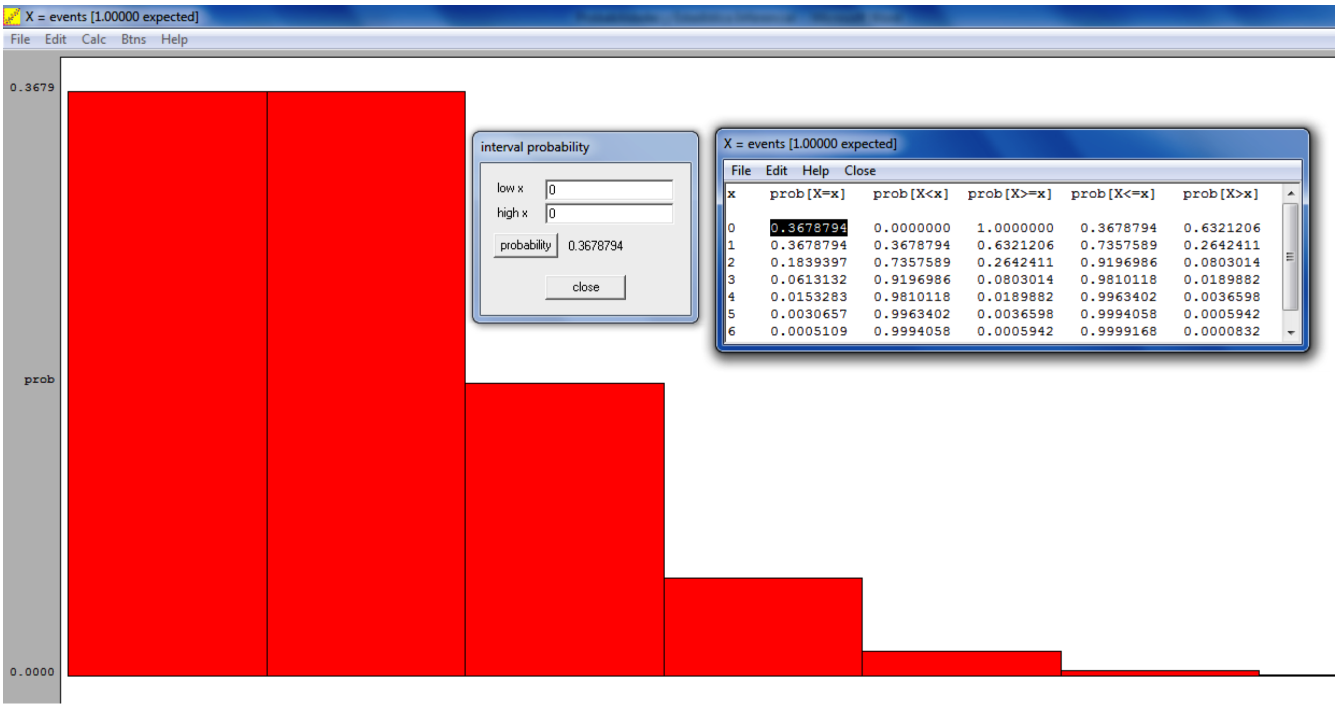
b) Clic en Aceptar. En la ventana de Argumentos de la función, en X seleccionar B2 en Media escribir o seleccionar B1 y en Acumulado escribir Falso.



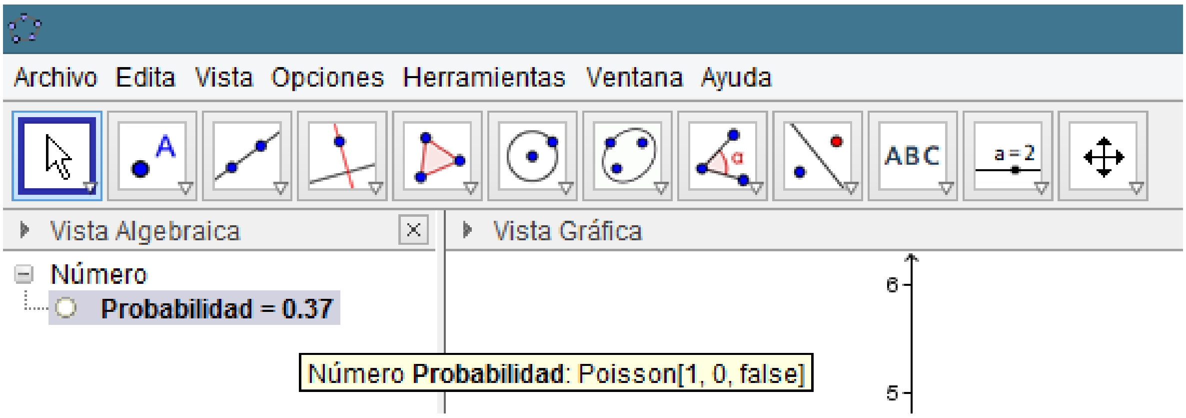
c) Clic en Aceptar



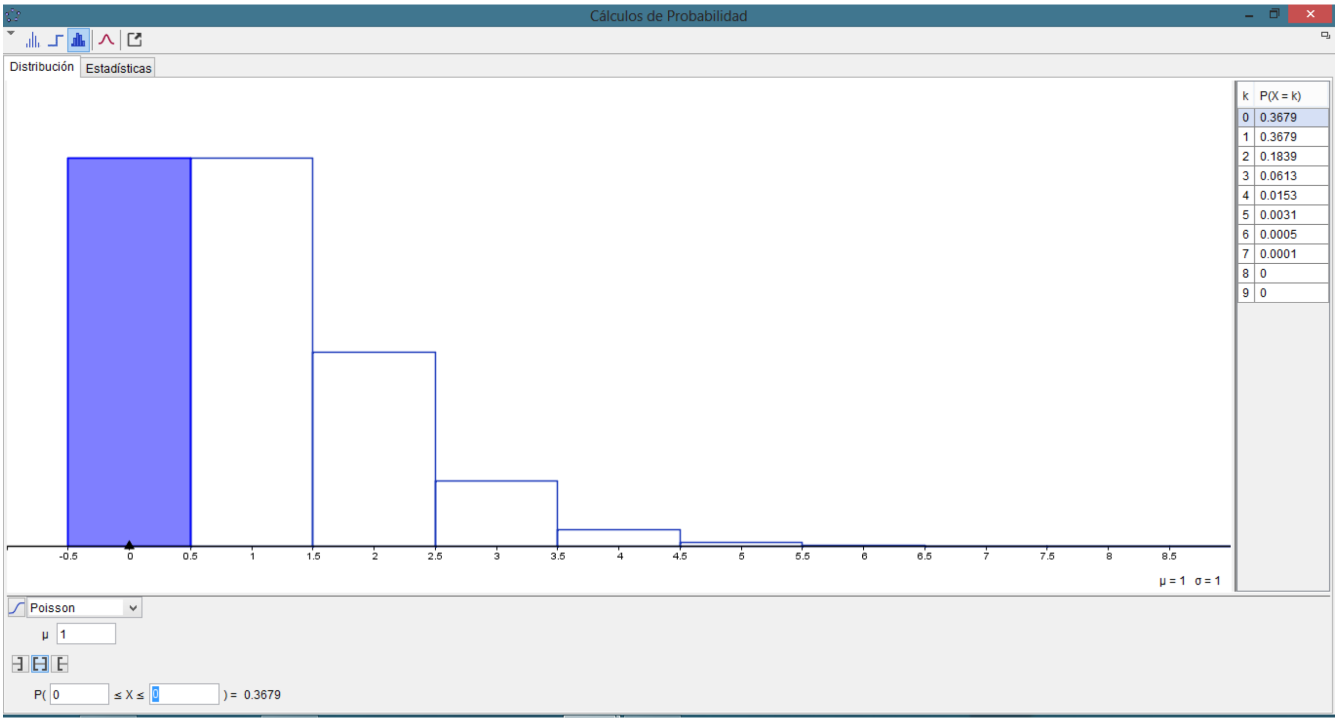
Los cálculos en Winstats se muestran en la siguiente figura:

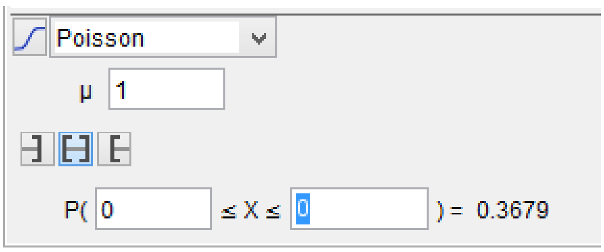


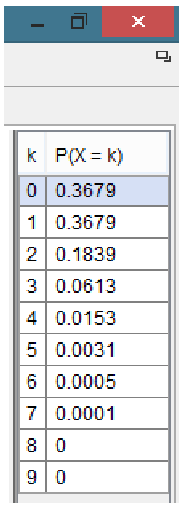
Los cálculos en GeoGebra se muestran en la siguiente figura:



Nota:
Escoger la opción Poisson[<Media>, <Valor de Variable>, <Acumulativa Booleana>]
Escribir 1 en <Media>, 0 en <Valor de Variable>, false en <Acumulativa Booleana>
Para P(X= n), siendo n el número de eventos o ensayos, en <Acumulada Booleana> se escribe false
Para P(X≤ n), siendo n el número de eventos o ensayos, en <Acumulada Booleana> se escribe true







2) Suponga una distribución con $\lambda = 5$. Determine $P(X \geq 10)$

Solución:
$P(X \geq 10) = 1 - P(X \leq 9)$
$P(X \leq 9) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + \cdots + P(X = 9)$

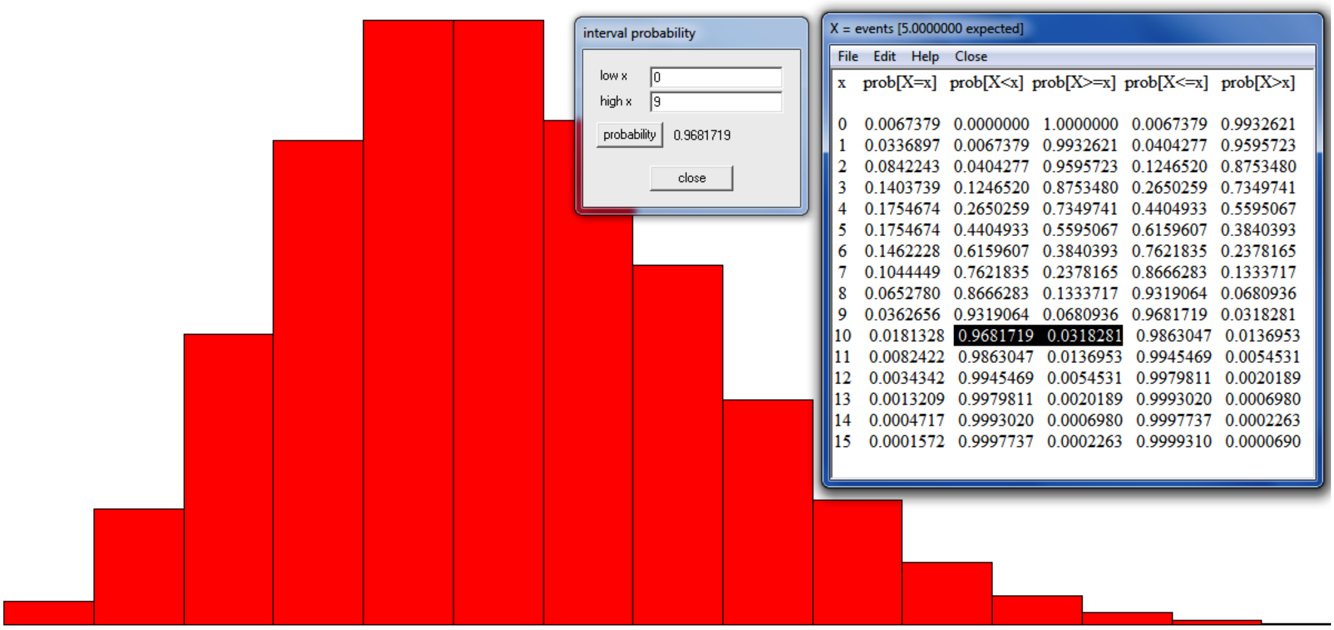
Aplicando la fórmula o con lectura en la tabla de la distribución de Poisson se obtiene:
$P(X \leq 9) =$ 0,0067 + 0,0337 + 0,0842 + 0,1404 + 0,1755 + 0,1755 +0,1462 + 0,1044 +0,0653 + 0,0363
$P(X \leq 9) = 0{,}9682$

Entonces:
$P(X \geq 10) = 1 - P(X \leq 9) = 1 - 0{,}9682 = 0{,}0318$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	λ	5			
2	X	9			
3	P(X≤ 9)	0,9681719	=POISSON(B2;B1;VERDADERO)		
4	P(X ≥10)	0,0318281	=1-B3		

Los cálculos en Winstats se muestran en la siguiente figura:



x	prob[X=x]	prob[X<x]	prob[X>=x]	prob[X<=x]	prob[X>x]
0	0.0067379	0.0000000	1.0000000	0.0067379	0.9932621
1	0.0336897	0.0067379	0.9932621	0.0404277	0.9595723
2	0.0842243	0.0404277	0.9595723	0.1246520	0.8753480
3	0.1403739	0.1246520	0.8753480	0.2650259	0.7349741
4	0.1754674	0.2650259	0.7349741	0.4404933	0.5595067
5	0.1754674	0.4404933	0.5595067	0.6159607	0.3840393
6	0.1462228	0.6159607	0.3840393	0.7621835	0.2378165
7	0.1044449	0.7621835	0.2378165	0.8666283	0.1333717
8	0.0652780	0.8666283	0.1333717	0.9319064	0.0680936
9	0.0362656	0.9319064	0.0680936	0.9681719	0.0318281
10	0.0181328	0.9681719	0.0318281	0.9863047	0.0136953
11	0.0082422	0.9863047	0.0136953	0.9945469	0.0054531
12	0.0034342	0.9945469	0.0054531	0.9979811	0.0020189
13	0.0013209	0.9979811	0.0020189	0.9993020	0.0006980
14	0.0004717	0.9993020	0.0006980	0.9997737	0.0002263
15	0.0001572	0.9997737	0.0002263	0.9999310	0.0000690

TAREA DE INTERAPRENDIZAJE N° 8

1) Defina con sus propias palabras:
1.1) Área de oportunidad
1.2) Distribución de Poisson
1.3) Número *e*

2) Escriba 5 ejemplos de área de oportunidad

3) Investigue sobre la biografía de Siméon Denis Poisson y realice un organizador gráfico de la misma.

4) Calcule empleando la fórmula de la distribución de Poisson, Excel , GeoGebra y Winstats

4.1) $P(X = 8)\ si\ \lambda = 8$

0,1396

4.2) $P(X = 1)\ si\ \lambda = 0,5$

0,3033

4.3) $P(X = 0)\ si\ \lambda = 3,7$

0,0247

5) Calcule empleando la tabla y Excel. Suponga una distribución con $\lambda = 5$.

5.1) $P(X = 1)$

0,0337

5.2) $P(X < 1)$

0,0067

5.3) $P(X > 1)$

0,9596

5.4) $P(X \leq 1)$

0,0404

6) Suponga que la media de clientes que llega a un banco por minuto durante la hora que va del mediodía a la 1 pm es igual a 3. Calcular empleando la tabla y Winstats.

6.1) ¿Cuál es la probabilidad de que lleguen exactamente dos clientes durante un minuto?

0,2240

6.2) ¿Y cuál es la probabilidad de que lleguen más de dos clientes durante un minuto dado?

0,5768

7) El gerente de control de calidad de una empresa que elabora galletas inspecciona un lote de galletas con chispas de chocolate que se acaban de preparar. Si el proceso de producción está bajo control, la media de chispas de chocolate por galleta es 6. Calcular empleando la tabla y Excel. ¿Cuál es la probabilidad de que en cualquier galleta inspeccionada.

7.1) Se encuentre menos de cinco chispas?

0,2851

7.2) Se encuentren exactamente cinco chispas?

0,1606

7.3) Se encuentren cinco o más chispas?

0,7149

8) El departamento de transporte registra las estadísticas de las maletas maltratadas por cada 1000 pasajeros. En 2003, una empresa de transporte tuvo 3,21 maletas maltratadas por cada 1000 pasajeros. Calcular empleando la fórmula y GeoGebra. ¿Cuál es la probabilidad de que, en los próximos 1000 pasajeros, aquella empresa tenga

8.1) Ninguna maleta maltratada?

0,0404

8.2) Al menos una maleta maltratada?

0,9596

8.3) Al menos dos maletas maltratadas?

0,8301

9) Plantee y resuelva 3 ejercicios de aplicación sobre la distribución de Poisson de manera manual, empleando Excel, GoeGebra y Winstats.

E) DISTRIBUCIÓN HIPERGEOMÉTRICA

i) Definición

La distribución binomial es apropiada sólo si la probabilidad de un éxito permanece constante. Esto ocurre si el muestreo se realiza con reemplazo en una población grande. Sin embrago, si la población es pequeña y ocurre sin reemplazo, la probabilidad de éxito variará, y la distribución hipergeométrica es que se utiliza.

ii) Fórmula

Se calcula empleando la siguiente fórmula:

$$P(X) = \frac{C_X^r \cdot C_{n-X}^{N-r}}{C_n^N}$$

Donde:
C = combinación
N = tamaño de la población
r = número de éxitos en la población
n = tamaño de la muestra
X = número de éxitos en la muestra

Notas:

- Si se selecciona una muestra sin reemplazo de una población grande conocida y contiene una proporción relativamente grande de la población, de manera que la probabilidad de éxito varía de una selección a la siguiente, debe utilizarse la distribución hipergeométrica.

- Cuando tamaño de la población (N) es muy grande, la distribución hipergeométrica tiende aproximarse a la binomial.

Ejemplo ilustrativo

Si se extraen juntas al azar 3 bolas de una urna que contiene 6 bolas rojas y 4 blancas. ¿Cuál es la probabilidad de que sean extraídas 2 bolas rojas?.

Solución:

Los datos son:
N =10; r = 6; n = 3 y X= 2

Aplicando la fórmula se obtiene:

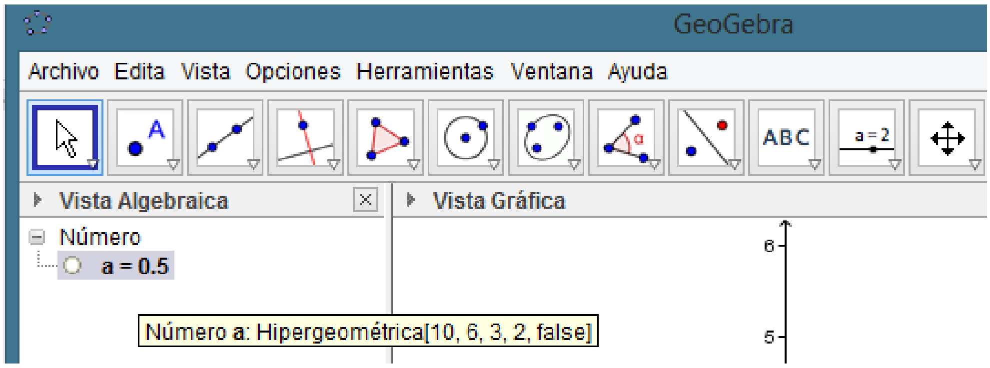
$$P(X) = \frac{C_X^r \cdot C_{n-X}^{N-r}}{C_X^N}$$

$$P(X=2) = \frac{C_2^6 \cdot C_{3-2}^{10-6}}{C_3^{10}} = \frac{C_2^6 \cdot C_1^4}{C_3^{10}} = \frac{\frac{6!}{2!\,(6-2)!} \cdot \frac{4!}{1!\,(4-1)!}}{\frac{10!}{3!\,(10-3)!}} = \frac{\frac{6!}{2!\,4!} \cdot \frac{4!}{1!\,3!}}{\frac{10!}{3!\,7!}} = \frac{15 \cdot 4}{120} = 0{,}5$$

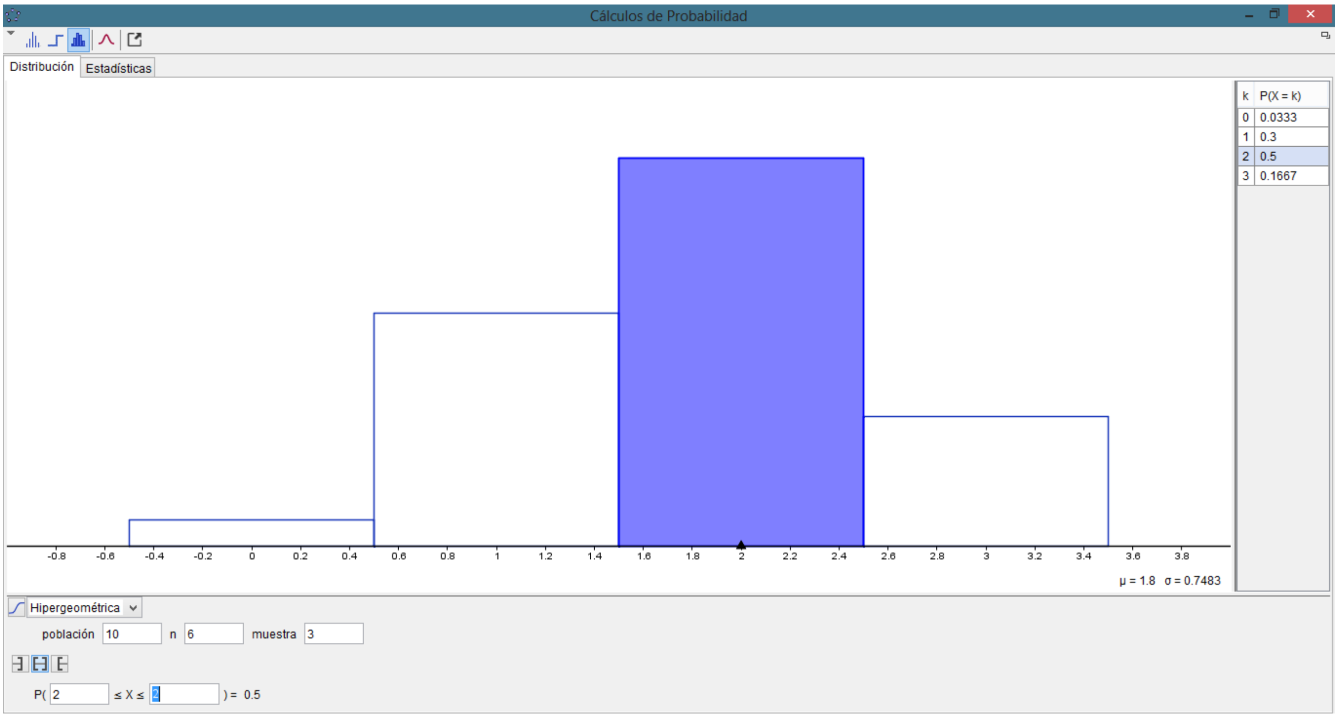
El cálculo de P(X=2) en Excel se muestra en la siguiente figura:

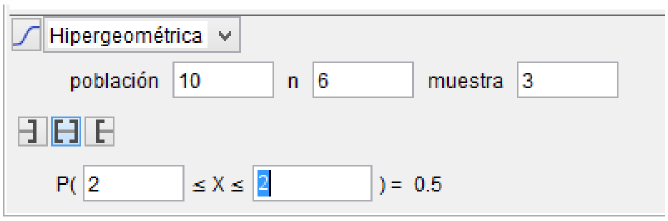
	A	B	C	D	E	F
1	N	10				
2	r	6				
3	n	3				
4	X	2				
5	P(X=2)	0,5	=DISTR.HIPERGEOM.N(B4;B3;B2;B1;FALSO)			

El cálculo de P(X=2) en GeoGebra se muestra en la siguiente figura:



Hipergeométrica[<Tamaño de Población>, <Número de Éxitos>, <Tamaño de Muestra>, <Valor de Variable>, <Acumulativa Booleana>]





TAREA DE INTERAPRENDIZAJE N° 9

1) ¿En qué se diferencia la distribución binomial con la distribución hipergeométrica?

2) Realice un organizador gráfico de la distribución hipergeométrica.

3) Resolver los siguientes ejercicios de manera manual, empleando Excel y GeoGebra

3.1) En un local de venta de automóviles existen 20 vehículos de los cuales 8 son de la preferencia de Mathías. Si Mathías selecciona 3 vehículos al azar, ¿cuál es la probabilidad de que un vehículo sea de su preferencia?.

0,4632

3.2) En un aula de 40 estudiantes hay 16 hombres. ¿Cuál es la probabilidad de seleccionar una muestra de 12 en la cual 8 sean hombres?

0,0245

3.3) De un grupo de 9 personas 4 son mujeres. ¿Cuál es la probabilidad de seleccionar una muestra de 3 personas en la cual no más que una mujer sea seleccionada?

0,5952

3.4) De 100 establecimientos educativos, 70 disponen de canchas deportivas. Si se pregunta si disponen de canchas deportivas a una muestra aleatoria de 20 establecimientos educativos, calcular la probabilidad de que:

a) Exactamente 8 dispongan de canchas deportivas

0,00152

b) Exactamente 8 no dispongan de canchas deportivas

0,11618

3.5) De 40 estudiantes de una clase de Estadística, a 30 les gusta la asignatura. Si se pregunta por la preferencia a esta asignatura a una muestra aleatoria de 8 estudiantes, calcular la probabilidad de que:

a) Exactamente a 6 les gusta la asignatura

0,34744

b) Mínimo a 6 les gusta la asignatura

0,68826

c) Exactamente a 6 no les gusta la asignatura

0,00119

d) Mínimo a 6 no les gusta la asignatura

0,00124

4) Plantee y resuelva 3 ejercicios de aplicación sobre la distribución hipergeométrica de manera manual, empleando Excel y GeoGebra

2.2 DISTRIBUCIONES CONTINUAS

A) INTRODUCCIÓN

Una distribución de probabilidad es continua cuando los resultados posibles del experimento son obtenidos de variables aleatorias continuas, es decir, de variables cuantitativas que pueden tomar cualquier valor, y que resultan principalmente del proceso de medición.

Ejemplos de variables aleatorias continuas son:

La estatura de un grupo de personas
El tiempo dedicado a estudiar
La temperatura en una ciudad

B) DISTRIBUCIÓN EXPONENCIAL

i) Definición

La distribución de Poisson calcula el número de eventos sobre alguna área de oportunidad (intervalo de tiempo o espacio), la distribución exponencial mide el paso del tiempo entre tales eventos. Si el número de eventos tiene una distribución de Poisson, el lapso entre los eventos estará distribuido exponencialmente.

ii) Fórmula

La probabilidad de que el lapso de tiempo sea menor que o igual a cierta cantidad x es:

$$P(X \leq x) = 1 - e^{-\lambda \cdot t}$$

Donde:
t =Lapso de tiempo
e = Base del logaritmo natural aproximadamente igual a 2,718281828
λ =Tasa promedio de ocurrencia

Ejemplo ilustrativo

Los buses interprovinciales llegan al terminal a una tasa promedio de 10 buses por hora.

1) ¿Cuál es la probabilidad de que llegue un bus en no más de 5 minutos?
2) ¿Cuál es la probabilidad de que llegue un bus en no más de 10 minutos?
3) ¿Cuál es la probabilidad de que llegue un bus entre 5 minutos y 10 minutos?
4) ¿Cuál es la probabilidad de que llegue un bus en más de 5 minutos?

Solución:

λ = 10 por una hora

1) Como la tasa promedio está dada por hora, y el problema se plantea en minutos, se calcula el porcentaje que representa 5 minutos de una hora (60 minutos), el cual es:

$$\frac{5}{60} = \frac{1}{12} = 0{,}0833$$

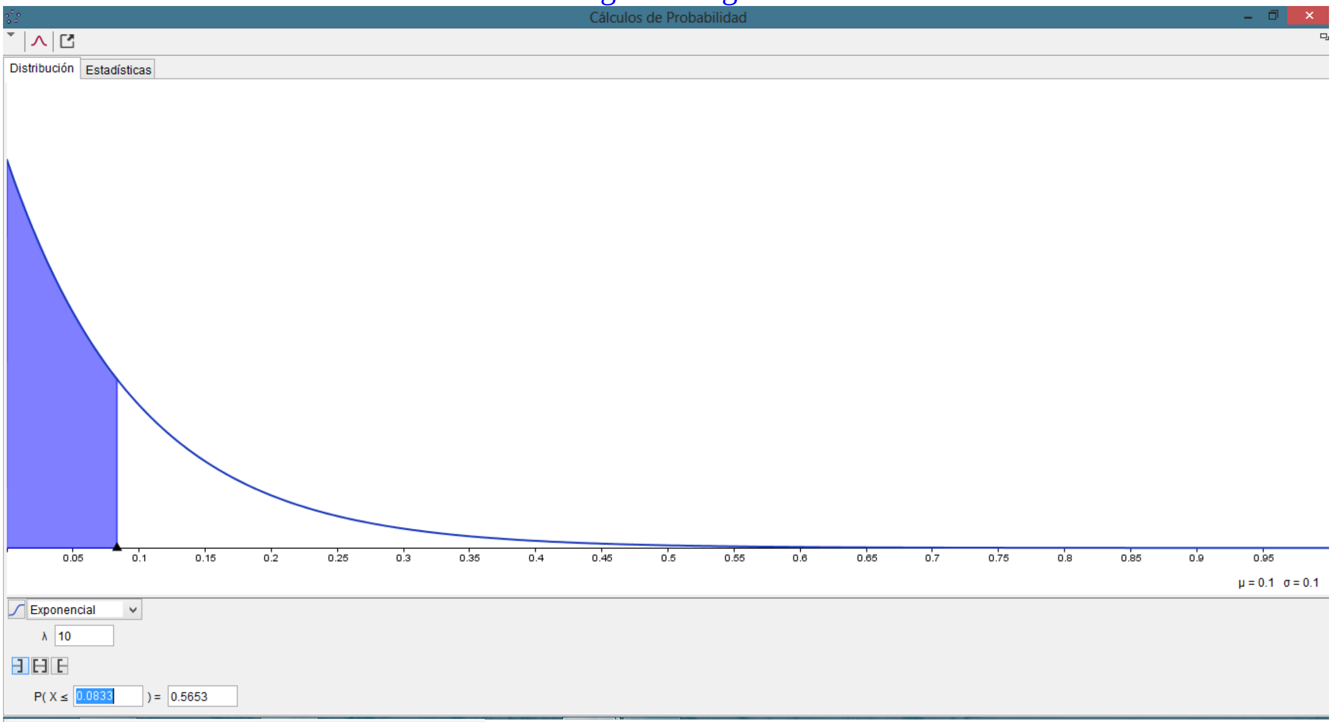
Reemplazado valores de la fórmula se obtiene:

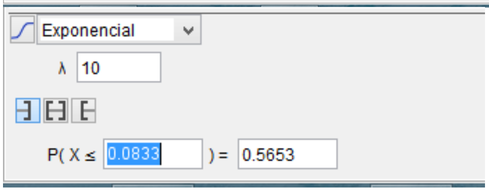
$$P(X \leq x) = 1 - e^{-\lambda \cdot t}$$

$$P(X \leq 5) = 1 - e^{-10 \cdot \frac{1}{12}} = 0{,}5654$$

Interpretación: Existe un 56,54% de probabilidad de que el segundo bus llegue al terminal en 5 minutos o menos del primero si la tasa promedio de llegada es de 10 buses por hora.

Los cálculos en GeoGebra se muestran en la siguiente figura





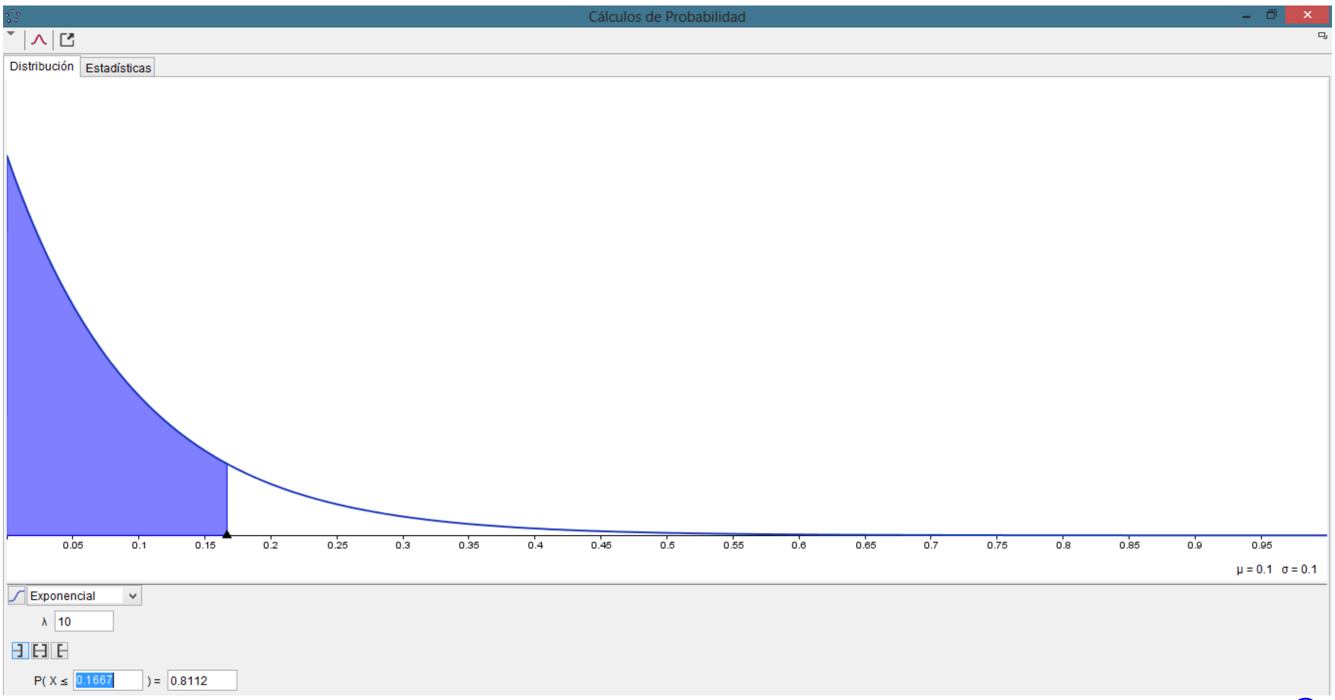
2) El porcentaje que representa 10 minutos de una hora (60 minutos) es:

$$\frac{10\ min}{60\ min} = \frac{1}{6} = 0{,}1667$$

Remplazado valores de la fórmula se obtiene: $P(X \leq x) = 1 - e^{-\lambda \cdot t}$

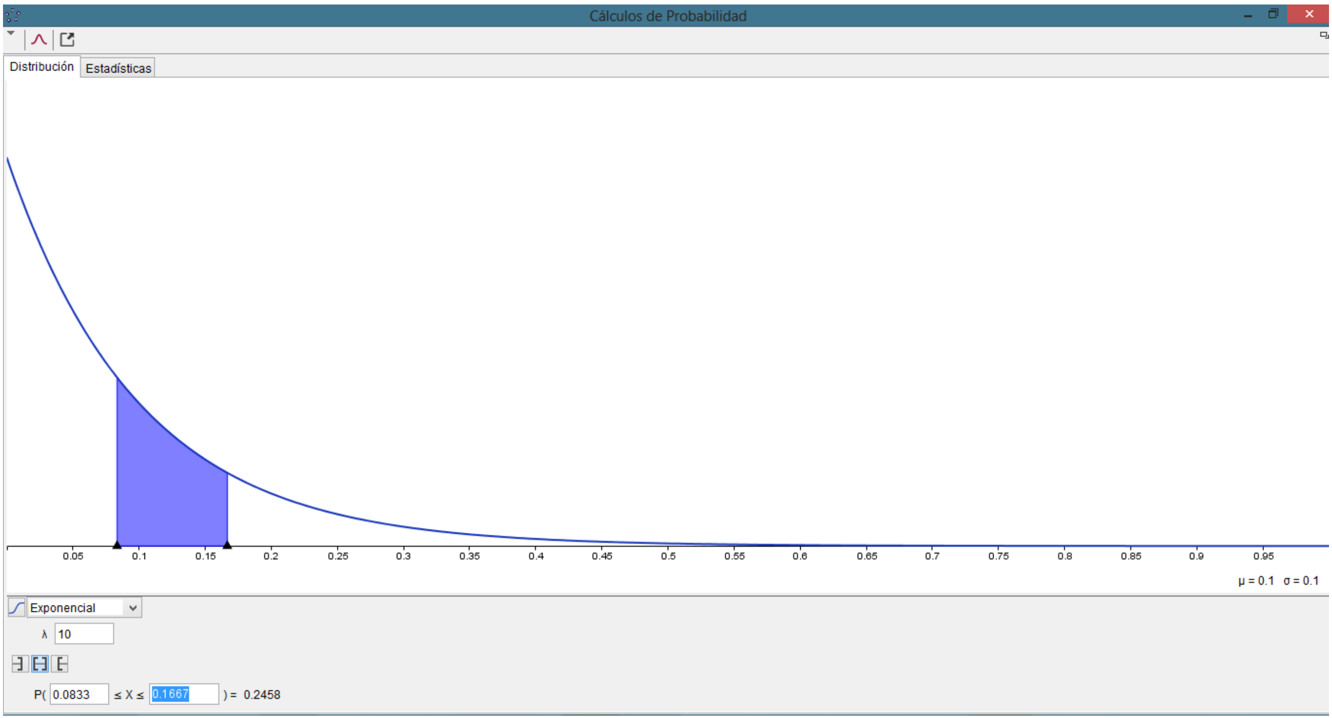
$$P(X \leq 10) = 1 - e^{-10 \cdot \frac{1}{6}} = 0{,}8111$$

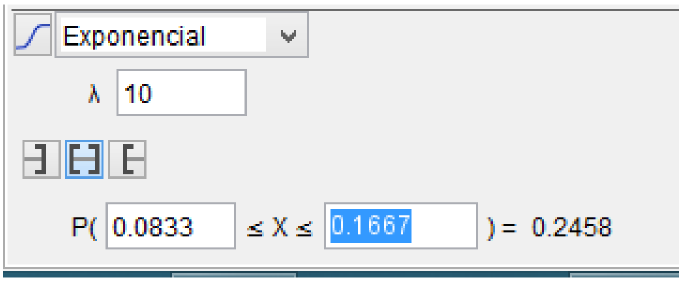
En GeoGebra:



3) P(5≤X≤10) = P(X≤10) – P(X≤5)

P(5≤X≤10) = 0,8111- 0,5654 = 0,2457





4) P(X>5) = 1 – P(X≤5)
P(X>5) = 1 – 0,5654 = 0,4346

En los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	t	1/12	=5/60		
2	λ	10			
3	P(X≤5)	0,5654	=DISTR.EXP(B1;B2;VERDADERO)		
4					
5	t	1/6	=10/60		
6	λ	10			
7	P(X≤10)	0,8111	=DISTR.EXP(B5;B6;VERDADERO)		
8					
9	P(5≤X≤10)	0,2457	=B7-B3		
10					
11	P(X>5)	0,4346	=1-B3		

TAREA DE INTERAPRENDIZAJE N° 10

1) Demuestre que $P(X > x) = e^{-\lambda \cdot t}$

2) Resuelva los siguientes ejercicios de aplicación empleando la fórmula, mediante Excel y GeoGebra.

2.1) La tasa promedio de llegada de clientes que compran en un local comercial es de 1,5 por hora. Calcular la probabilidad de que no más de dos horas transcurran entre llegadas de los mencionados clientes?

0,9502

2.2) Los barcos llegan a un puerto en una tasa promedio de 8 por hora. ¿Cuál es la probabilidad de que transcurran más de 15 minutos entre la llegada de 2 barcos?

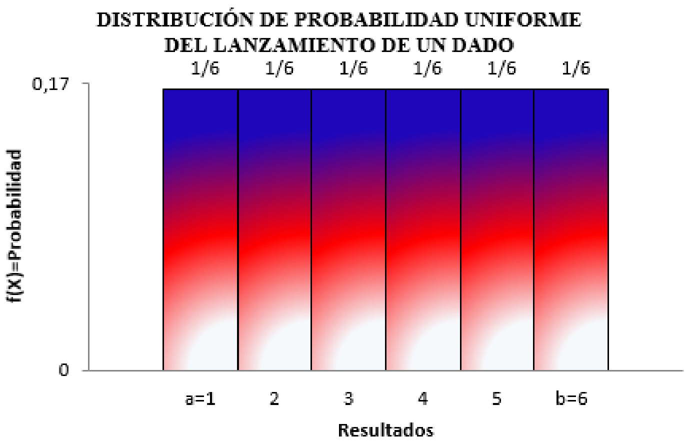
0,1353

2.3) Plantee y resuelva dos ejercicios de aplicación similares a los anteriores empleando datos reales sobre cualquier tema de su preferencia.

2.4) Plantee y resuelva un ejercicio de aplicación similar al ejemplo ilustrativo de la distribución exponencial empleando datos reales sobre cualquier tema de su preferencia.

C) DISTRIBUCIÓN UNIFORME

i) Definición.- Es una distribución en el intervalo $[a, b]$ en la cual las probabilidades son las mismas para todos los posibles resultados, desde el mínimo de **a** hasta el máximo de **b**. El experimento de lanzar un dado es un ejemplo que cumple la distribución uniforme, ya que todos los 6 resultados posibles tienen 1/6 de probabilidad de ocurrencia.



ii) Función de densidad de una distribución uniforme (altura de cada rectángulo en la gráfica anterior) es:

$$\mathrm{f(X)} = \mathrm{Altura} = \frac{1}{\mathrm{b} - \mathrm{a}}$$

Donde:
a = mínimo valor de la distribución
b = máximo valor de la distribución
b – a = Rango de la distribución

iii) La media, valor medio esperado o esperanza matemática de una distribución uniforme se calcula empleando la siguiente fórmula:

$$\mathrm{E(X)} = \mu = \frac{\mathrm{a} + \mathrm{b}}{2}$$

iv) La varianza de una distribución uniforme se calcula empleando la siguiente fórmula:

$$\sigma^2 = \frac{(b-a)^2}{12}$$

De donde la desviación estándar es $\sigma = \sqrt{\sigma^2}$

v) La probabilidad de que una observación caiga entre dos valores se calcula de la siguiente manera:

$$P(X_1 \leq X \leq X_2) = \frac{X_2 - X_1}{b - a}$$

Ejemplo ilustrativo

Sea X el momento elegido al azar en que un estudiante recibe clases en un determinado día entre las siguientes horas: 7:00 - 8:00 - 9:00 - 10:00 - 11:00 - 12:00 - 13:00

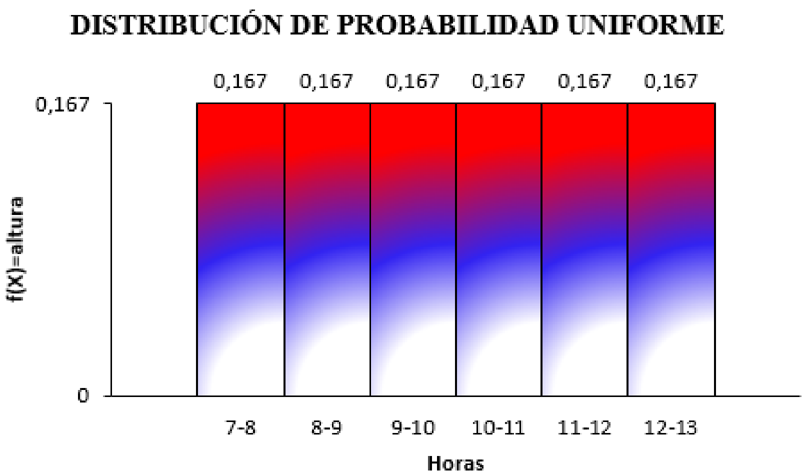
1) ¿Cuál es la función de densidad de la variable X?
2) Elaborar un gráfico de la distribución de probabilidades
3) Calcular el valor medio esperado
4) Calcular la desviación estándar
5) Calcular la probabilidad de que llegue en la primera media hora
6) Si recibe clases de Estadística Aplicada de 10:00 a 12:15, calcular la probabilidad de recibir esta asignatura.

Solución:
1) a = 7 y b = 13
Reemplazando valores en la ecuación de la función de densidad se obtiene:

$$f(X) = \text{Altura} = \frac{1}{b-a}$$

$$f(X) = \text{Altura} = \frac{1}{13-7} = \frac{1}{6} = 0{,}167$$

2) Elaborando el gráfico de la distribución de probabilidad empleando Excel se obtiene:



Interpretación:

Cada rectángulo tiene 1 de base y 1/6 = 0,167 de altura.

El área de cada rectángulo es:

$$A\square = base \cdot altura = 1 \cdot \frac{1}{6} = \frac{1}{6}$$

El área total (rectángulo de base el intervalo 7-13 y altura 1/6=0,167) representa a la suma de todas las probabilidades, y es igual a uno:

$$A\square = base \cdot altura = (13-7) \cdot \frac{1}{6} = 6 \cdot \frac{1}{6} = 1$$

3) Reemplazando valores en la fórmula del valor esperado se obtiene:

$$\mathrm{E(X)} = \mu = \frac{\mathrm{a+b}}{2}$$

$$\mathrm{E(X)} = \mu = \frac{7+13}{2} = 10$$

4) Reemplazando valores en la fórmula de la varianza se obtiene:

$$\sigma^2 = \frac{(\mathrm{b-a})^2}{12}$$

$$\sigma^2 = \frac{(13-7)^2}{12} = \frac{(6)^2}{12} = \frac{36}{12} = 3$$

Por lo tanto la desviación estándar es: $\sigma = \sqrt{\sigma^2} = \sqrt{3} = 1{,}732$

5) Llegar en la primera media hora significa que llega a la 7:30. Por lo tanto se debe calcular la probabilidad entre las 7:00 y las 7:30.

Como 7:30 = 7 horas + 30 minutos, y el porcentaje que representa 30 minutos de una hora es:

$$\frac{30}{60} = 0{,}5 \Rightarrow 7{:}30 = 7{,}5\ horas$$

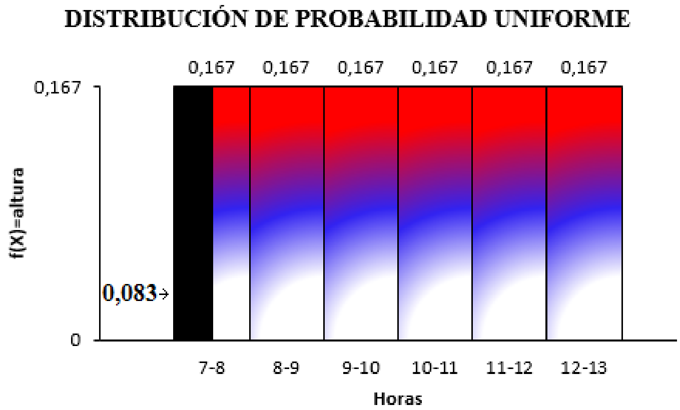
Por lo tanto se debe calcular la probabilidad entre 7 y 7,5

Aplicando la fórmula de la probabilidad entre dos valores se obtiene:

$$P(X_1 \leq X \leq X_2) = \frac{X_2 - X_1}{b-a}$$

$$P(7 \leq X \leq 7{,}5) = \frac{7{,}5-7}{13-7} = \frac{0{,}5}{6} = 0{,}0833$$

En el siguiente gráfico se muestra la probabilidad calculada:



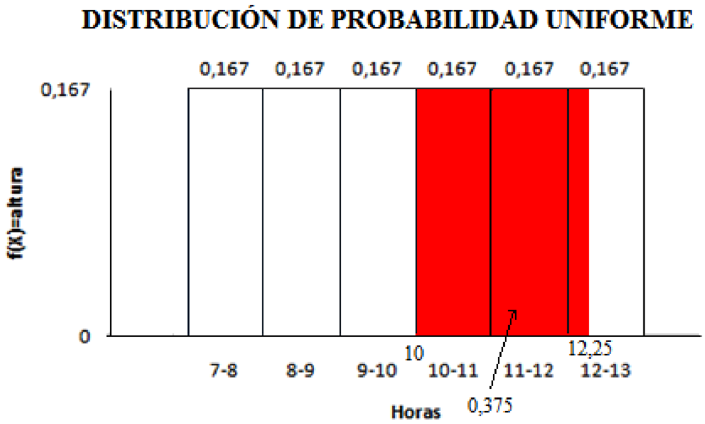
6) Se debe calcular la probabilidad entre las 10:00 y las 12:15

Como 12:15 = 12horas + 15 minutos, y el porcentaje que representa 15 minutos de una hora es:
$$\frac{15}{60} = 0{,}25 \Rightarrow 12{:}15 = 12{,}25\ horas$$

Por lo tanto de debe calcular la probabilidad entre 10 y 12,25

Aplicando la fórmula de la probabilidad entre dos valores se obtiene:
$$P(X_1 \le X \le X_2) = \frac{X_2 - X_1}{b - a}$$
$$P(10 \le X \le 12{,}25) = \frac{12{,}25 - 10}{13 - 7} = \frac{2{,}25}{6} = 0{,}375$$

En el siguiente gráfico se muestra la probabilidad calculada:



Los cálculos en Excel se muestran en la siguiente figura:

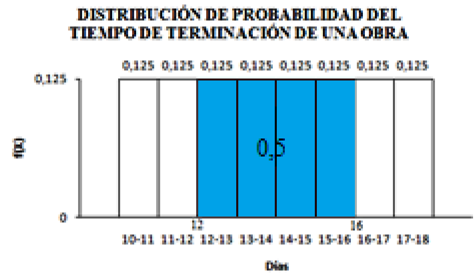
	A	B	C	D	E
1	7		a	7	=MIN(A1:A7)
2	8		b	13	=MAX(A1:A7)
3	9				
4	10		$\text{Altura} = \frac{1}{b-a}$	0,167	=1/(D2-D1)
5	11				
6	12		$\mu = \frac{a+b}{2}$	10	=PROMEDIO(D1:D2)
7	13				
8			$\sigma^2 = \frac{(b-a)^2}{12}$	3	=(D2-D1)^2/12
9					
10			$\sigma = \sqrt{\sigma^2}$	1,732	=RAIZ(D8)
11					
12			X_1	7	
13			X_2	7,5	
14	$P(X_1 \le X \le X_2) = \frac{X_2 - X_1}{b - a}$			0,083	=(D13-D12)/(D2-D1)
15					
16					
17			X_1	10	
18			X_2	12,25	
19	$P(X_1 \le X \le X_2) = \frac{X_2 - X_1}{b - a}$			0,375	=(D18-D17)/(D2-D1)
20					

TAREA DE INTERAPRENDIZAJE N° 11

1) Realice un organizador gráfico de la distribución uniforme

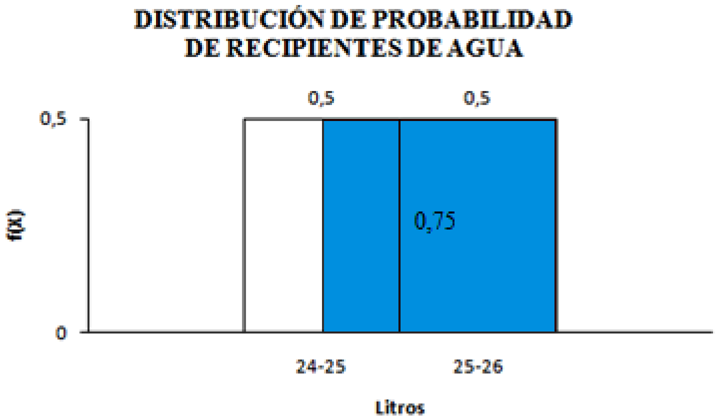
2) Los tiempos de terminación de una obra varían entre 10 días y 18 días. ¿Cuál es la probabilidad de que se requiera entre 12 y 16 días para realizar la mencionada obra?. Realice un gráfico que ilustre la probabilidad calculada. Resuelva el ejercicio de manera manual y empleando Excel.

0,5



3) Ciertos recipientes contienen agua con un volumen uniformemente distribuido de media igual a 25 litros y un rango de 2 litros. Calcule la probabilidad de seleccionar un recipiente que contenga entre 24,5 y 26 litros. Realice un gráfico que ilustre la probabilidad calculada. Resuelva el ejercicio de manera manual y empleando Excel.

0,75

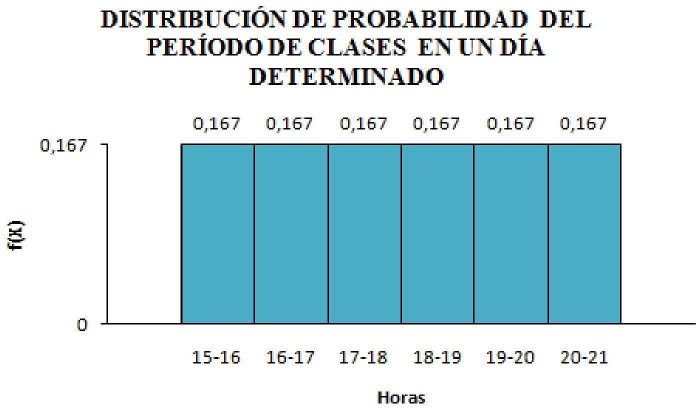


4) Sea X el momento elegido al azar en que un estudiante recibe clases en un determinado día entre las siguientes horas: 15:00 - 16:00 - 17:00 - 18:00 - 19:00 - 20:00 - 21:00. Resuelva el ejercicio de manera manual y empleando Excel.

4.1) ¿Cuál es la función de densidad de la variable X?

0,167

4.2) Elabore un gráfico de la distribución de probabilidades



4.3) Calcule el valor medio esperado

18

4.4) Calcule la desviación estándar

1,732

4.5) Calcule la probabilidad de que llegue en los primeros 15 minutos. Realice un gráfico que ilustre la probabilidad calculada.

0,042

4.6) Si recibe clases de Estadística Aplicada de 19:30 a 21:00, calcular la probabilidad de recibir esta asignatura. Realice un gráfico que ilustre la probabilidad calculada.

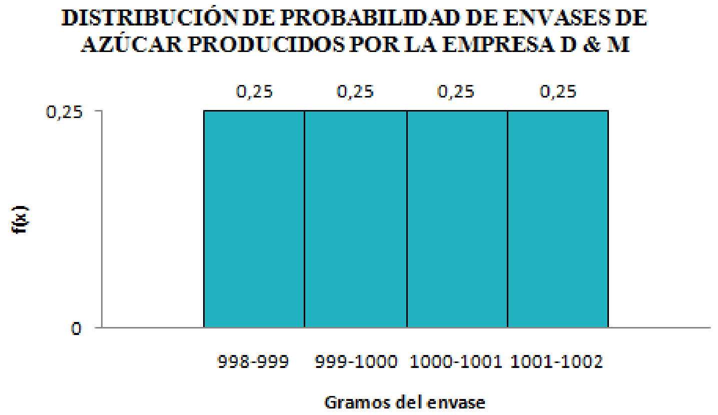
0,25

5) Sea X el contenido de envases de azúcar producidos por la empresa D & M elegido al azar. El contenido de los envases varía entre 998 y 1002 gramos. Resuelva el ejercicio de manera manual y empleando Excel.

5.1) ¿Cuál es la función de densidad de la variable X?.

0,25

5.2) Elaborar un gráfico de la distribución de probabilidades.



5.3) Calcular el valor medio esperado.

1000

5.4) Calcular la desviación estándar.

1,155

5.5) Calcular la probabilidad de que un envase pese entre la esperanza matemática y 1000,5 gramos. Realice un gráfico que ilustre la probabilidad calculada.

0,125

6) Plantee y resuelva 2 ejercicios de aplicación sobre la distribución uniforme empleando datos reales sobre cualquier tema de su preferencia. Resuelva de manera manual y empleando Excel.

D) DISTRIBUCIÓN NORMAL

i) Reseña histórica

Abrahan De Moivre (1733) fue el primero en obtener la ecuación matemática de la curva normal. Kart Friedrich Gauss y Márquez De Laplece (principios del siglo diecinueve) desarrollaron más ampliamente los conceptos de la curva. La curva normal también es llamada curva de error, curva de campana, curva de Gauss, distribución gaussiana o curva de De Moivre.

Su altura máxima se encuentra en la media aritmética, es decir su ordenada máxima corresponde a una abscisa igual a la media aritmética. La asimetría de la curva normal es nula y por su grado de apuntamiento o curtosis se clasifica en mesocúrtica.

ii) Ecuación

Su ecuación matemática de la función de densidad es:

$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

Donde:

$\sigma = desviación\ estándar$

$\sigma^2 = varianza$

$\pi = 3{,}141592654 \ldots\ldots constante\ matemática$

$e = 2{,}7182818 \ldots \ldots constante\ matemática$

X = valor en el eje horizontal

Y = altura de la curva para cualquier valor de x

$\mu = media\ aritmética$

Cuando se expresa la variable x en unidades estándar (fórmula de estandarización)

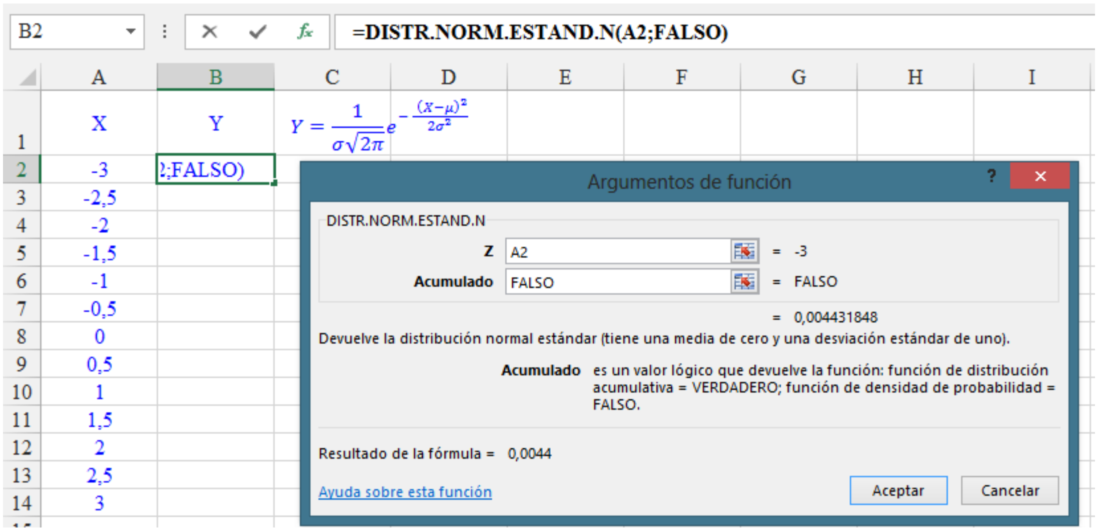
$$Z = \frac{X-\mu}{\sigma}$$

La ecuación anterior es remplazada por la llamada forma canónica, la cual es

$$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}$$

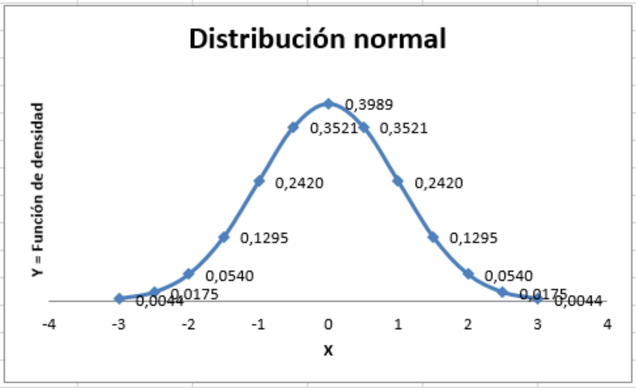
Para calcular Y en Excel se procede de la siguiente manera:

a) Se ubica valores para X del -3 hasta el 3. Se insertar la función DISTR.NORM.ESTAND.N. En la ventana de argumentos de función, en Z se seleccionada A2 que representa al -3, y en Acumulado es escribe FALSO. Clic en Aceptar. Se arrastra con el mouse para obtener los demás valores.



b) Para obtener la gráfica se inserta gráfico de dispersión.

	A	B	C
1	X	Y	$Y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
2	-3	0,0044	=DISTR.NORM.ESTAND.N(A2;FALSO)
3	-2,5	0,0175	=DISTR.NORM.ESTAND.N(A3;FALSO)
4	-2	0,0540	=DISTR.NORM.ESTAND.N(A4;FALSO)
5	-1,5	0,1295	=DISTR.NORM.ESTAND.N(A5;FALSO)
6	-1	0,2420	=DISTR.NORM.ESTAND.N(A6;FALSO)
7	-0,5	0,3521	=DISTR.NORM.ESTAND.N(A7;FALSO)
8	0	0,3989	=DISTR.NORM.ESTAND.N(A8;FALSO)
9	0,5	0,3521	=DISTR.NORM.ESTAND.N(A9;FALSO)
10	1	0,2420	=DISTR.NORM.ESTAND.N(A10;FALSO)
11	1,5	0,1295	=DISTR.NORM.ESTAND.N(A11;FALSO)
12	2	0,0540	=DISTR.NORM.ESTAND.N(A12;FALSO)
13	2,5	0,0175	=DISTR.NORM.ESTAND.N(A13;FALSO)
14	3	0,0044	=DISTR.NORM.ESTAND.N(A14;FALSO)
15			



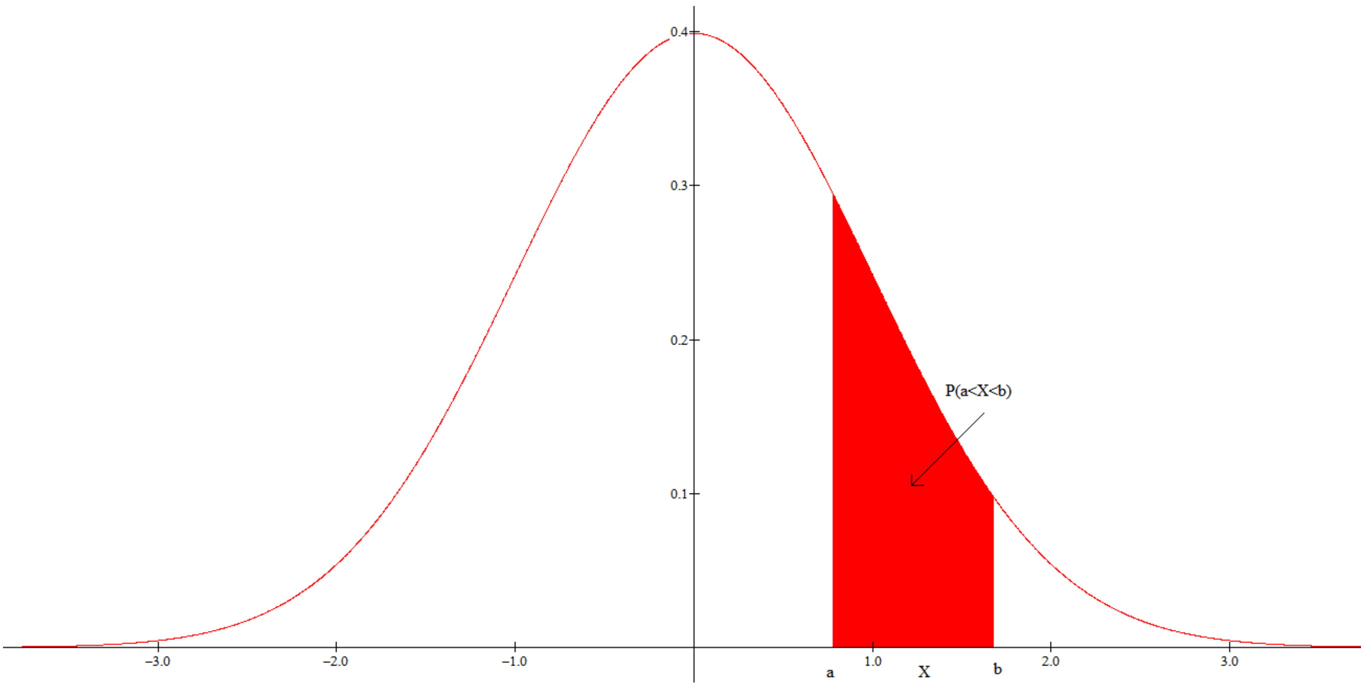
Nota: No existe una única distribución normal, sino una familia de distribuciones con una forma común, diferenciadas por los valores de su media y su varianza. De entre todas ellas, la más utilizada es la **distribución normal estándar**, que corresponde a una distribución con una media aritmética de 0 y una desviación típica de 1.

iii) Área bajo la curva

El área total limitada por la curva y el eje "X" es 1, por lo tanto, el área bajo la curva entre X = a y X = b, con a < b, representa la probabilidad de que X esté entre a y b. Esta probabilidad se denota por:

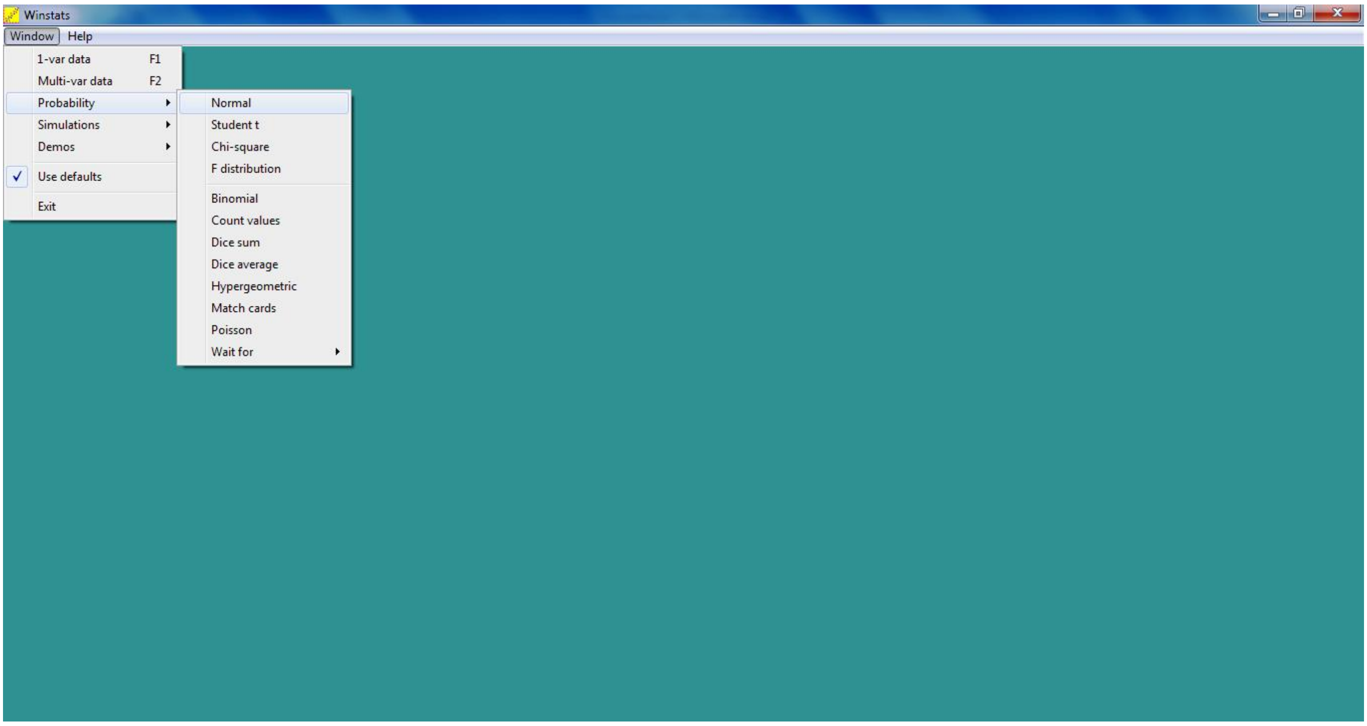
$$P(a < X < b)$$

Esta probabilidad se ilustra en el siguiente gráfico elaborado con el programa Winstats.

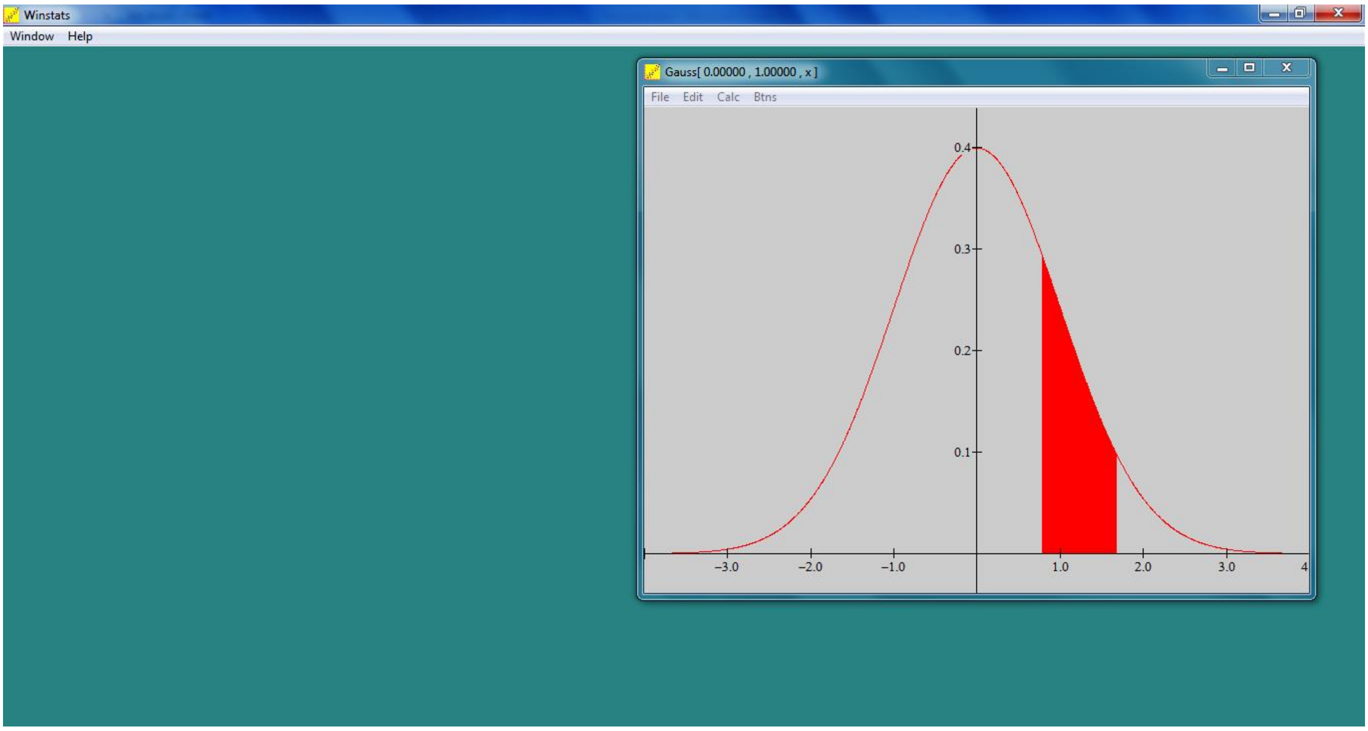


Para elaborar el gráfico se procede de la siguiente manera:

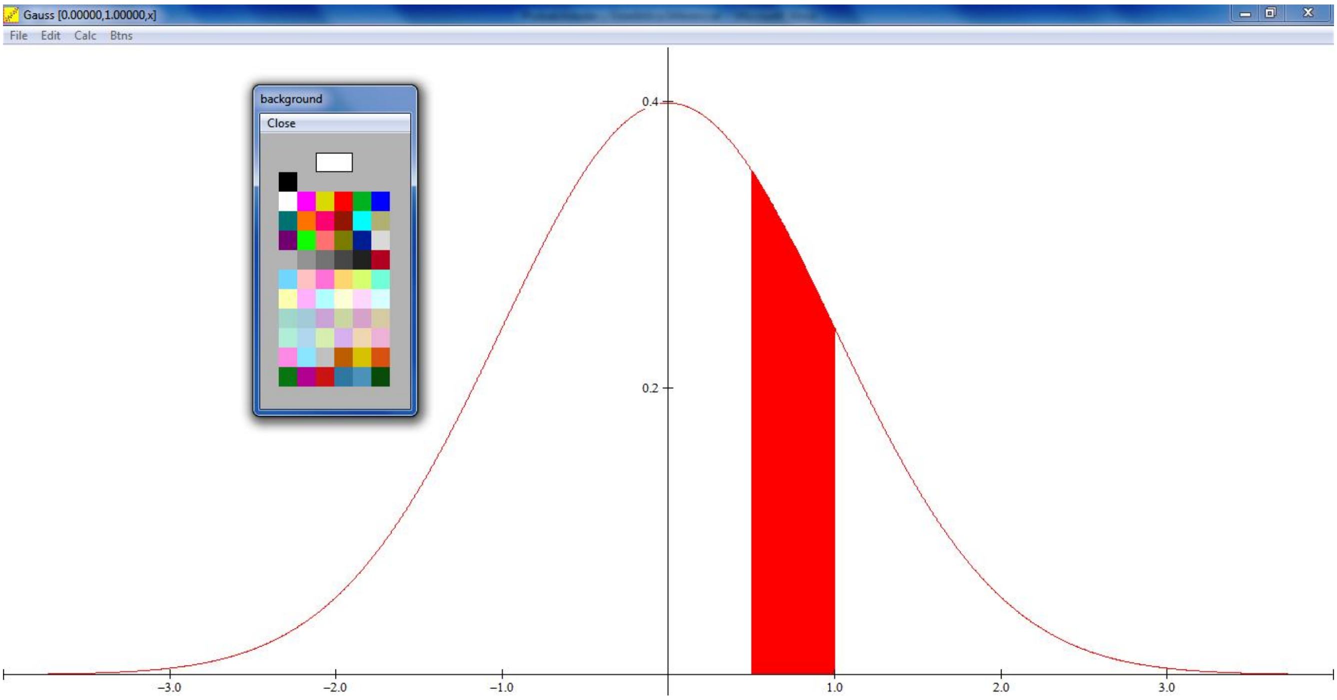
a) Se abre el programa. Clic en Window- Probability



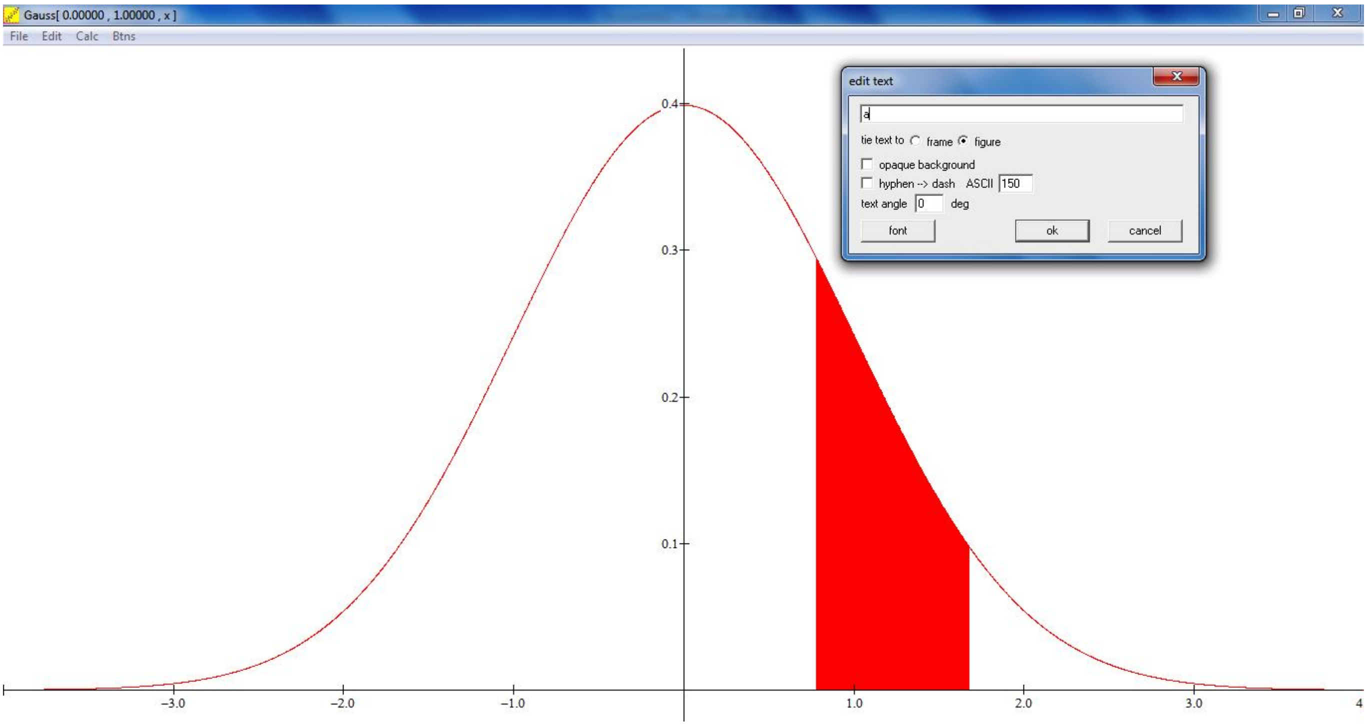
b) Clic en Normal



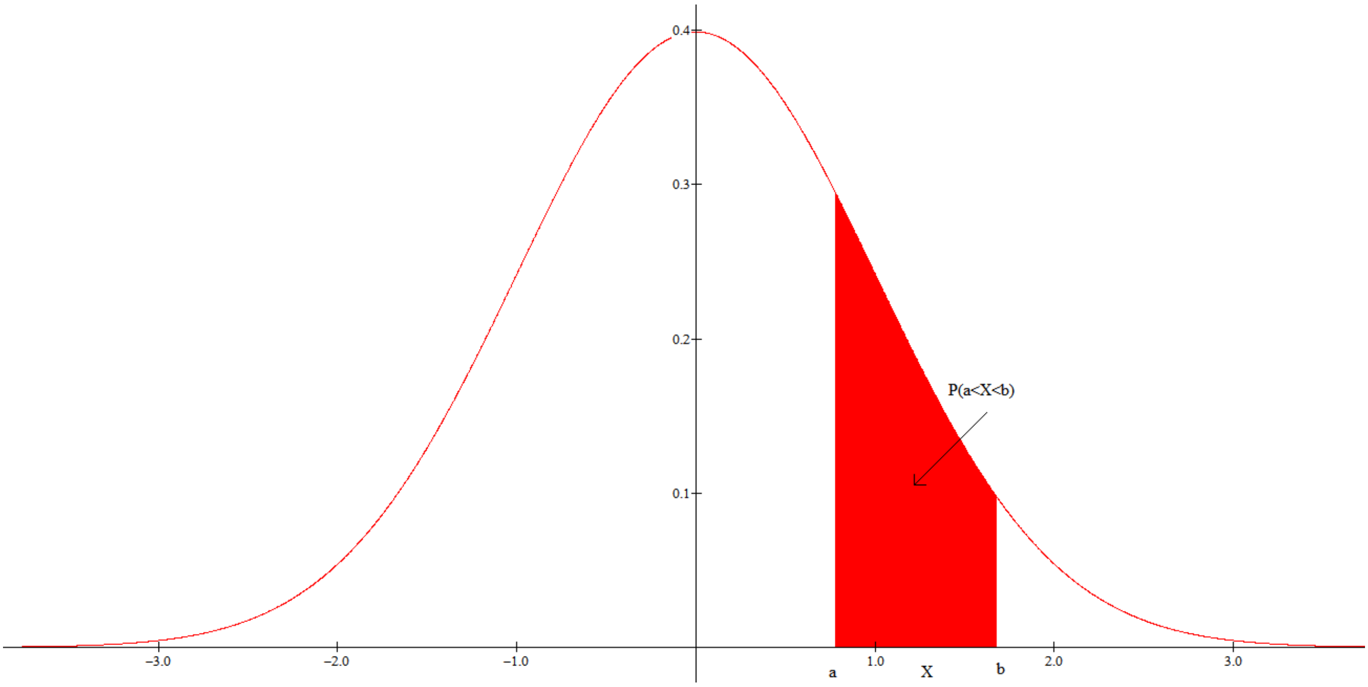
c) Para cambiar el color del fondo, maximizar la ventana de la curva. Clic en Edit-Colors y luego en Window background. Seleccionar el color blanco para el fondo.



d) Para escribir, clic en Btns y luego en Text mode. Clic derecho en cualquier parte de la pantalla. Luego escribir en la venta edit text. Clic en ok



e) Se obtiene el siguiente gráfico

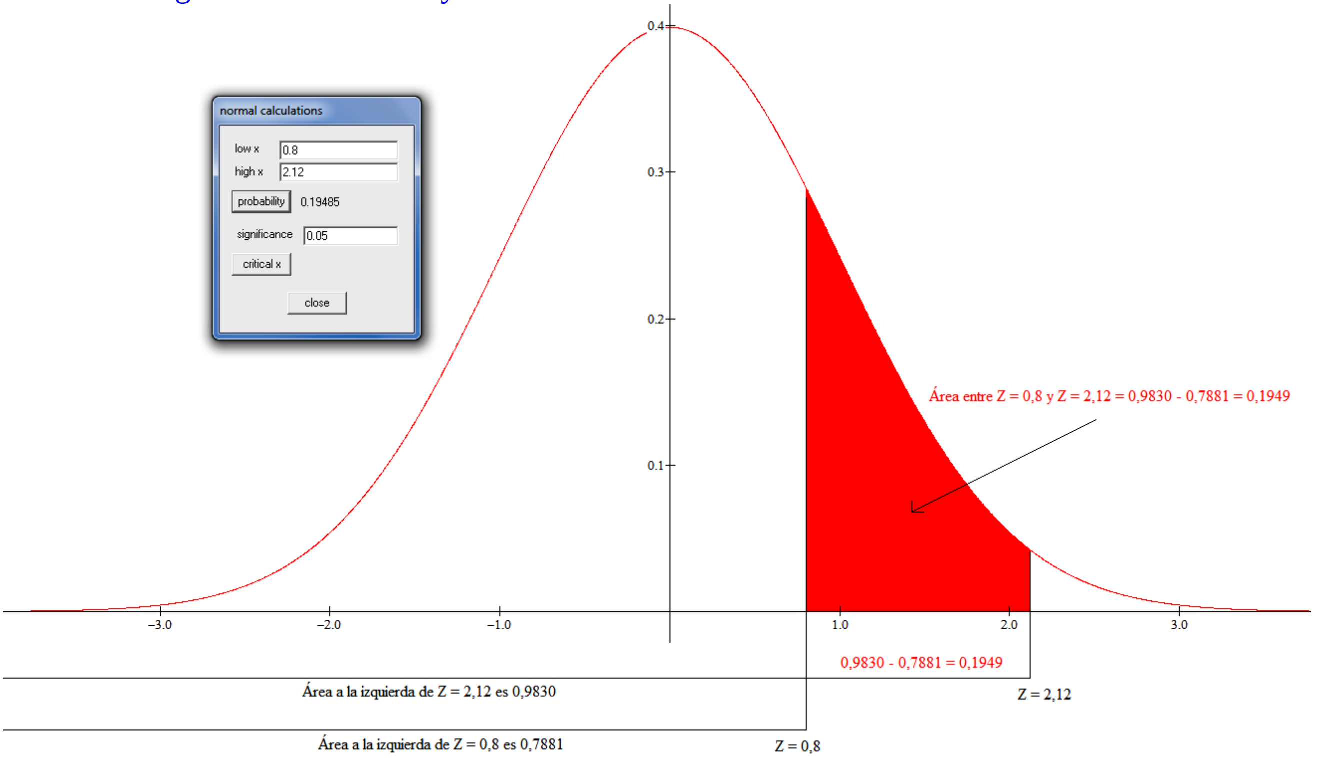


Ejemplos ilustrativos

1) Averigüe el área bajo la curva de distribución normal entre Z = 0,8 y Z = 2,12

Solución:

Realizando el gráfico en Winstats y Paint se obtiene:

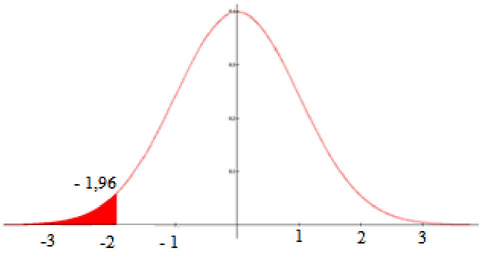


El área a la izquierda de Z = 0,8 con lectura en la tabla de la distribución normal es 0,7881

El área a la izquierda de Z = 2,12 con lectura en la tabla de la distribución normal es 0,9830

TABLA Nº 3
DISTRIBUCIÓN NORMAL

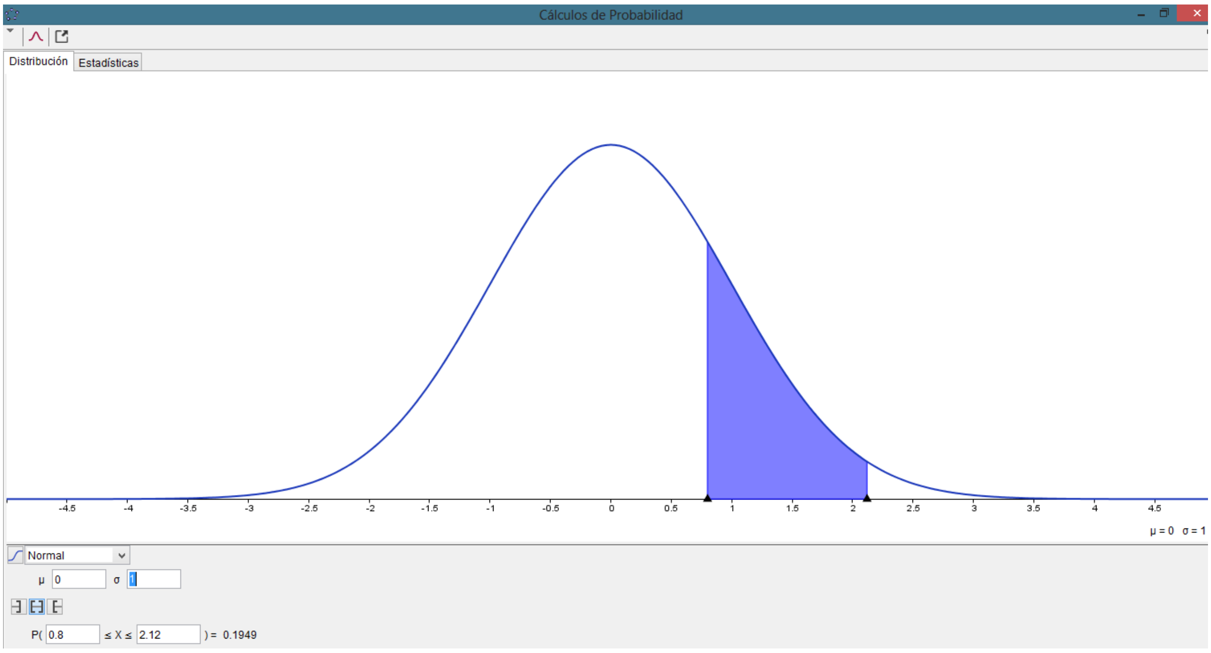
$$Z = \frac{X-\mu}{\sigma}$$

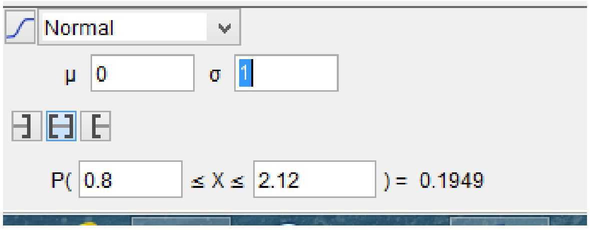


Ejemplo:
$P(Z \leq -1{,}96) = 0{,}0250$

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
	⋮		⋮							
0,8	**0,7881**	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	**0,9830**	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857

El área Z = 0,8 y Z = 2,12 es 0,9830 – 0,7881 = 0,1949
Los cálculos en GeoGebra se presentan en la siguiente figura:





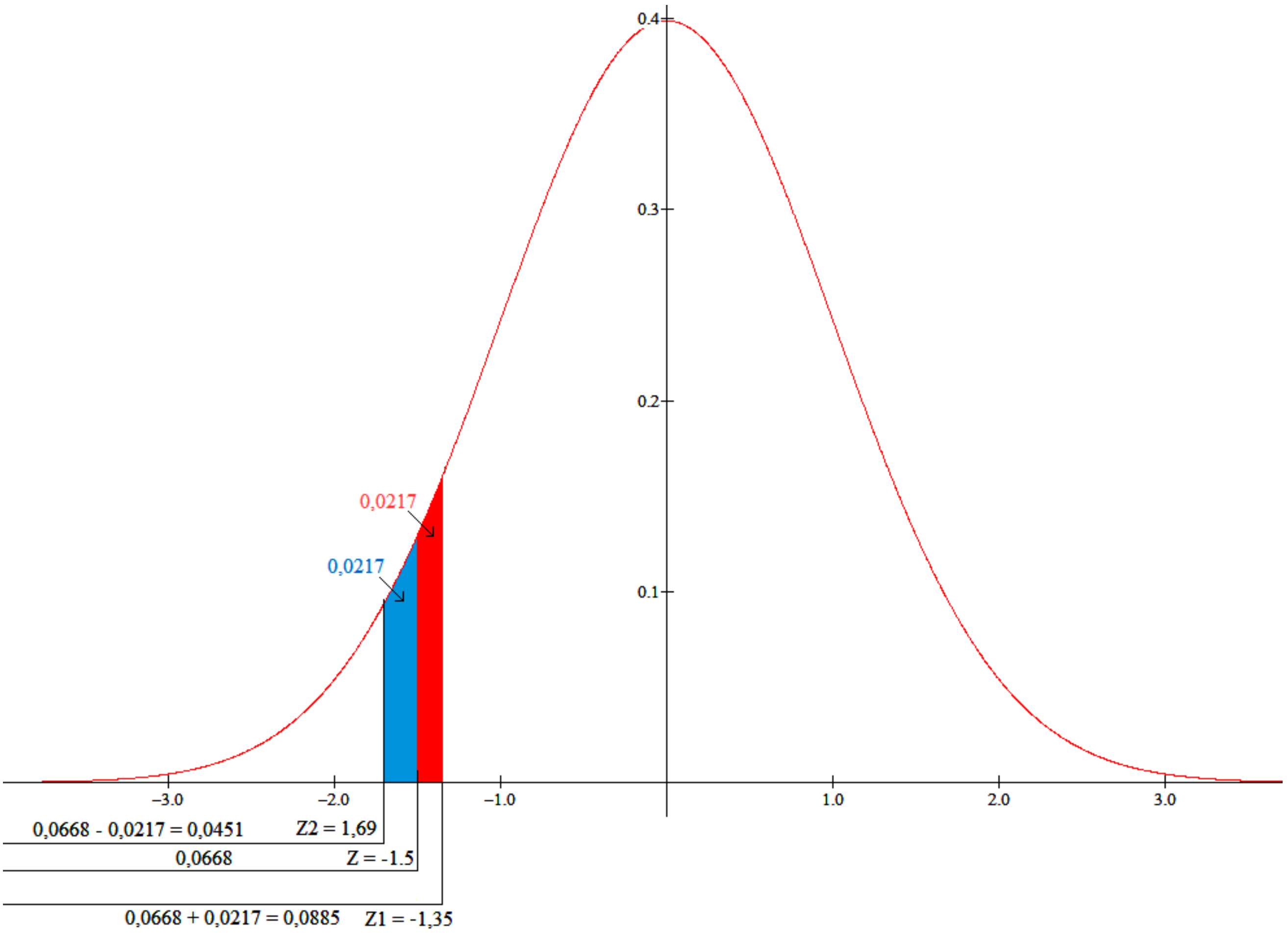
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	Z_1	0,8				
2	Z_2	2,12				
3	Área a la izquierda de Z_1	0,7881	=DISTR.NORM.ESTAND.N(B1;VERDADERO)			
4	Área a la izquierda de Z_2	0,9830	=DISTR.NORM.ESTAND.N(B2;VERDADERO)			
5	$P(0,8 \leq X \leq 2,12)$	0,1949	=B4-B3			

2) Halle Z si el área entre -1,5 y Z es 0,0217

Solución:

Realizando un gráfico ilustrativo en Winstats y Paint se obtiene:



Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	Z	-1,5				
2	Área entre Z y Z_1	0,0217				
3	Área a la izquierda de Z	0,0668	=DISTR.NORM.ESTAND.N(B1;VERDADERO)			
4	Área a la izquierda de Z_1	0,0885	=B2+B3			
5	Z_1	-1,3500	=INV.NORM.ESTAND(B4)			
6	Área a la izquierda de Z_2	0,0451	=B3-B2			
7	Z_2	-1,6943	=INV.NORM.ESTAND(B6)			

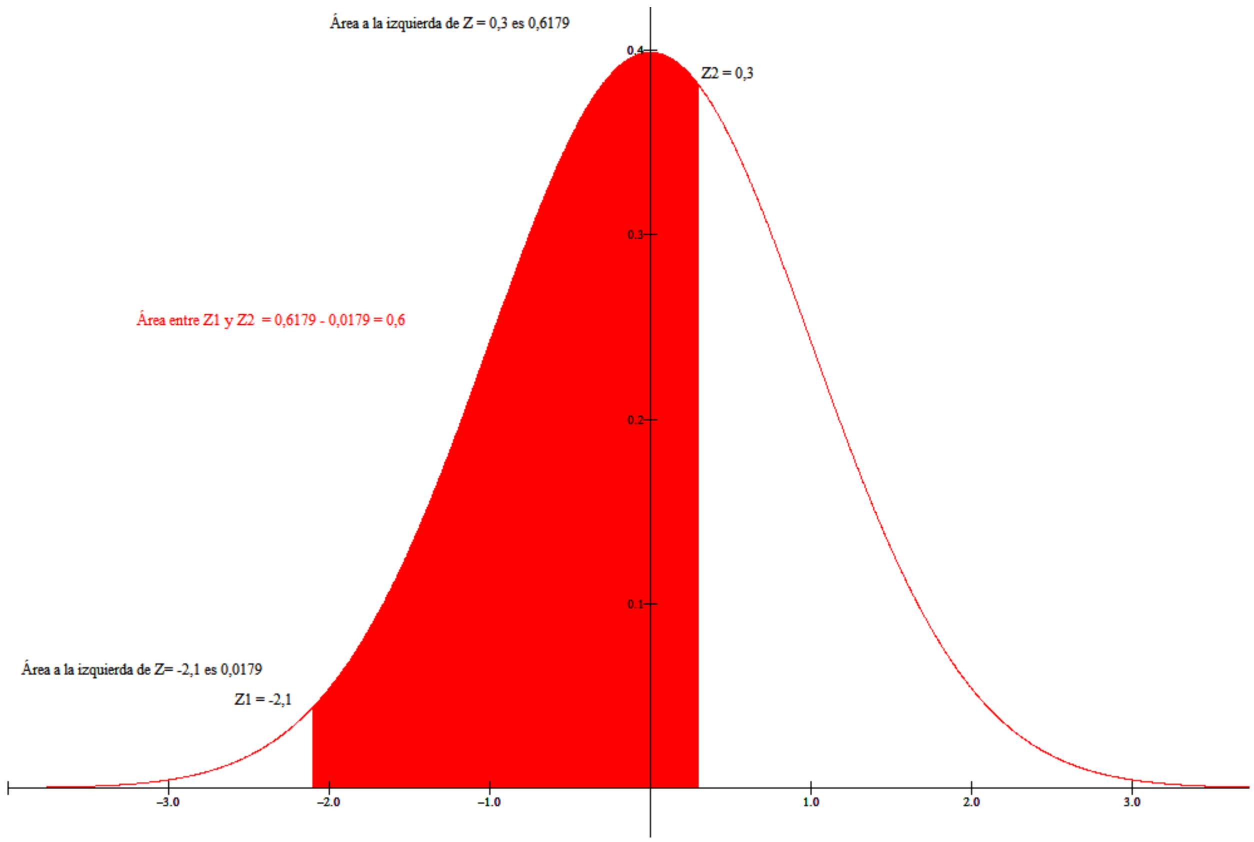
3) El peso promedio de 200 estudiantes varones de cierta universidad es 151 libras con una desviación típica de 15 libras. Si los pesos están distribuidos normalmente, calcular la probabilidad y el número de estudiantes que pesan Entre 120 y 155 libras

Solución: La curva normal corresponde a una función continua (valor decimal). Para resolver estos problemas se emplea los límites inferior y superior según sea el caso, es decir, para este problema es entre 119,5 y 155,5 libras

Para saber la ubicación de los datos dentro de una distribución estadística nos valemos de los puntajes Z. Para calcular los puntajes Z se normaliza o estandariza los datos de la siguiente manera:

$$Z = \frac{X-\mu}{\sigma} \quad ; \quad Z_1 = \frac{X_1-\mu}{\sigma} = \frac{119{,}5-151}{15} = -2{,}1 \quad ; \quad Z_2 = \frac{X_2-\mu}{\sigma} = \frac{155{,}5-151}{15} = 0{,}3$$

Graficando se obtiene:



El área a la izquierda de Z = 0,3 con lectura en la tabla de la distribución normal es 0,6179
El área a la izquierda de Z = -2,1 con lectura en la tabla de la distribución normal es 0,0179
El área entre -2,1 y 0,3 es 0,6179 – 0,0179 = 0,6 = 60%
El número de estudiantes es 0,6 x 200 = 120

Los cálculos en Excel se muestran en la siguiente figura:

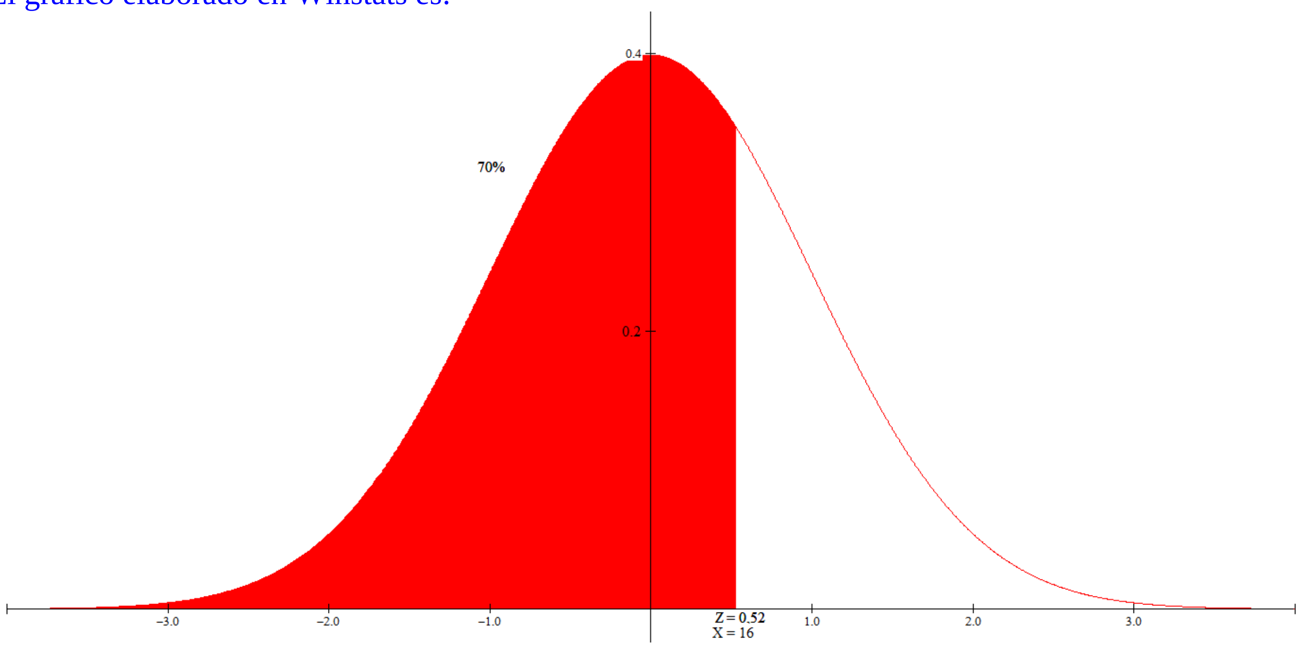
	A	B	C	D	E	F
1	n	200				
2	μ	151				
3	σ	15				
4	X_1	119,5				
5	X_2	155,5				
6	Z_1	-2,1	=NORMALIZACION(B4;B2;B3)		$Z = \frac{X-\mu}{\sigma}$	
7	Z_2	0,3	=NORMALIZACION(B5;B2;B3)			
8	Área a la izquierda de Z_1	0,0179	=DISTR.NORM.ESTAND.N(B6;VERDADERO)			
9	Área a la izquierda de Z_2	0,6179	=DISTR.NORM.ESTAND.N(B7;VERDADERO)			
10	Área entre Z_1 y Z_2	0,60	=B9-B8			
11	Número de estudiantes	120	=B10*B1			

4) Las calificaciones que obtienen los alumnos en un examen siguen una distribución normal, siendo la media igual a 14. El 70% de los alumnos obtienen una calificación inferior a 16.
4.1) Calcule la desviación típica de las calificaciones
4.2) Se escoge un alumno al azar, calcule el porcentaje de obtener una calificación superior a 18

Solución
4.1)
Si el área es inferior al 70%, entonces con lectura en la tabla se obtiene el valor de Z = 0,52
Reemplazando valores en la fórmula y realizando las operaciones se obtiene:

$$Z = \frac{X-\mu}{\sigma} \Rightarrow 0{,}52 = \frac{16-14}{\sigma} \Rightarrow \sigma = \frac{16-14}{0{,}52} = 3{,}8$$

El gráfico elaborado en Winstats es:



4.2)
Remplazando valores en la fórmula se obtiene el siguiente puntaje o número Z:

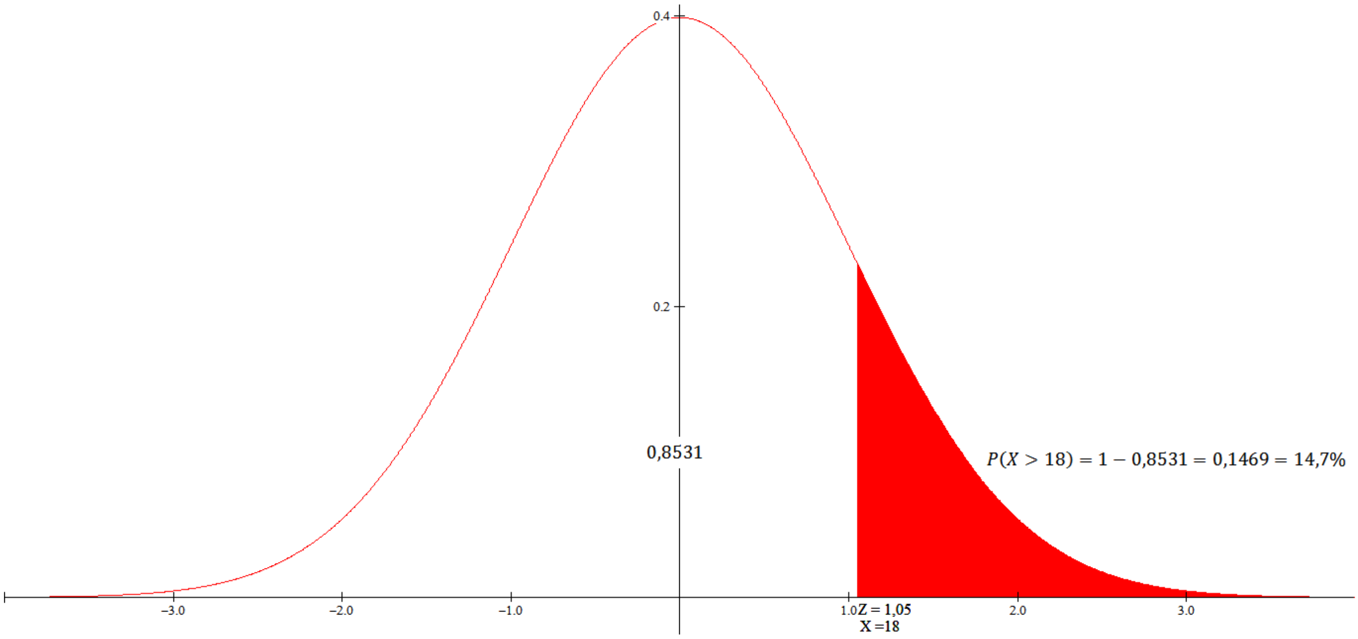
$$Z = \frac{X-\mu}{\sigma} \Rightarrow Z = \frac{18-14}{3,8} \Rightarrow Z = 1,05$$

Con lectura en la tabla para Z = 1,05 se obtiene un área de 0,8531, la cual representa una probabilidad inferior a la calificación de 18

Para calcular la probabilidad de obtener una calificación superior a 18 se realiza la siguiente operación:

$$P(X > 18) = 1 - 0,8531 = 0,1469 = 14,7\%$$

El gráfico elaborado en Winstats y en Paint es:



Por lo tanto existe una probabilidad de 14,7% de obtener una calificación superior a 18

TAREA DE INTERAPRENDIZAJE N° 12

1) Realice un organizador gráfico sobre la distribución normal

2) Empleando Excel grafique la campana de la distribución normal calculando previamente los respectivos valores de función de densidad.

3) Calcule el área bajo la curva de distribución normal empleando la tabla de la distribución normal, Excel y GeoGebra entre:

3.1) Z = 0 y Z = 1,2

0,3849

3.2) Z = 1,2 y Z = 2,2

0,1012

3.3) Z = -2,2 y Z = -1,2

0,1012

3.4) Z = -2 y Z = 1

0,8186

3.5) Z = -1 y Z = 1

0,6827

3.6) Z = -2 y Z = 2

0,9545

3.7) Z = -3 y Z = 3

0,9973

3.8) A la izquierda de Z = 1,54

0,9382

3.9) A la izquierda de Z = -1,54

0,0618

3.10) A la izquierda de Z = -0,6

0,2743

3.11) A la izquierda de Z = 0,6

0,7257

3.12) Plantee y resuelta un ejercicio similar a los anteriores

4) Halle Z de forma manual, empleando Excel y Winstats, si:

4.1) El área entre 0 y Z es 0,3531

±1.05

4.2) El área entre 0 y Z es 0.3972 y Z es positivo

1,266

4.3) El área a la derecha de Z es 0,115

1,20

4.4) El área a izquierda de Z es 0.8621

1,09

4.5) El área a la izquierda de Z = 0,6692

0,44

4.6) El área entre -1 y Z es 0,8186

2

4.7) El área entre -1 y Z es 0,1499

-2,38 y -0,5

4.8) El área entre -0,5 y Z es 0,2313

-1,42 y 0,1

4.9) El área entre 0,5 y Z es 0,1450

0,12 y 0,98

4.10) Plantee y resuelva un ejercicio similar a los anteriores

5) En cierta área, un conductor promedio recorre una distancia de 1200 millas al mes (1,609 Km al mes), con una desviación estándar de 150 millas. Suponga que el número de millas se aproxima mediante una curva normal, encuentre la probabilidad de todos los automovilistas que recorren entre 1200 y 1600 millas por mes. Resuelva empleando la tabla de la distribución normal, Excel y Winstats

49,6%

6) Se sabe que la longitud de cierto mueble se distribuye en forma normal con una media de 84 cm y una desviación estándar 0,4. Calcular el porcentaje de muebles que cumplen con especificaciones 84±1. Resuelva empleando la tabla de la distribución normal, Excel y Winstats

98,76%

7) Las calificaciones que obtienen alumnos universitarios en un examen siguen una distribución normal, siendo la media igual a 7. El 80% de los alumnos obtienen una calificación inferior a 8. Resuelva empleando la tabla de la distribución normal, Excel y Winstats

7.1) Calcule la desviación típica de las calificaciones

1,2

7.2) Se escoge un alumno al azar, calcule el porcentaje de obtener una calificación superior a 9

4,6%

8) La altura de los árboles de un bosque sigue una distribución normal con una altura media de 17 m. Se selecciona un árbol al azar. La probabilidad de que la altura del árbol seleccionado sea mayor que 24 metros es 6%. Si se selecciona al azar 100 árboles. Utilice Excel y GeoGebra.

8.1) Calcule la desviación típica de las alturas de los árboles

4,5

8.2) Calcular el número esperado de árboles cuyas alturas varían entre 17m y 24m.

44

9) Las calificaciones de los estudiantes de un colegio siguen una distribución normal con una media de 15. Se selecciona un estudiante al azar. La probabilidad de que la calificación del estudiante sea mayor que 17 es del 8%. Resuelva empleando la tabla de la distribución normal, Excel y Winstats

9.1) Calcule la desviación típica de las calificaciones

1,4

9.2) Calcular la probabilidad de que el estudiante seleccionado tenga una calificación menor que 17

0,92

9.3) La probabilidad de que el estudiante tenga una calificación menor que X es 0,08. Calcular el valor de X.

13

9.4) Se selecciona al azar 100 estudiantes. Calcule el número esperado de estudiantes cuyas calificaciones varían entre 1 y 13

8

10) Las calificaciones que obtienen los alumnos en un examen, cuya máxima nota es 10, siguen una distribución normal, siendo la media igual a 7. El 70% de los alumnos obtienen una calificación inferior a 8,5. Se escoge un alumno al azar, éste alumno tiene la misma probabilidad de obtener una calificación inferior a 9 que una calificación superior a X. Emplear la tabla de la distribución normal, Excel y Winstats

10.1) Calcule la desviación típica de las calificaciones

2,86

10.2) Calcule el porcentaje de obtener una calificación inferior a 9

75,78%

10.3) Calcule el valor de X

5

11) Plantee y resuelva un problema similar al anterior con datos obtenidos en un examen o prueba de cualquier asignatura de su preferencia. Emplear la tabla de la distribución normal, Excel y GeoGebra.

12) Investigue en la biblioteca o internet sobre el empleo o aplicación de los puntajes Z (números Z). Presente la consulta empleando un organizador gráfico.

CAPÍTULO III

ESTIMACIÓN DE INTERVALOS DE CONFIANZA

RESULTADOS DE APRENDIZAJE DEL CAPÍTULO

Al finalizar el presente capítulo el lector podrá evidenciar que:

- ✓ Calcula e interpreta estimaciones de intervalos de confianza de manera manual, empleando Excel, Winstats y GeoGebra.
- ✓ Calcula el tamaño de una muestra a partir de situaciones concretas de la vida cotidiana de manera manual, empleando Excel y Winstats.
- ✓ Plantea y resuelve ejercicios de aplicación sobre intervalos de confianza y del tamaño de una muestra de manera manual, utilizando Excel,Winstats y GeoGebra.

CONTENIDOS

- ✓ Estimación del intervalo de confianza para la media (σ conocida)
- ✓ Estimación del intervalo de confianza para la media (σ desconocida).- Distribución t de Student
- ✓ Estimación del intervalo de confianza para una proporción
- ✓ Determinación del tamaño de la muestra

ESTIMACIÓN DE INTERVALOS DE CONFIANZA

La estadística inferencial es el proceso de uso de los resultados derivados de las muestras para obtener conclusiones acerca de las características de una población. La estadística inferencial nos permite estimar características desconocidas como la media de la población o la proporción de la población. Existen dos tipos de estimaciones usadas para estimar los parámetros de la población: la estimación puntual y la estimación de intervalo. Una estimación puntual es el valor de un solo estadístico de muestra. Una estimación del intervalo de confianza es un rango de números, llamado intervalo, construido alrededor de la estimación puntual. El intervalo de confianza se construye de manera que la probabilidad del parámetro de la población se localice en algún lugar dentro del intervalo conocido.

Suponga que quiere estimar la media de todos los alumnos en su universidad. La media para todos los alumnos es una media desconocida de la población, simbolizada como μ. Usted selecciona una muestra de alumnos, y encuentra que la media es de 5,8. La muestra de la media $\bar{x} = 5{,}8$ es la estimación puntual de la media poblacional μ. ¿Qué tan preciso es el 5,8? Para responder esta pregunta debe construir una estimación del intervalo de confianza.

Recuerde que la media de la muestra $\bar{x}$ es una estimación puntual de la media poblacional μ. Sin embargo, la media de la muestra puede variar de una muestra a otra porque depende de los elementos seleccionados en la muestra. Tomando en cuenta la variabilidad de muestra a muestra, se aprenderá a desarrollar la estimación del intervalo para la media poblacional. El intervalo construido tendrá una confianza especificada de la estimación correcta del valor del parámetro poblacional μ . En otras palabras, existe una confianza especificada de que μ se encuentre en algún lugar en el rango de números definidos por el intervalo.

En general, el *nivel de confianza* se simboliza con $(1-\alpha)\cdot 100\%$, donde α es la proporción de las colas de la distribución que están fuera del intervalo de confianza. La proporción de la cola superior e inferior de la distribución es $\alpha/2$

Ejemplo ilustrativo
Calcular la proporción de cola superior e inferior para un intervalo del 95% de confianza

Solución:
$$Nivel\ de\ confianza = (1-\alpha)\cdot 100\%$$

Remplazando valores en la fórmula anterior del novel de confianza se obtiene:
$$95\% = (1-\alpha)\cdot 100\% \Rightarrow \frac{95\%}{100\%} = (1-\alpha) \Rightarrow \alpha = 1-\frac{95\%}{100\%} \Rightarrow \alpha = 1-0{,}95 = 0{,}05$$

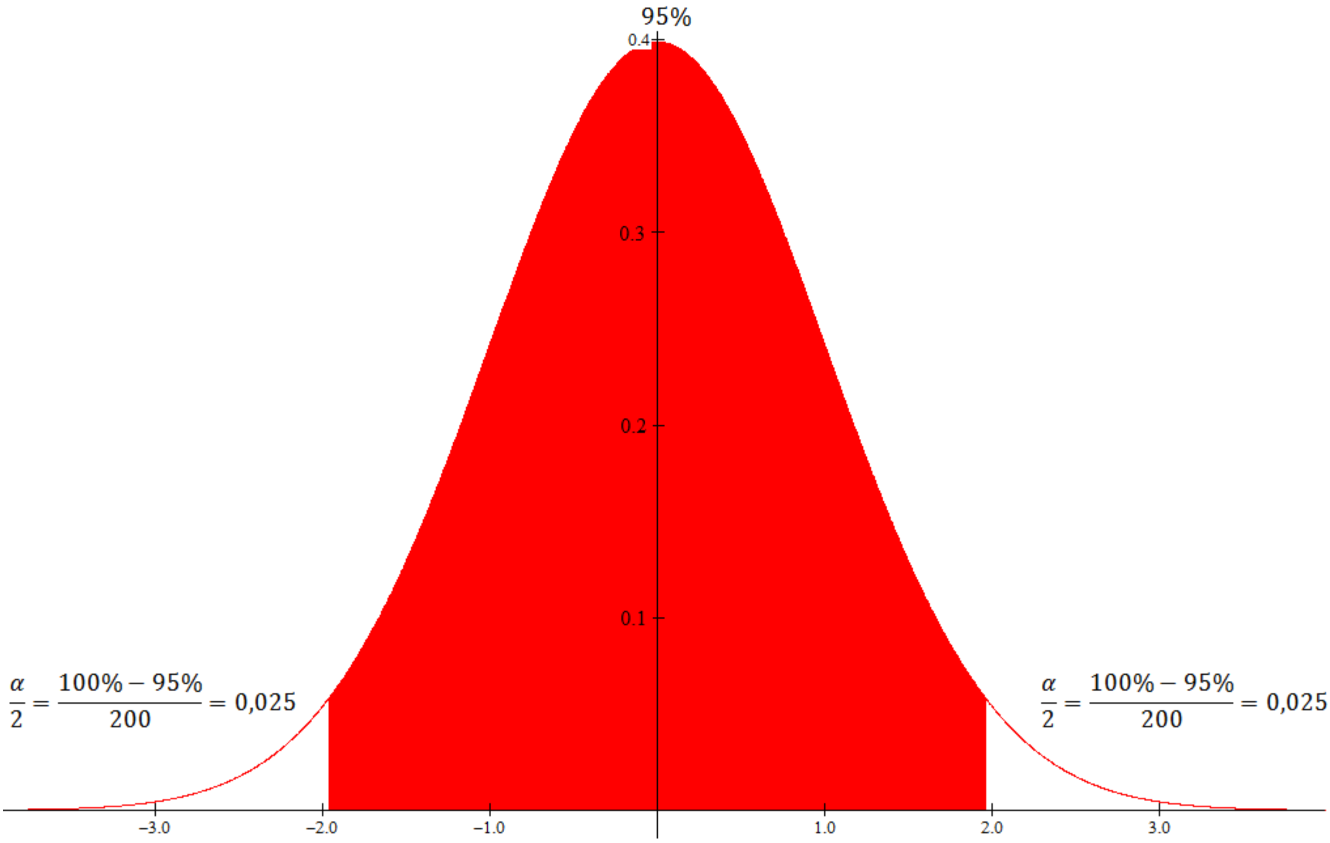
La proporción de la cola superior e inferior de la distribución es:
$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

Otra forma para calcular la proporción de la cola superior e inferior de la distribución es aplicando la siguiente fórmula:

$$\frac{\alpha}{2} = \frac{100\% - Nivel\ de\ confianza}{200}$$

$$\frac{\alpha}{2} = \frac{100\% - 95\%}{200} = 0{,}025$$

El siguiente gráfico ilustra lo calculado:



3.1)ESTIMACIÓN DEL INTERVALO DE CONFIANZA PARA LA MEDIA (σ CONOCIDA)

Se emplea la siguiente fórmula:

$$\bar{x} - Z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z\frac{\sigma}{\sqrt{n}}$$

Donde:
Z = valor crítico de la distribución normal estandarizada

Se llama *valor crítico* al valor de Z necesario para construir un intervalo de confianza para la distribución. El 95% de confianza corresponde a un valor α de 0,05. El valor crítico Z correspondiente al área acumulativa de 0,975 es 1,96 porque hay 0,025 en la cola superior de la distribución y el área acumulativa menor a Z = 1,96 es 0,975.

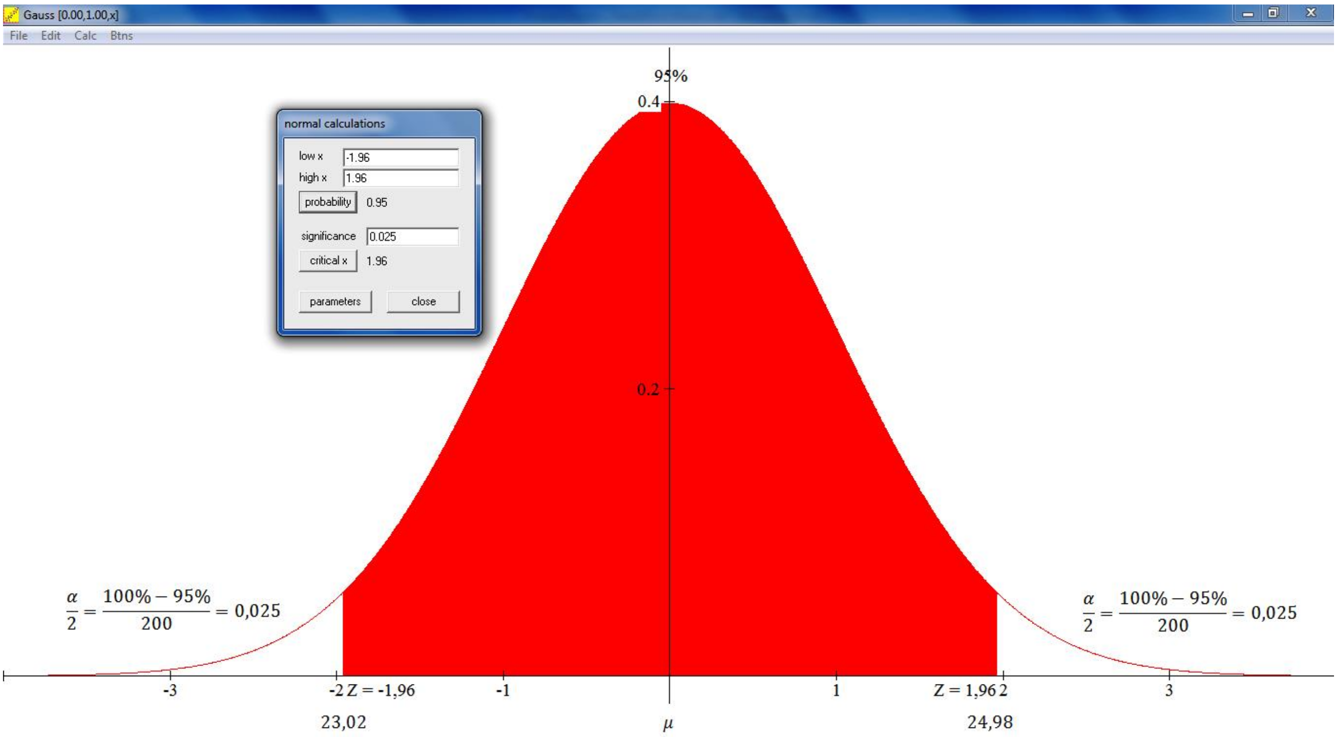
Un nivel de confianza del 95% lleva a un valor Z de 1,96. El 99% de confianza corresponde a un valor α de 0,01. El valor de Z es aproximadamente 2,58 porque el área de la cola alta es 0,005 y el área acumulativa menor a Z = 2,58 es 0,995.

Ejemplo ilustrativo

Si $\bar{x} = 24\,; \sigma = 3\ y\ n = 36$ construya para la media poblacional μ una estimación de intervalo de confianza del 95%

Solución:

Realizando un gráfico ilustrativo en Winstats y Paint se obtiene:



Con lectura en la tabla de la distribución normal para un área de 0,025 se obtiene Z = -1,96. Por simetría se encuentra el otro valor Z = 1,96

Remplazando valores y realizando lo cálculos se obtiene:

$$\bar{x} - Z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z\frac{\sigma}{\sqrt{n}}$$

$$24 - 1{,}96\frac{3}{\sqrt{36}} \leq \mu \leq 24 + 1{,}96\frac{3}{\sqrt{36}}$$

$$23{,}02 \leq \mu \leq 24{,}98$$

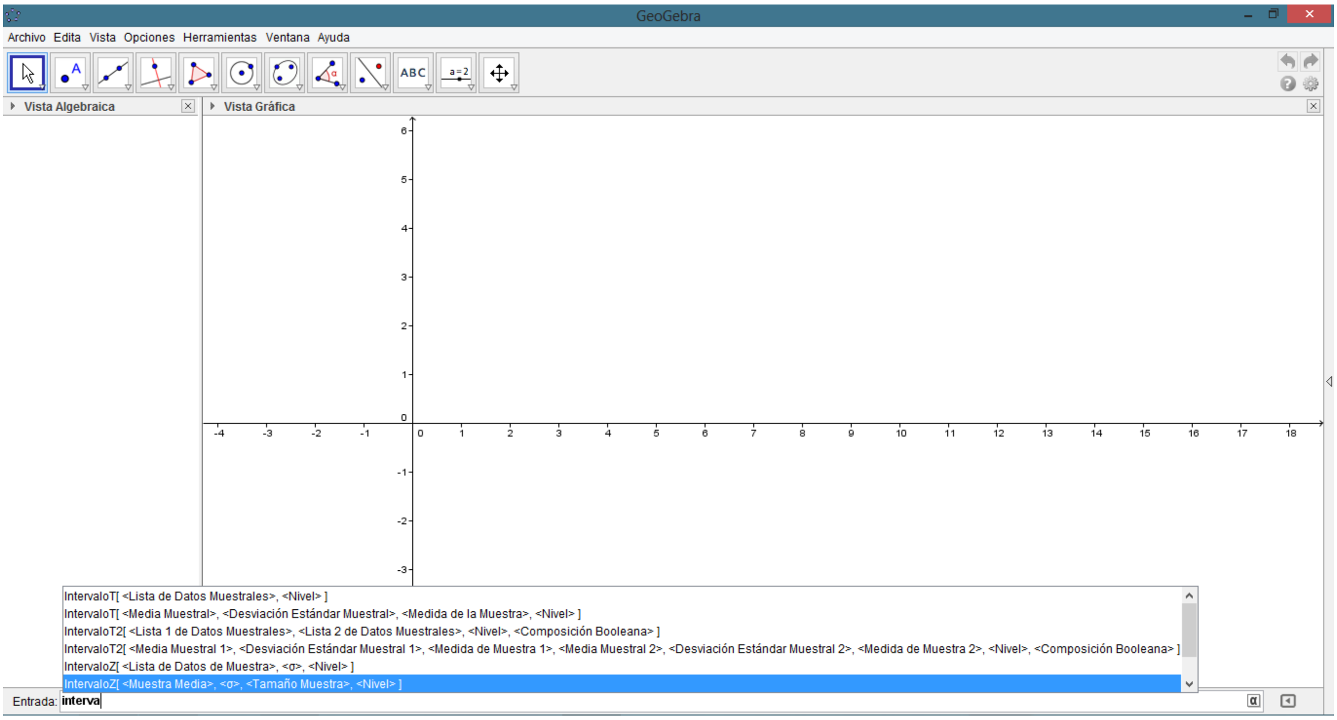
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	$\bar{X}$	24				
2	σ	3				
3	n	36				
4	Confianza	95				
5	α	0,05	=(100-B4)/100			
6	$Z\frac{\sigma}{\sqrt{n}}$	0,98	=INTERVALO.CONFIANZA.NORM(B5;B2;B3)			
7						
8	23,02	$\leq \mu \leq$	24,98			
9	=B1-B6		=B1+B6			

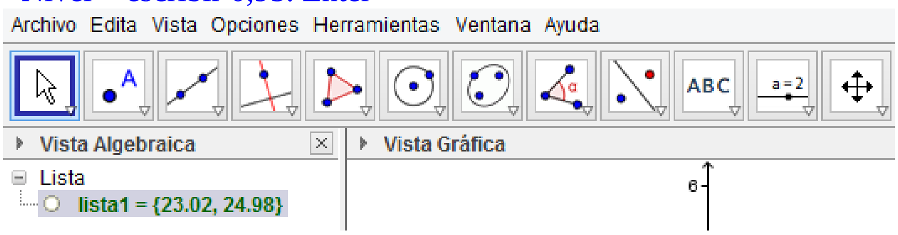
Interpretación: Existe un 95% de confianza de que la media poblacional se encuentre entre 23,02 y 24,98

En Geogebra se sigue los siguientes pasos:

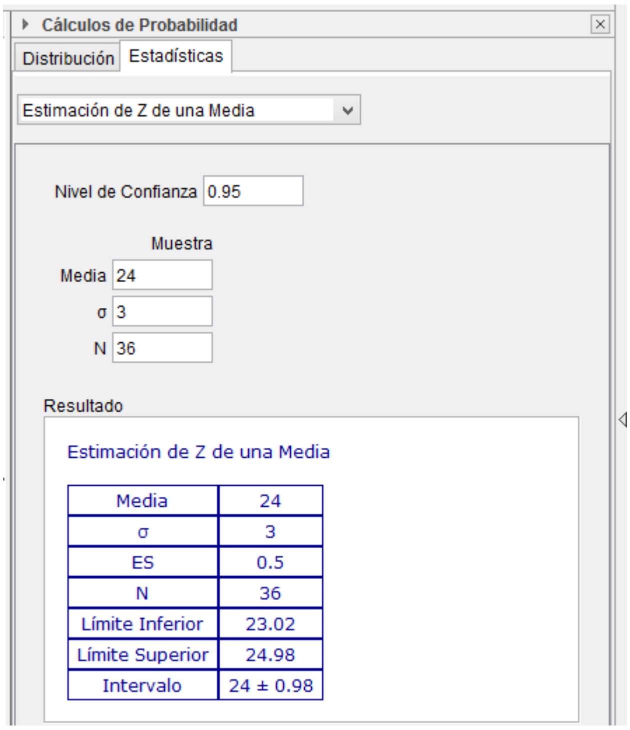
a) En Entrada, seleccione IntervaloZ[<Muestra Media>, <σ>, <Tamaño Muestra>, <Nivel>]



b) En <Muestra Media> escribir 24, en <σ> escribir 3, en <Tamaño Muestra> escribir 24 , y en <Nivel> escribir 0,95. Enter



O también



TAREA DE INTERAPRENDIZAJE Nº 13

1) Conteste:

1.1) ¿Qué es estadística inferencial?
1.2) ¿Qué permite estimar la estadística inferencial?
1.3) ¿Qué es estimación puntual?
1.4) ¿Qué es estimación de intervalo?
1.5) ¿Para qué se sirve construir una estimación del intervalo de confianza?
1.6) ¿Qué es la media de la muestra?
1.7) ¿Por qué la media de la muestra puede variar?
1.8) ¿Por qué no se sabe con certeza si el intervalo específico incluye la media poblacional o no?
1.9) ¿Qué es nivel de confianza?
1.10) ¿Qué es valor crítico?
1.11) ¿El valor Z para el 95% de confianza es?
1.12) ¿El valor Z para el 99% de confianza es?

Resolver los siguientes ejercicios de manera manual, con Excel y GeoGebra para realizar los cálculos y con Winstats para los gráficos

2) Si $\bar{x} = 24{,}2\,;\sigma = 3\ y\ n = 36$ construya para la media poblacional μ una estimación de intervalo de confianza del

2.1) 95%

$23{,}22 \leq \mu \leq 25{,}18$

2.2) 99%

$22{,}91 \leq \mu \leq 25{,}49$

3) Un fabricante de papel para computadora tiene un proceso de producción que opera continuamente a lo largo del turno. Se espera que el papel tenga una media de longitud de 11 pulgadas y una desviación estándar de 0,02 pulgadas. Se selecciona una muestra de 100 hojas con una media de longitud del papel de 10,998 pulgadas. Calcule la estimación del intervalo de confianza del

3.1) 95%

$10{,}99408 \leq \mu \leq 11{,}00192$

3.2) 60%

$10{,}99632 \leq \mu \leq 10{,}99968$

4) El gerente de control de calidad de una fábrica de focos necesita estimar la media de vida de un gran embarque de focos. La desviación estándar es de 100 horas. Una muestra aleatoria de 64 focos indicó que la vida media de la muestra es de 350 horas. Calcule la estimación del intervalo de confianza para la media poblacional de vida de los focos de este embarque del

4.1) 95%

$325{,}5 \leq \mu \leq 374{,}5$

4.2) 98%

$320{,}875 \leq \mu \leq 379{,}125$

5) Las medias de los diámetros de una muestra aleatoria de 200 bolas de rodamientos producidas por una máquina es 0,824 cm. Se sabe la desviación estándar de la población es 0,042 cm. Calcule el intervalo de confianza a un 90%

$0{,}819 \leq \mu \leq 0{,}83$

6) Plantee y resuelva un ejercicio de aplicación similar a cualquiera de los anteriores.

3.2) ESTIMACIÓN DE INTERVALO DE CONFIANZA PARA LA MEDIA (σ DESCONOCIDA)

Así como la media poblacional μ suele ser desconocida, rara vez se conoce la desviación estándar real de la población σ. Por lo tanto, se requiere desarrollar una estimación del intervalo de confianza de μ usando sólo los estadísticos de muestra $\bar{x}$ y S.
Se emplea la siguiente fórmula:

$$\bar{x} - t_{n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1}\frac{S}{\sqrt{n}}$$

Donde t_{n-1} es el valor crítico de la distribución t con n-1 grados de libertad para un área de α/2 en la cola superior

La distribución t supone que la población está distribuida normalmente. Esta suposición es particularmente importante para n < 30. Pero cuando la población es finita y el tamaño de la muestra constituye más del 5% de la población, se debe usar el factor finito de corrección para modificar las desviaciones estándar. Por lo tanto si cumple:

$$\frac{n}{N} \cdot 100\% > 5\%$$

Se aplica la ecuación

$$\bar{x} - t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Siendo N el tamaño de la población y n el tamaño de la muestra

Antes de seguir continuando es necesario estudiar la distribución *t de Student*, por lo que a continuación se presenta una breve explicación de esta distribución.

Al comenzar el siglo XX, un especialista en Estadística de la Guinness Breweries en Irlanda llamado William S. Gosset deseaba hacer inferencias acerca de la media cuando la σ fuera desconocida. Como a los empleados de Guinness no se les permitía publicar el trabajo de investigación bajo sus propios nombres, Gosset adoptó el seudónimo de "Student". La distribución que desarrolló se conoce como la distribución t de Student.

Si la variable aleatoria X se distribuye normalmente, entonces el siguiente estadístico tiene una distribución t con n - 1 grados de libertad.

$$t = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

Esta expresión tiene la misma forma que el estadístico Z en la ecuación para la distribución muestral de la media con la excepción de que S se usa para estimar la σ desconocida.

Entre las principales propiedades de la distribución t se tiene:

En apariencia, la distribución t es muy similar a la distribución normal estandarizada. Ambas distribuciones tienen forma de campana. Sin embargo, la distribución t tiene mayor área en los extremos y menor en el centro, a diferencia de la distribución normal. Puesto que el valor de σ es desconocido, y se emplea S para estimarlo, los valores t son más variables que los valores Z.
Los grados de libertad n - 1 están directamente relacionados con el tamaño de la muestra n. A medida que el tamaño de la muestra y los grados de libertad se incrementan, S se vuelve una mejor estimación de σ y la distribución t gradualmente se acerca a la distribución normal estandarizada hasta que ambas son virtualmente idénticas. Con una muestra de 120 o más, S estima σ con la suficiente precisión como

para que haya poca diferencia entre las distribuciones t y Z. Por esta razón, la mayoría de los especialistas en estadística usan Z en lugar de t cuando el tamaño de la muestra es igual o mayor de 30.

Como se estableció anteriormente, la distribución t supone que la variable aleatoria X se distribuye normalmente. En la práctica, sin embargo, mientras el tamaño de la muestra sea lo suficientemente grande y la población no sea muy sesgada, la distribución t servirá para estimar la media poblacional cuando σ sea desconocida.

Los grados de libertad de esta distribución se calculan con la siguiente fórmula

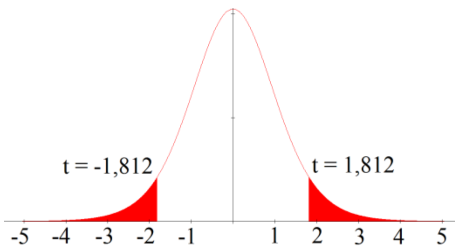
$$n-1$$

Donde n = tamaño de la muestra

Ejemplo: Imagínese una clase con 40 sillas vacías, cada uno elige un asiento de los que están vacíos. Naturalmente el primer alumno podrá elegir de entre 40 sillas, el segundo de entre 39, y así el número irá disminuyendo hasta que llegue el último alumno. En este punto no hay otra elección (grado de libertad) y aquel último estudiante simplemente se sentará en la silla que queda. De este modo, los 40 alumnos tienen 39 o n-1 grados de libertad.

Para leer en la tabla de la distribución t se procede de la siguiente manera:

TABLA Nº 4
DISTRIBUCIÓN t DE STUDENT



Ejemplos:
Para n-1 = 10 grados de libertad
$P(t > 1,812) = 0,05$
$P(t < -1,812) = 0,05$

α / n-1	0,25	0,2	0,15	0,1	**0,05**	0,025	0,01	0,005	0,0005
1	1,0000	1,3764	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	636,6192
2	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	31,5991
3	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	12,9240
4	0,7407	0,9410	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8688
6	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	5,4079
8	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,6998	0,8791	1,0931	1,3722	**1,8125**	2,2281	2,7638	3,1693	4,5869
11	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370

Usted encontrará los valores críticos de t para los grados de libertad adecuados en la tabla para la distribución t. Las columnas de la tabla representan el área de la cola superior de la distribución t. Cada fila representa el valor t determinado para cada grado de libertad específico. Por ejemplo, con 10 grados de libertad, si se quiere un nivel de confianza del 90%, se encuentra el valor t apropiado como se muestra en la tabla. El nivel de confianza del 90% significa que el 5% de los valores (un área de 0,05) se encuentran en cada extremo de la distribución. Buscando en la columna para un área de la cola superior y en la fila correspondiente a 10 grados de libertad, se obtiene un valor crítico para t de 1.812. Puesto que t es una distribución simétrica con una media 0, si el valor de la cola superior es +1.812, el valor para el área de la cola inferior (0,05 inferior) sería -1.812. Un valor t de -1.812 significa que la probabilidad de que t sea menor a -1.812, es 0,05, o 5% (vea la figura).

Ejemplos ilustrativos:

1) Determinar el valor crítico de t con lectura en la tabla, Excel y Winstats en cada una de las siguientes condiciones para $1-\alpha = 0{,}95\ ; n = 13$

Solución:

Con lectura en la tabla

$Si\ 1-\alpha = 0{,}95 \Rightarrow \alpha = 1-0{,}95 = 0{,}05$

Para leer en la tabla se necesita calcular el área de una cola, la cual es:

$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

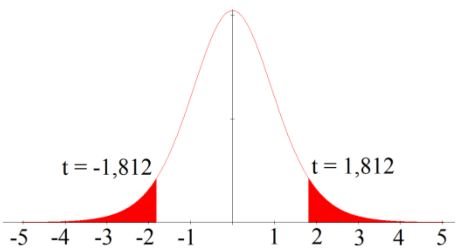
O también el área de una cola se calcula de la siguiente manera:

$$\frac{\alpha}{2} = \frac{1-(1-\alpha)}{2} \Rightarrow \frac{\alpha}{2} = \frac{1-0{,}95}{2} = 0{,}025$$

Calculando los grados de libertad se tiene:

$n-1 = 13-1 = 12$

TABLA Nº 4
DISTRIBUCIÓN t DE STUDENT



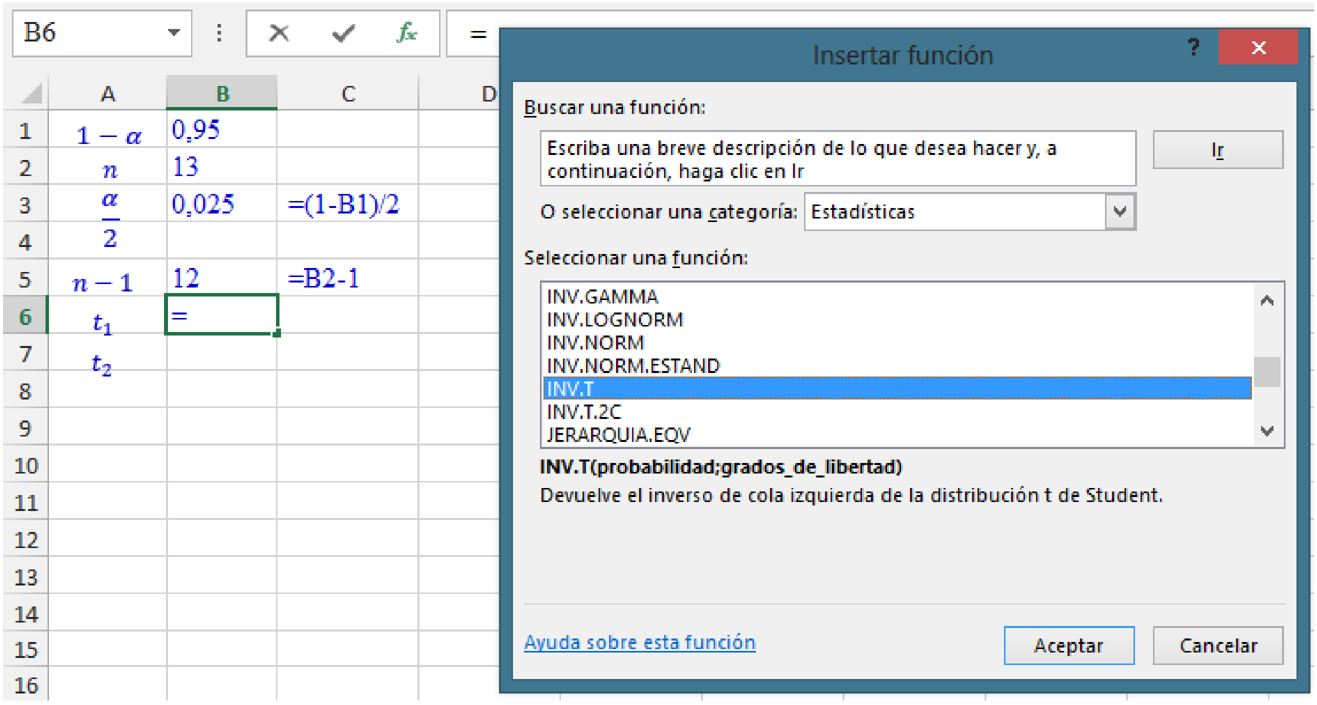
Ejemplos:
Para n-1 = 10 grados de libertad
$\boldsymbol{P(t \geq 1{,}812) = 0{,}05}$
$\boldsymbol{P(t \leq -1{,}812) = 0{,}05}$

α / n-1	0,25	0,2	0,15	0,1	0,05	**0,025**	0,01	0,005	0,0005
1	1,0000	1,3764	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	636,6192
2	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	31,5991
3	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	12,9240
4	0,7407	0,9410	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8688
6	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	5,4079
8	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,6998	0,8791	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5869
11	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370
12	0,6955	0,8726	1,0832	1,3562	1,7823	**2,1788**	2,6810	3,0545	4,3178
13	0,6938	0,8702	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2208

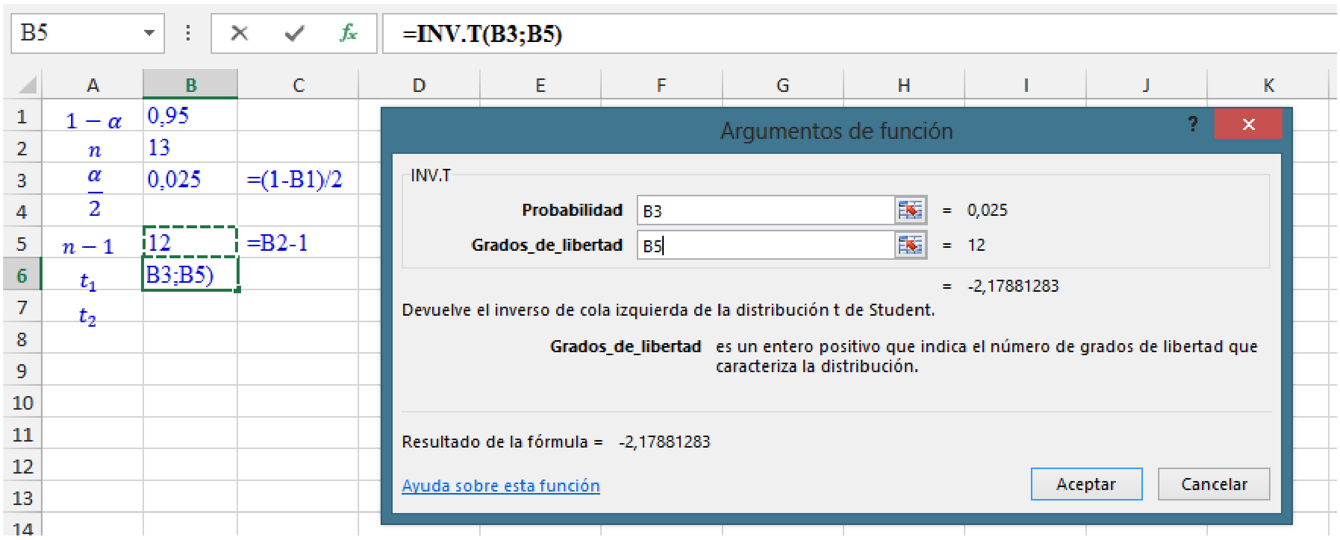
En la tabla con 12 grados de libertad y 0,025 de área se obtiene un valor de t =2,1788, y por simetría es igual también a t = -2,1788

Para realizar los cálculos en Excel se procede de la siguiente manera:

a) Llenar los datos y hacer los cálculos del área de una cola y de los grados de libertad. Luego insertar función. En la casilla seleccionar una categoría, seleccionar Estadísticas. Seleccionar la función INV.T.



b) Clic en Aceptar. En la ventana Argumentos de la función, en Probabilidad seleccionar B3, y en Grados de libertad seleccionar B6.

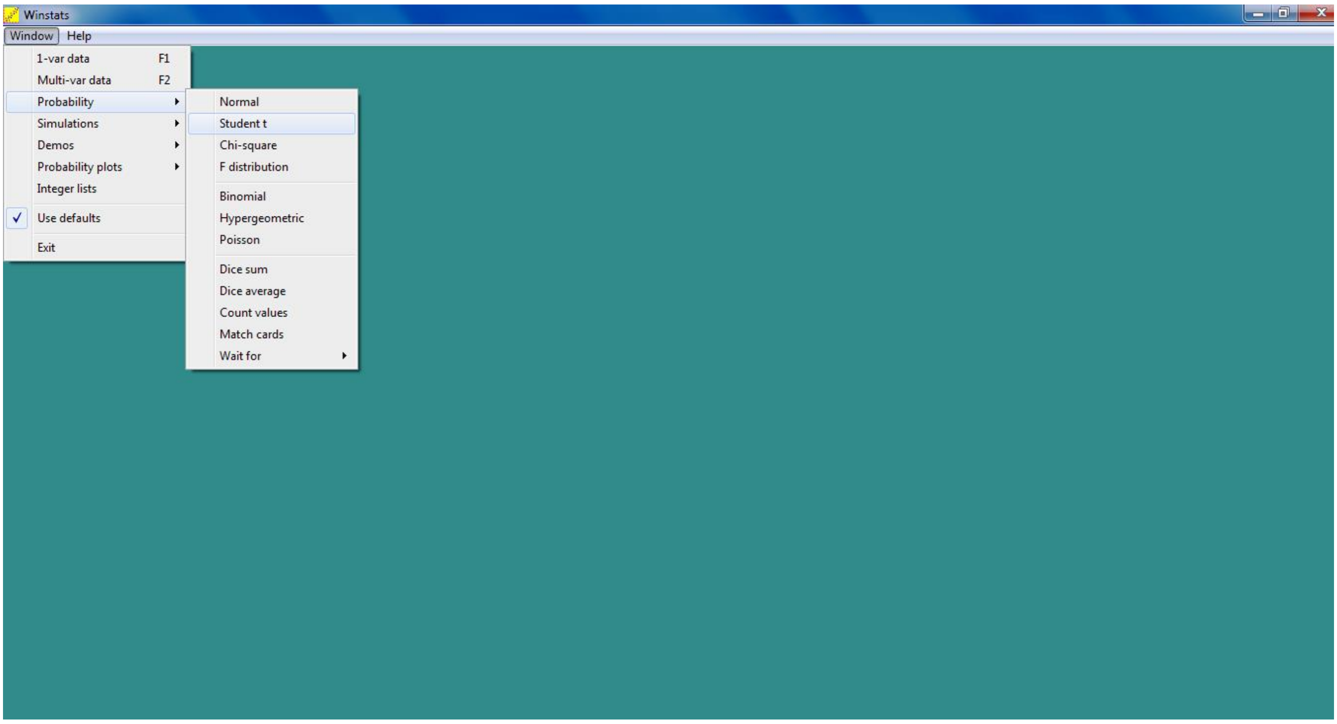


c) Clic en Aceptar. Los demás cálculos se muestran en la siguiente figura:

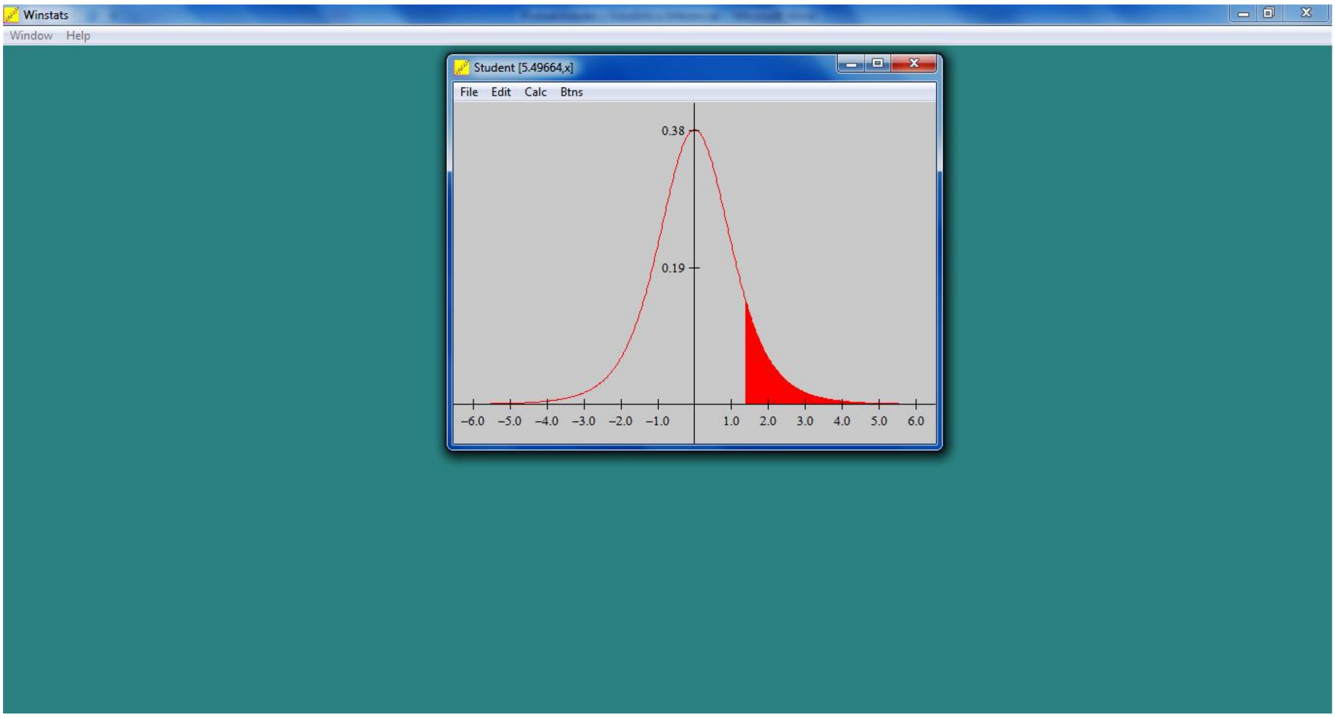
	A	B	C
1	$1-\alpha$	0,95	
2	n	13	
3	$\frac{\alpha}{2}$	0,025	=(1-B1)/2
4			
5	$n-1$	12	=B2-1
6	t_1	-2,1788	=INV.T(B3;B5)
7	t_2	2,17881	=B6*-1

Para resolver con Winstats se procede de la siguiente manera:

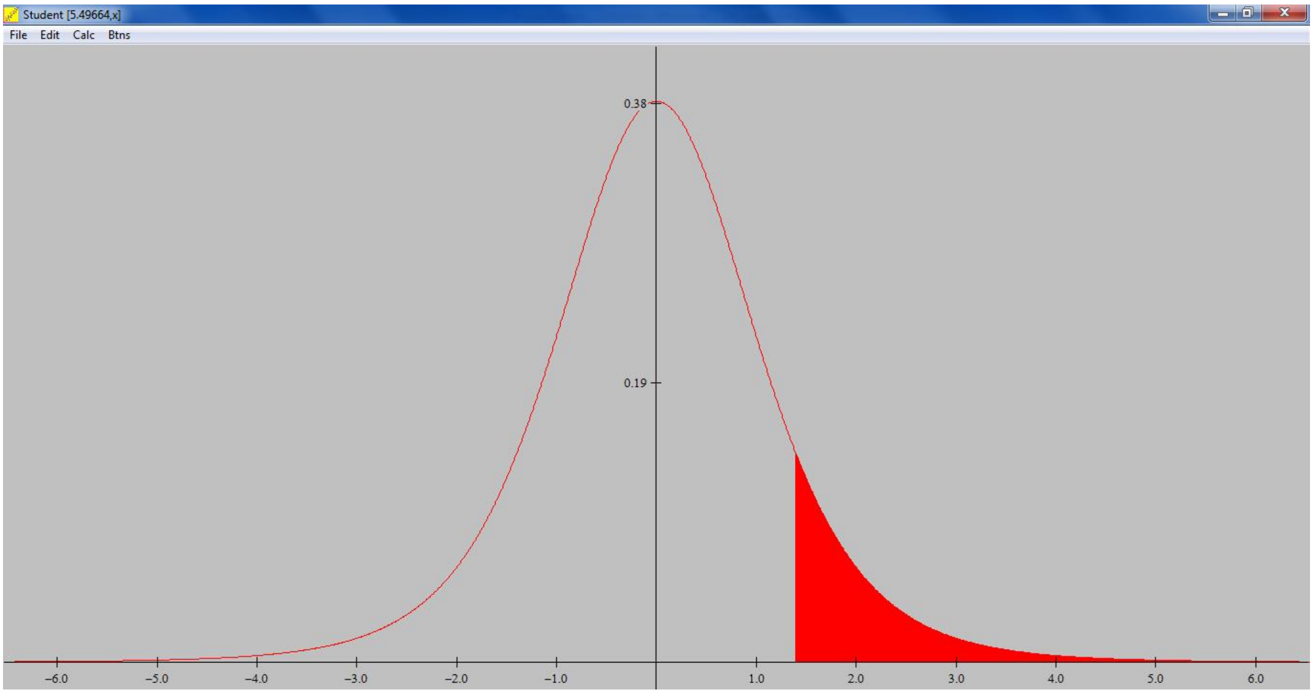
a) Clic en Window y luego en Probability seleccionar Student t



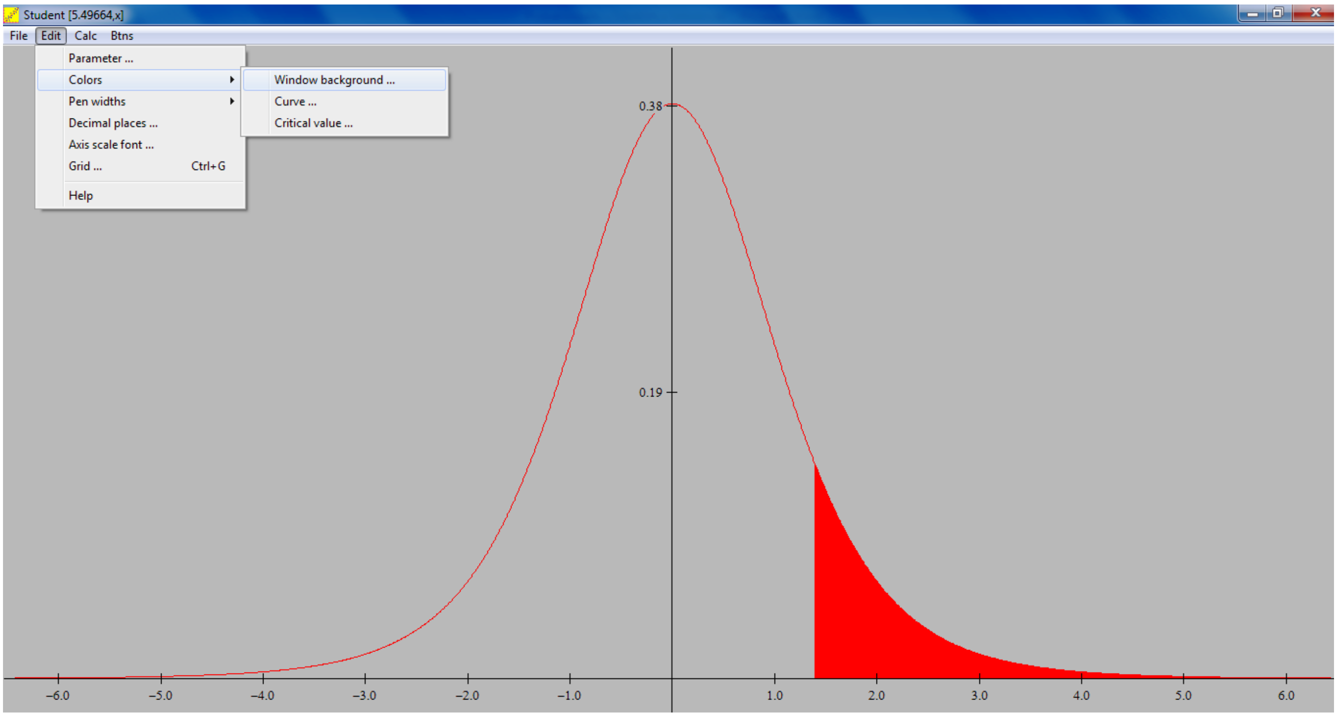
b) Clic en Student t



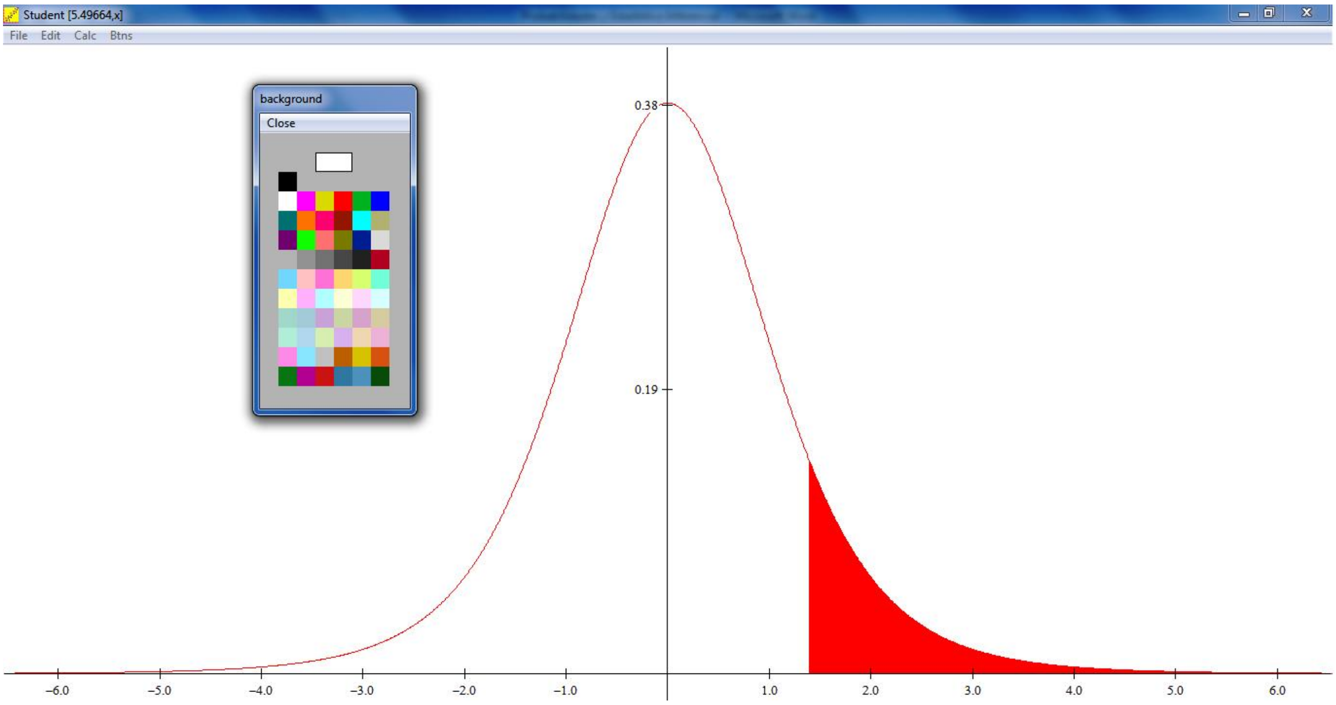
c) Maximizar la ventana de la distribución



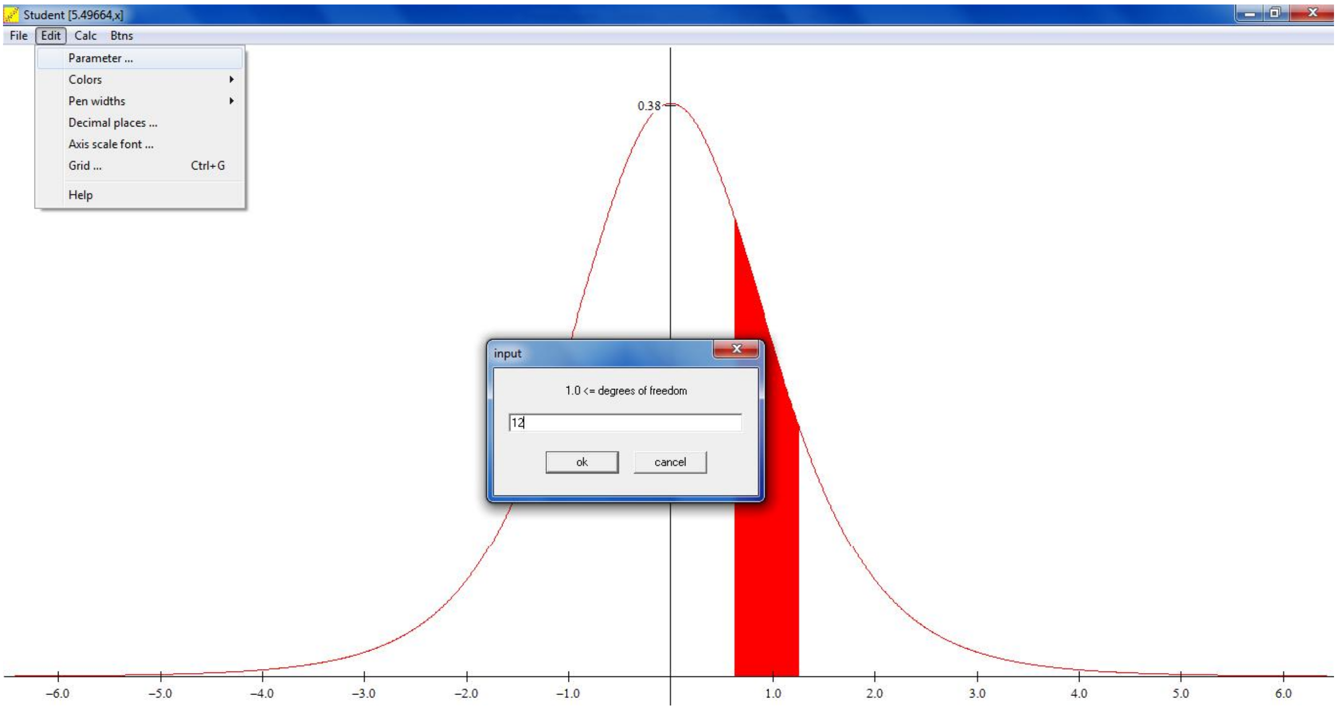
d) Para cambiar el color del fondo, clic en Edit + Colors + Window background



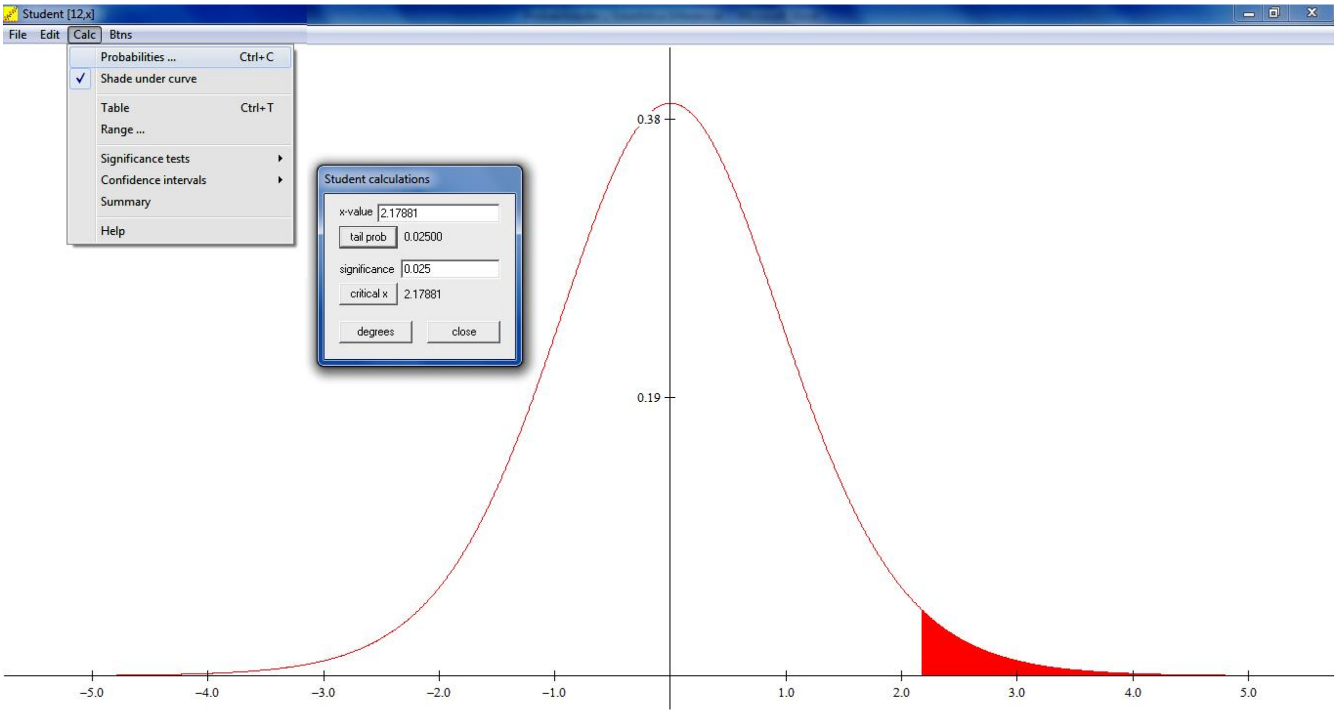
e) Clic Window background. En la venta de background seleccionar el color deseado, que este caso se seleccionó el color blanco. Luego clic en Close para cerrar la venta background.



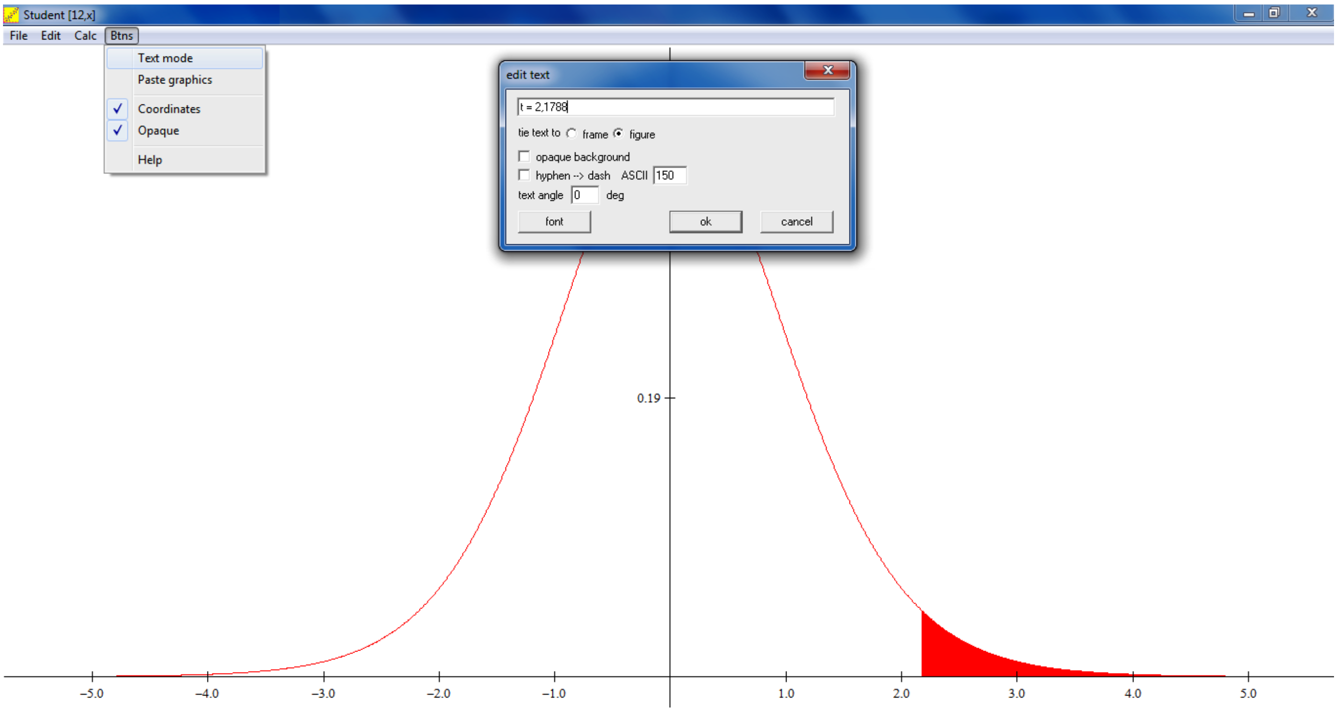
f) Para editar lo grados de libertad, clic en Edit + Parameter...(Parámetros). Clic en Parámetros. En la casilla de la ventana input escribir 12. Clic en ok



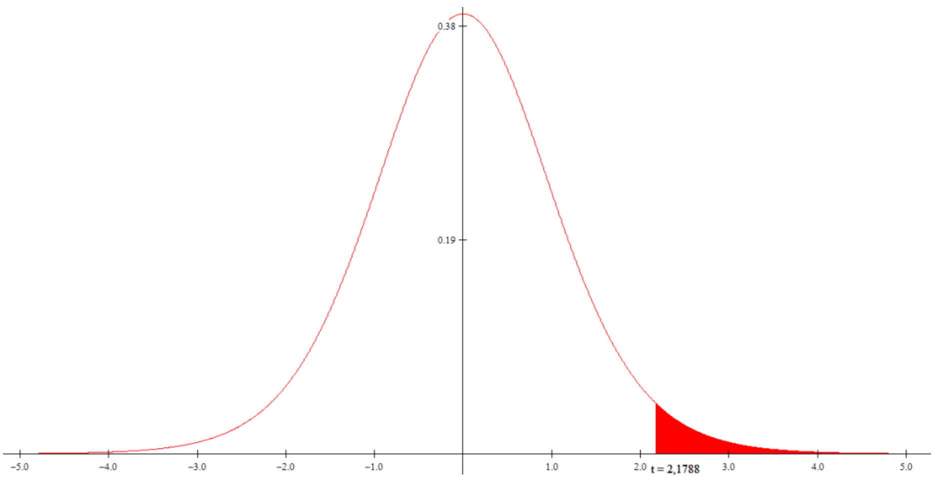
g) Para calcular el valor crítico de t, clic en Calc + Probabilities. En la ventana Student calculations, en significanse escribir 0,025 y luego clic en critical x. Clic en close para cerrar la ventana Student calculations.



h) Para escribir textos, clic en Btns. Luego clic derecho en cualquier parte de la ventana y aparece la ventana edit text. En la casilla de la ventana edit text escribir el texto deseado.

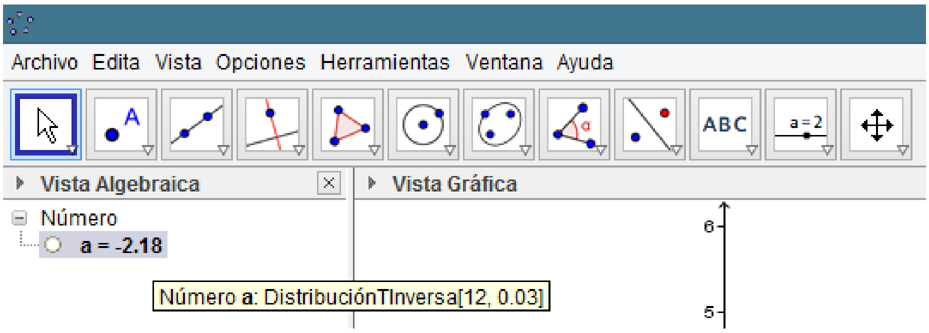


i) Clic en ok de la ventana edit text. Luego arrastar con el mause el texto al lugar deseado



En GeoGebra

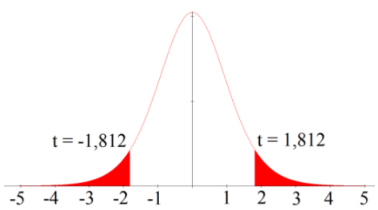
Seleccione DistribuciónTInversa[<Grados de Libertad>, <Probabilidad>]



2) Sea X = $t_{(10)}$ hallar el valor de P(X ≤ -1,3722) + P(X ≥ 2,7638) con lectura en la tabla, Excel y Winstats

Solución:

TABLA Nº 4
DISTRIBUCIÓN t DE STUDENT



Ejemplos:
Para n-1 = 10 grados de libertad
$\boldsymbol{P(t \geq 1,812) = 0,05}$
$\boldsymbol{P(t \leq -1,812) = 0,05}$

α / n-1	0,25	0,2	0,15	**0,1**	0,05	0,025	**0,01**	0,005	0,0005
1	1,0000	1,3764	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	636,6192
2	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	31,5991
3	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	12,9240
4	0,7407	0,9410	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8688
6	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	5,4079
8	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,6998	0,8791	1,0931	**1,3722**	1,8125	2,2281	**2,7638**	3,1693	4,5869
11	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370

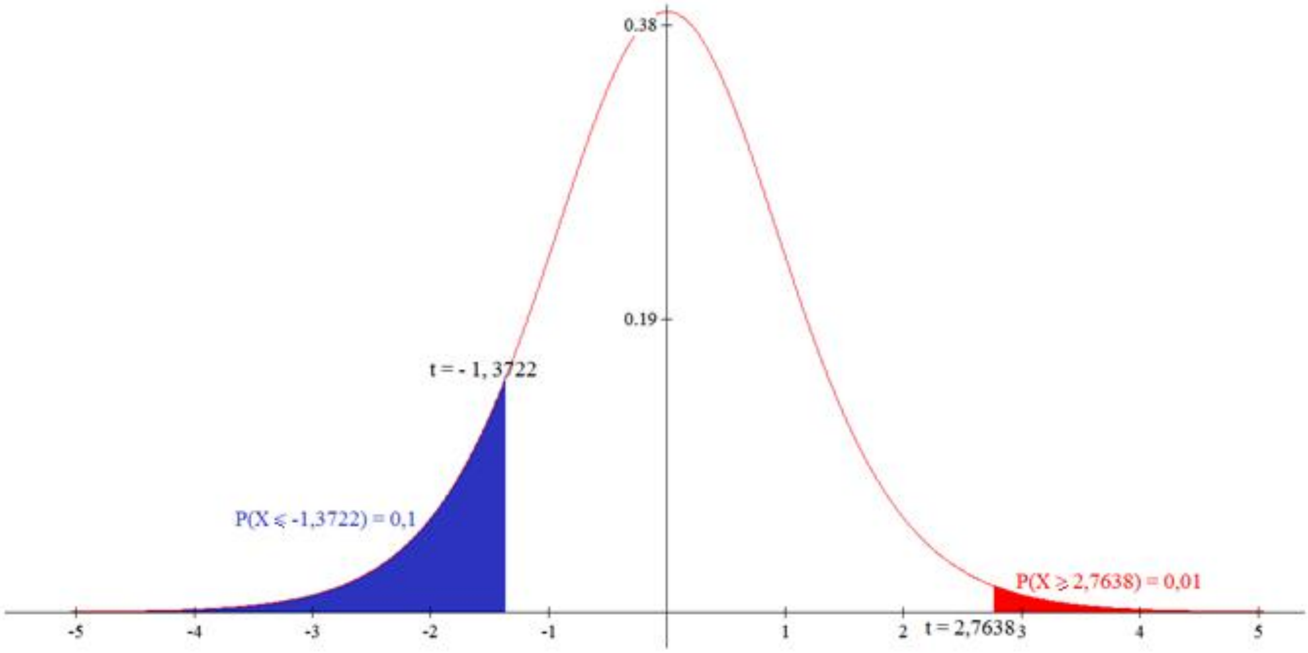
Con lectura en la tabla se obtiene:
P(X ≤ -1,3722) = 0,1 y P(X ≥ 2,7638) = 0,01

Entonces:
P(X<-1,3722) + P(X> 2,7638) = 0,1 + 0,01 = 0,11

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	t_1	-1,3722			
2	t_2	2,7638			
3	$n-1$	10			
4	$P(t_1 < -1{,}3722)$	0,1	=DISTR.T.N(B1;B3;VERDADERO)		
5	$P(t_2 > 2{,}7638)$	0,01	=DISTR.T.CD(B2;B3)		
6	$P(X < t_1) + P(X > t_2)$	0,11	=B4+B5		

El gráfico elaborado en Winstats y Paint se muestra en la siguiente figura:



3) Un fabricante de papel para computadora tiene un proceso de producción que opera continuamente a lo largo del turno. Se espera que el papel tenga una media de longitud de 11 pulgadas. De 500 hojas se selecciona una muestra de 29 hojas con una media de longitud del papel de 10,998 pulgadas y una desviación estándar de 0,02 pulgadas. Calcular la estimación del intervalo de confianza del 99%

Solución:
Los datos del problema son:
$\mu = 11$
$N = 500$
$n = 29$
$\bar{x} = 10{,}998$
$S = 0{,}02$
$Confianza = 99\%$

Como en los datos aparece el tamaño de la población, se debe verificar si el tamaño de la nuestra es mayor que el 5% para emplear la fórmula con el factor finito de corrección. Se remplaza valores en la siguiente fórmula:

$$\frac{n}{N}\cdot 100\% > 5\% \Longrightarrow \frac{29}{500}\cdot 100\% = 5{,}8\%$$

Por lo tanto se debe utilizar la fórmula con el factor finito de corrección.

Calculando la proporción de la cola superior e inferior de la distribución se obtiene:

$$Nivel\ de\ confianza = (1-\alpha)\cdot 100\%$$

$$\frac{\alpha}{2} = \frac{100\% - Nivel\ de\ confianza}{200} \Longrightarrow \frac{\alpha}{2} = \frac{100\% - 99\%}{200} = 0{,}005$$

Calculando los grados de libertad se obtiene:

$$n-1 = 29-1 = 28$$

Con lectura en la tabla para un área de 0,005 y 28 grados de libertad se obtiene t = ±2,7633

Remplazando valores y realizando lo cálculos se obtiene:

$$\bar{x} - t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \le \mu \le \bar{x} + t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$10{,}998 - 2{,}7633\frac{0{,}02}{\sqrt{29}}\sqrt{\frac{500-29}{500-1}} \le \mu \le 10{,}998 + 2{,}7633\frac{0{,}02}{\sqrt{29}}\sqrt{\frac{500-29}{500-1}}$$

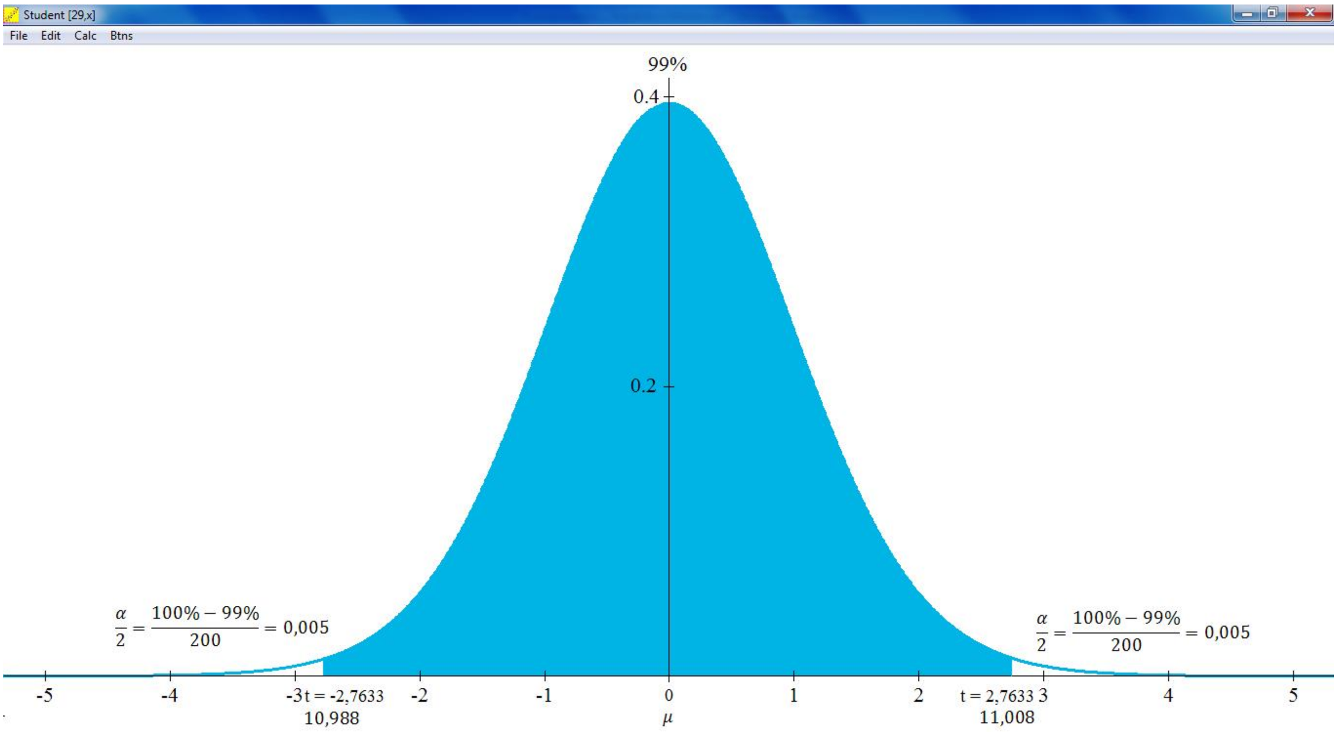
$$10{,}988 \le \mu \le 11{,}008$$

Los cálculos en Excel se muestran en la siguiente figura:

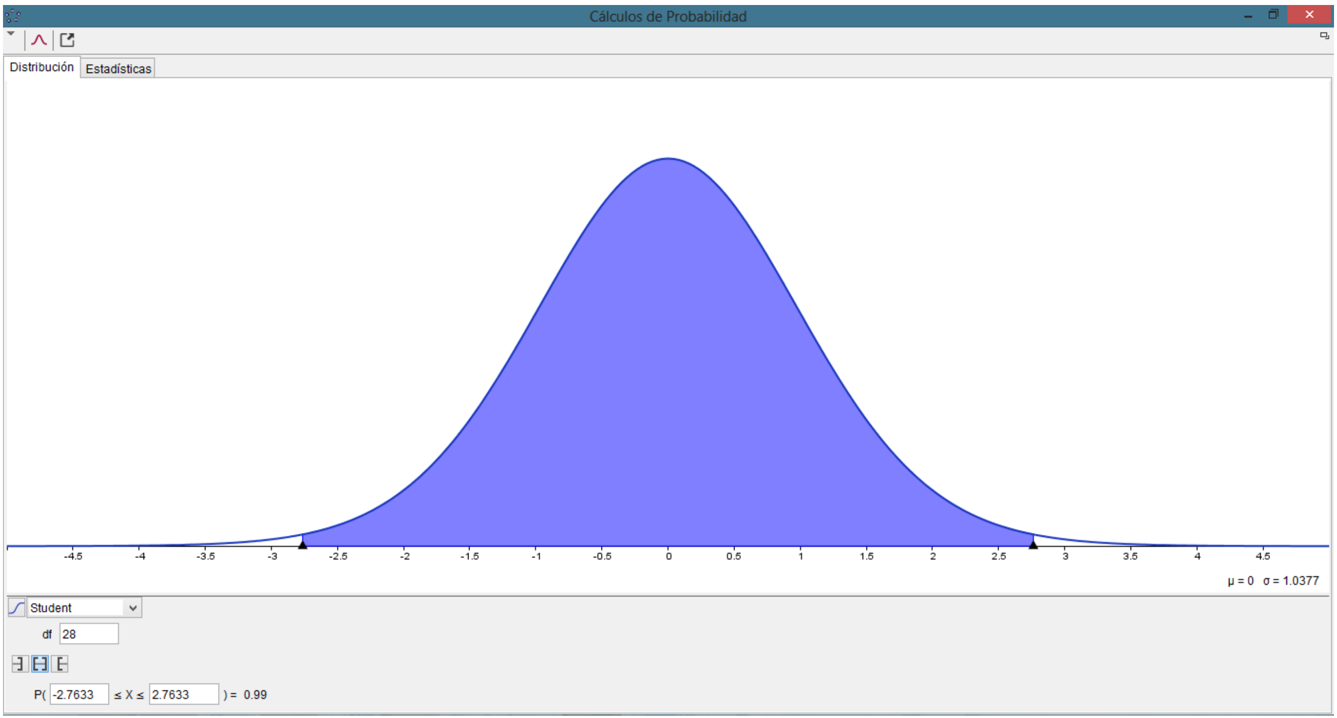
	A	B	C	D	E	F
1	μ	11				
2	N	500				
3	n	29				
4	$\bar{X}$	10,998				
5	S	0,02				
6	Confianza	99				
7	$\frac{n}{N}\cdot 100 > 5\%$	5,8	=(B3/B2)*100			
8						
9	α	0,01	=(100-B6)/100			
10	$t_{n-1}\frac{S}{\sqrt{n}}$	0,0103	=INTERVALO.CONFIANZA.T(B9;B5;B3)			
11						
12						
13	$\bar{x} - t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \le \mu \le \bar{x} + t_{n-1}\frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$					
14						
15	10,988	$\le \mu \le$	11,008			
16	=B4-B10*RAIZ((B2-B3)/(B2-1))		=B4+B10*RAIZ((B2-B3)/(B2-1))			

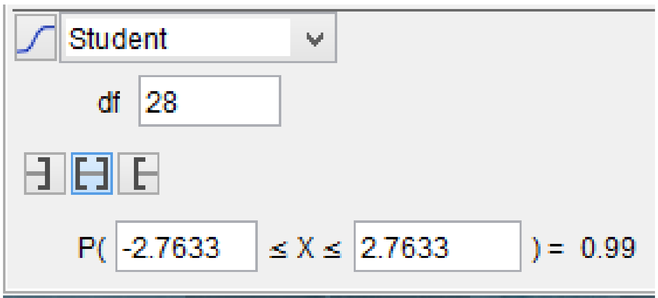
Interpretación: Existe un 99% de confianza de que la media poblacional se encuentra entre 10,998 y 11,008

El gráfico elaborado en Winstats y Paint se muestra en la siguiente figura:



El gráfico elaborado en GeoGebra se muestra en la siguiente figura:





TAREA DE INTERAPRENDIZAJE Nº 14

1) ¿Cuál fue el origen del nombre de la distribución t de Student
2) Describa las diferencias entre la distribución t y la distribución normal
3) ¿En qué caso la distribución t es virtualmente idéntica a la distribución normal?
4) Defina con sus propias palabras el concepto de grados de libertad. Ilustre con un ejemplo
5) Realice un organizador gráfico (cuadro sinóptico, mapa conceptual, etc.) sobre la distribución t de Student
6) Determinar el valor crítico de t con lectura en la tabla, Excel y Winstats en cada una de las siguientes condiciones:

6.1) $1-\alpha = 0{,}95$; n = 10

2,2622

6.2) $1-\alpha = 0{,}99$; n = 10

3,2498

7) Sea $X = t_{(10)}$. Calcule el valor de las siguientes probabilidades con lectura en la tabla, Excel y Winstats

7.1) $P(X \leq -1{,}8125) + P(X \geq 1{,}8125)$

0,1

7.2) $P(X \leq -1{,}3722) + P(X \geq 1{,}3722)$

0,2

8) Sea $X = t_{(10)}$ .Calcule el valor de las siguientes probabilidades con lectura en la tabla, Excel y GeoGebra

8.1) $P(X\leq 0{,}6998)$

0,75

8.2) $P(-1{,}0931\leq X\leq 2{,}7638)$

0,84

8.3) $P(-0{,}6998\leq X\leq 4{,}5869)$

0,75

9) Si $\bar{X} = 50$, S = 15, n=16, construya una estimación del intervalo de confianza del 99% de la media poblacional μ con lectura en la tabla, Excel y Winstats.

$38{,}95 \leq \mu \leq 61{,}05$

10) Si $\bar{X} = 30$, S = 6, n=15 construya una estimación del intervalo de confianza del 99,9% de la media poblacional μ con lectura en la tabla, Excel y GeoGebra.

$23{,}59 \leq \mu \leq 36{,}41$

11) Determine el intervalo de confianza de 95% con lectura en la tabla, Excel y Winstats para los casos siguientes:

11.1) Si $\bar{X} = 15$, S =2, n=16 y N= 200

$13{,}975 \leq \mu \leq 16{,}025$

11.2) Si $\bar{X} = 14$, S =2, n=16 y N= 150

$12{,}975 \leq \mu \leq 15{,}025$

14) Una empresa manufacturera produce aislantes eléctricos. Si los aislantes se rompen al usarse, muy posiblemente tendremos un corto circuito. Para probar la fuerza de los aislantes, se lleva a cabo una prueba destructiva para determinar cuánta fuerza se requiere para romperlos. Se mide la fuerza observando cuántas libras se aplican al aislante antes de que se rompa. La siguiente tabla lista 30 valores de este experimento de la fuerza en libras requerida para romper el aislante. Construya una

estimación del intervalo de confianza del 95% para la población media de fuerza requerida para romper al aislante empleando Excel y Winstats.

1870	1728	1656	1610	1634	1784	1522	1696	1592	1662
1866	1764	1734	1662	1734	1774	1550	1756	1762	1866
1820	1744	1788	1688	1810	1752	1680	1810	1652	1736

$1689{,}96 \leq \mu \leq 1756{,}84$

15) Plantee y resuelva un problema similar al anterior empleando Excel y GeoGebra

3.3) ESTIMACIÓN DEL INTERVALO DE CONFIANZA PARA UNA PROPORCIÓN

Sirve para calcular la estimación de la proporción de elementos en una población que tiene ciertas características de interés. La proporción desconocida de la población, se representa con la letra griega π. La estimación puntual para π es la proporción de la muestra, $p = \frac{X}{n}$, donde n es el tamaño de la muestra y X es el número de elementos en la muestra que tienen la característica de interés. La siguiente ecuación define la estimación del intervalo de confianza para la proporción de la población.

$$p - Z\sqrt{\frac{p(1-p)}{n}} \leq \pi \leq p + Z\sqrt{\frac{p(1-p)}{n}}$$

Donde:

$p = proporción\ de\ la\ muestra = \frac{X}{n} = \frac{número\ de\ elementos\ con\ característica\ de\ interés}{tamaño\ de\ la\ muestra}$

$\pi = proporción\ de\ la\ población$

$Z = valor\ crítico\ para\ la\ distribución\ normal\ estandarizada$

$n = tamaño\ de\ la\ muestra$

Cuando la población es finita (N) y el tamaño de la muestra (n) constituye más del 5% de la población, se debe usar el factor finito de corrección. Por lo tanto si cumple:

$$\frac{n}{N} \cdot 100\% > 5\%$$

Se aplica la ecuación

$$p - Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} \leq \pi \leq p + Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$

Ejemplo ilustrativo

En un almacén se está haciendo una auditoria para las facturas defectuosas. De 500 facturas de venta se escoge una muestra de 30, de las cuales 5 contienen errores. Construir una estimación del intervalo de confianza del 95%.

Solución:

Los datos del problema son:

$N = 500$

$n = 30$

$X = 5$

$Confianza = 95\%$

Como en los datos aparece el tamaño de la población, se debe verificar si el tamaño de la nuestra es mayor que el 5% para emplear la fórmula con el factor finito de corrección. Se remplaza valores en la siguiente fórmula:

$$\frac{n}{N}\cdot 100\% > 5\% \Longrightarrow \frac{30}{500}\cdot 100\% = 6\%$$

Por lo tanto se debe utilizar la fórmula con el factor finito de corrección.

Calculando la proporción de la cola superior e inferior de la distribución se obtiene:

$$Nivel\ de\ confianza = (1-\alpha)\cdot 100\%$$

$$\frac{\alpha}{2} = \frac{100\% - Nivel\ de\ confianza}{200} \Longrightarrow \frac{\alpha}{2} = \frac{100\% - 95\%}{200} = 0{,}025$$

Con lectura en la tabla de la distribución normal para un área de 0,025 se obtiene Z = -1,96, y por simetría Z =1,96

Calculando la proporción de la muestra se obtiene:

$$p = \frac{X}{n} = \frac{5}{30} = 0{,}167$$

Calculando el intervalo de confianza se obtiene:

$$p - Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} \leq \pi \leq p + Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$

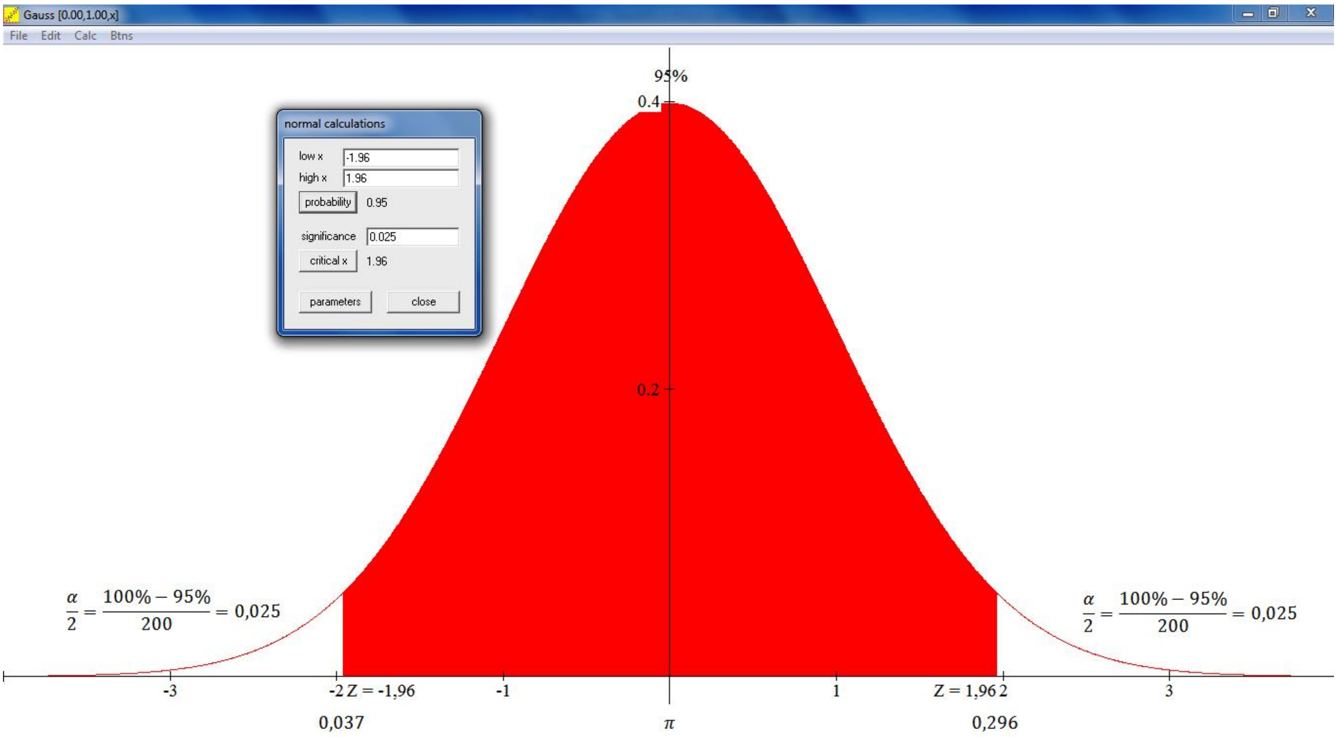
$$0{,}167 - 1{,}96\sqrt{\frac{0{,}167(1-0{,}167)}{30}}\sqrt{\frac{500-30}{500-1}} \leq \pi \leq 0{,}167 + 1{,}96\sqrt{\frac{0{,}167(1-0{,}167)}{30}}\sqrt{\frac{500-30}{500-1}}$$

$$0{,}037 \leq \pi \leq 0{,}296$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H
1	N	500						
2	n	30						
3	X	5						
4	Confianza	95						
5	$\frac{n}{N}\cdot 100\% > 5\%$	6						
6								
7	$\frac{\alpha}{2}$	0,025						
8								
9	Z	-1,96	=INV.NORM.ESTAND(B7)					
10	Z	1,96	=B9*-1					
11	$p = \frac{X}{n}$	0,167	=B3/B2					
12								
13	$p - Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} \leq \pi \leq p + Z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$							
14								
15		0,037	$\leq \mu \leq$	0,296				
16	=B11-B10*RAIZ(B11*(1-B11)/B2)*RAIZ((B1-B2)/(B1-1))			=B11+B10*RAIZ(B11*(1-B11)/B2)*RAIZ((B1-B2)/(B1-1))				

El gráfico elaborado en Winstats y Paint se muestra en la siguiente figura:



TAREA DE INTERAPRENDIZAJE Nº 15

1) Si n = 200 y X = 50, construya una estimación del intervalo de confianza con lectura en la tabla, Excel y Winstats para la proporción de la población con confianza del

1.1) 95%

$0{,}19 \leq \pi \leq 0{,}31$

1.2) 99%

$0{,}171 \leq \pi \leq 0{,}3$

2) Si n = 400 y X = 25, construya una estimación del intervalo de confianza con lectura en la tabla, Excel y Winstats para la proporción de la población con confianza del

2.1) 90%

$0{,}0426 \leq \pi \leq 0{,}0824$

2.2) 99%

$0{,}031 \leq \pi \leq 0{,}094$

3) En un almacén se está haciendo una auditoria para las facturas defectuosas. Se escoge una muestra de 100 facturas de venta, 10 de ellas contienen errores.

3.1) Construir una estimación del intervalo de confianza del 95% para la proporción de la población con lectura en la tabla, Excel y Winstats

$0{,}041 \leq \pi \leq 0{,}159$

3.2) ¿Cómo interpreta el resultado obtenido en el inciso 3.1?

4) Plantee y resuelva un ejercicio similar al anterior con lectura en la tabla, Excel y Winstats

5) El editor de un periódico desea estimar la proporción de periódicos impresos con algún defecto, tal como borraduras en exceso, disposición errónea de las hojas, páginas faltantes o duplicadas. Se selecciona una muestra aleatoria de 200 periódicos, 35 de ellos contienen algún tipo de defecto. Realice e interprete un intervalo de confianza del 90% para la proporción de periódicos impresos que tienen defectos con lectura en la tabla, Excel y Winstats.

$13,1\% \leq \pi \leq 21,9\%$

6) Una empresa telefónica desea estimar la proporción de hogares en los que se contrataría una línea telefónica adicional. Se seleccionó una muestra aleatoria de 500 hogares. Los resultados indican que a un costo reducido, 135 de los hogares contratarían una línea telefónica adicional. Construya e interprete una estimación del intervalo de confianza del 99% de la proporción poblacional de hogares que contratarían una línea telefónica adicional con lectura en la tabla, Excel y Winstats.

$21,9\% \leq \pi \leq 32,1\%$

7) Plantee y resuelva un ejercicio similar al anterior con lectura en la tabla, Excel y GeoGebra.

8) Miles de ecuatorianos organizan sus planes de viaje al exterior por Internet. En una encuesta reciente, se reportó que el 25% compra boletos de avión en Internet. Suponga que la encuesta se basó en 180 ecuatorianos que respondieron.

8.1) Construya una estimación del intervalo de confianza del 95% para la proporción poblacional de ecuatorianos que compran boletos de avión en Internet con lectura en la tabla, Excel y Winstats.

$18,7\% \leq \pi \leq 31,3\%$

8.2) Construya una estimación del intervalo de confianza del 90% para la proporción poblacional de ecuatorianos que compran boletos de avión en Internet con lectura en la tabla, Excel y Winstats.

$19,7\ \% \leq \pi \leq 30,3\ \%$

8.3) ¿Cuál intervalo es más amplio?. Explique por qué esto es cierto

9) Se hace una encuesta entre mujeres trabajadoras en Ecuador. De 1000 mujeres encuestadas, el 55% piensa que las empresas deben reservar los puestos de trabajo durante seis meses o menos para aquellas con permiso de maternidad, y el 45% considera que deberían reservar sus puestos durante más de seis meses.

9.1) Con lectura en la tabla, Excel y Winstats construya un intervalo de confianza del 95% para la proporción de las mujeres trabajadoras de Ecuador quienes creen que las empresas deberían reservar los puestos de trabajo durante seis meses o menos para aquellas con permiso de maternidad.

$51,9\ \% \leq \pi \leq 58,1\ \%$

9.2) Con lectura en la tabla, Excel y Winstats construya un intervalo de confianza del 95% para la proporción de las mujeres trabajadoras de Ecuador quienes creen que las empresas deberían reservar los puestos de trabajo durante más de seis meses para aquellas con permiso de maternidad.

$41,9\ \% \leq \pi \leq 48,1\ \%$

9.3) ¿Cuál intervalo es más amplio?. Explique por qué sucede

10) Plantee y resuelva un problema similar al anterior con lectura en la tabla, Excel y GeoGebra

3.4) DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

En cada ejemplo de la estimación del intervalo de confianza, usted seleccionó un tamaño de muestra sin considerar la amplitud del intervalo de confianza. En el mundo de los negocios, determinar el tamaño adecuado de la muestra es un procedimiento complicado, ya que está sujeto a las restricciones de presupuesto, tiempo y cantidad aceptada de error de muestreo. Si por ejemplo usted desea estimar la media de la cantidad de ventas en dólares de la facturas de ventas o la proporción de facturas de ventas que contienen errores, debe determinar por anticipado cuán grande sería el error de muestreo para estimar cada uno de los parámetros. También debe determinar por anticipado el nivel del intervalo de confianza a usar al estimar el parámetro poblacional.

Recordemos los siguientes conceptos básicos

Población.- Llamado también universo o colectivo, es el conjunto de todos los elementos que tienen una característica común. Una población puede ser finita o infinita. Es *población finita* cuando está delimitada y conocemos el número que la integran, así por ejemplo: Estudiantes de la Universidad UTN. Es *población infinita* cuando a pesar de estar delimitada en el espacio, no se conoce el número de elementos que la integran, así por ejemplo: Todos los profesionales universitarios que están ejerciendo su carrera.

Muestra.- La muestra es un subconjunto de la población. Ejemplo: Estudiantes de 2do Semestre de la Universidad UTN.

Sus principales características son:

Representativa.- Se refiere a que todos y cada uno de los elementos de la población tengan la misma oportunidad de ser tomados en cuenta para formar dicha muestra.

Adecuada y válida.- Se refiere a que la muestra debe ser obtenida de tal manera que permita establecer un mínimo de error posible respecto de la población.

Para que una muestra sea fiable, es necesario que su tamaño sea obtenido mediante procesos matemáticos que eliminen la incidencia del error.

Elemento o individuo.- Unidad mínima que compone una población. El elemento puede ser una entidad simple (una persona) o una entidad compleja (una familia), y se denomina unidad investigativa.

A) DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA PARA LA MEDIA

Para desarrollar una fórmula que permita determinar el tamaño apropiado de la muestra para construir una estimación del intervalo de confianza para la media, recuerde la ecuación

$$\bar{X} - Z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z\frac{\sigma}{\sqrt{n}}$$

La cantidad sumada o sustraída de $\bar{X}$ es igual a la mitad de la amplitud del intervalo. Esta cantidad representa la cantidad de imprecisión en la estimación que resulta del error de muestreo. El error de muestreo e (también conocido como margen de error) se define como:

$$e = Z\frac{\sigma}{\sqrt{n}}$$

Al despejar n se obtiene el tamaño de muestra necesario para construir una estimación del intervalo de confianza adecuado para la media. “Adecuado” significa que el intervalo resultante tendrá una cantidad aceptable de error de muestreo.

El proceso de despejar n se muestra a continuación:

Elevando al cuadrado ambos miembros de la igualdad:

$$(e)^2 = \left(Z\frac{\sigma}{\sqrt{n}}\right)^2$$

Resolviendo la potencia:

$$e^2 = Z^2\frac{\sigma^2}{n}$$

Transponiendo n a la izquierda de la igualdad:

$$e^2 n = Z^2\sigma^2$$

Transponiendo e^2 a la derecha de la igualdad y ordenando se obtiene la fórmula buscada:

$$n = \frac{\sigma^2 Z^2}{e^2}$$

Donde:

σ = Desviación estándar de la población que rara vez conoce su valor. En algunas ocasiones es posible estimar la desviación estándar a partir de datos pasados. En otras situaciones, puede hacer una cuidadosa conjetura tomando en cuenta el rango y la distribución de la variable. Por ejemplo, si supone que hay una distribución normal, el rango es aproximadamente igual a 6σ (es decir, $\pm 3\sigma$ alrededor de la media). Generalmente cuando no se tiene su valor, suele utilizarse un valor constante de 0,5.

Z = Valor obtenido mediante niveles de confianza. Es un valor constante que, si no se tiene su valor, se lo toma en relación al 95% de confianza equivale a 1,96 (como más usual) o en relación al 99% de confianza equivale 2,58, valor que queda a criterio del investigador.

e = Límite aceptable de error muestral que, generalmente cuando no se tiene su valor, suele utilizarse un valor que varía entre el 1% (0,01) y 9% (0,09), valor que queda a criterio del encuestador, sin embargo, se aconseja emplear el 5% o el 1%, por ser valores que guardan relación con el 95% y 99% de confianza, respectivamente.

Cuando en los datos se tiene el tamaño de la población se utiliza la siguiente fórmula:

$$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2}$$

La fórmula anterior se obtiene de la fórmula de la estimación del intervalo de confianza para la media, la cual es:

$$\bar{X} - Z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + Z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

De donde el error es:

$$e = Z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

De esta fórmula del error de la estimación del intervalo de confianza para la media se despeja la n, para lo cual se sigue el siguiente proceso:

Elevando al cuadrado a ambos miembros de la fórmula se obtiene:

$$(e)^2 = \left(Z\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\right)^2 \Longrightarrow e^2 = Z^2\frac{\sigma^2}{n}\frac{N-n}{N-1}$$

Multiplicando fracciones y eliminando denominadores:

$$e^2 = \frac{Z^2\sigma^2(N-n)}{n(N-1)} \Longrightarrow e^2 n(N-1) = Z^2\sigma^2(N-n)$$

Eliminando paréntesis y transponiendo n a la izquierda:

$$e^2 nN - e^2 n = Z^2\sigma^2 N - Z^2\sigma^2 n \Longrightarrow e^2 nN - e^2 n + Z^2\sigma^2 n = Z^2\sigma^2 N$$

Factor común de n y Despejando n:

$$n(e^2N - e^2 + Z^2\sigma^2) = Z^2\sigma^2 N \Longrightarrow\ n = \frac{Z^2\sigma^2 N}{e^2N - e^2 + Z^2\sigma^2}$$

Sacando factor común e^2 y ordenando se obtiene la fórmula para calcular el tamaño de la muestra para la media cuando se conoce el tamaño de la población:

$$n = \frac{N\sigma^2Z^2}{(N-1)e^2 + \sigma^2Z^2}$$

B) DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA PARA LA PROPORCIÓN

Para determinar el tamaño de la muestra necesario para estimar la proporción poblacional (π), se utiliza un método similar al método para calcular la media poblacional. Al desarrollar el tamaño de la muestra para un intervalo de confianza para la media, el error de muestreo se define por:

$$e = Z\frac{\sigma}{\sqrt{n}}$$

Cuando estima la proporción, se remplaza $\sigma = \sqrt{\pi(1-\pi)}$. Así, el error de muestreo es:

$$e = Z\frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}} \Rightarrow e = Z\sqrt{\frac{\pi(1-\pi)}{n}}$$

Al despejar n se obtiene el tamaño de la muestra necesario para desarrollar la estimación del intervalo de confianza para la proporción. El tamaño de la muestra n es igual al cuadrado del valor Z multiplicado por la proporción poblacional π, multiplicado por 1 menos la proporción poblacional π, y dividido por el cuadrado del error de muestreo e

$$n = \frac{\pi(1-\pi)Z^2}{e^2}$$

Al determinar π se tiene dos alternativas. En muchas situaciones se cuenta con información anterior o experiencias relevantes que proporcionan una estimación cuidadosa de π. 0 también, si no se cuenta con información anterior o experiencias relevantes, se trata de proporcionar un valor para π que nunca subestime el tamaño de muestra necesario. Con referencia a la ecuación anterior, observe que la cantidad $\pi(1-\pi)$aparece en el numerador. Por eso, se requiere determinar el valor de π que haga la cantidad $\pi(1-\pi)$lo más grande posible. Cuando $\pi = 0{,}5$, el producto $\pi(1-\pi)$ logra su resultado máximo. Para mostrar esto, los valores de π junto con los productos que los acompañan de $\pi(1-\pi)$ son como sigue:

$$Cuando\ \pi = 0{,}9 \Rightarrow \pi(1-\pi) = 0{,}9(1-0{,}9) = 0{,}09$$
$$Cuando\ \pi = 0{,}7 \Rightarrow \pi(1-\pi) = 0{,}7(1-0{,}7) = 0{,}21$$
$$Cuando\ \pi = 0{,}5 \Rightarrow \pi(1-\pi) = 0{,}5(1-0{,}5) = 0{,}25$$
$$Cuando\ \pi = 0{,}3 \Rightarrow \pi(1-\pi) = 0{,}3(1-0{,}3) = 0{,}21$$
$$Cuando\ \pi = 0{,}1 \Rightarrow \pi(1-\pi) = 0{,}1(1-0{,}1) = 0{,}09$$

Por lo tanto, cuando no se tiene conocimiento previo o una estimación de la proporción poblacional π, se debería usar $\pi = 0{,}5$ para determinar el tamaño de la muestra. Esto produce el tamaño de muestra más grande posible y deriva en el mayor costo posible del muestreo.

Cuando en los datos se tiene el tamaño de la población se utiliza la siguiente fórmula:

$$n = \frac{N\pi(1-\pi)Z^2}{(N-1)e^2 + \pi(1-\pi)Z^2}$$

Ejemplos ilustrativo

Calcular el tamaño de la muestra de una población de 500 elementos con un nivel de confianza del 95%

Solución: Se tiene N=500, para el 95% de confianza Z = 1,96, y como no se tiene los demás valores se tomará $\sigma = 0{,}5$, y e = 0,05.

Remplazando valores de la fórmula se tiene:

$$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2} \Longrightarrow n = \frac{500 \cdot 0{,}5^2 \cdot 1{,}96^2}{(500-1)0{,}05^2 + 0{,}5^2 \cdot 1{,}96^2} = 217$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	N	500				
2	e	0,05				
3	σ	0,5				
4	Confianza	95				
5	Área a la izquierda de -Z	0,025	=(100-B4)/200			
6	-Z	-1,96	=INV.NORM.ESTAND(B5)			
7	Z	1,96	=B6*-1			
8	$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2}$	217,49	=B1*B3^2*B7^2/((B1-1)*B2^2+B3^2*B7^2)			
9						

TAREA DE INTERAPRENDIZAJE Nº 16

1) Realice un organizador gráfico sobre el tamaño de la muestra

2) Proponga 3 ejemplos de población, de muestra y de elemento.

3) Contestar

3.1) ¿Qué significa “adecuado” para una estimación del intervalo de confianza que permite calcular el tamaño de la muestra?

3.2) ¿A partir de qué ecuación se obtiene el error de muestreo para la media cuando se conoce el tamaño de la población?

3.3) ¿Con qué otro nombre se le conoce al error de muestreo?

3.4) ¿Por qué se emplea la distribución Z en lugar de la distribución t para determinar el nivel de confianza deseado del tamaño de la muestra?

3.5) ¿Qué valores suelen utilizarse para Z y e cuando no se conocen sus valores?

3.6) ¿Por qué cuando no se tiene conocimiento previo o una estimación de la proporción poblacional π se debe empelar $\pi = 0{,}5$?

4) Calcular el valor de σ si se sabe $\pi = 0{,}5$

0,5

5) Escriba el proceso para despejar n de la siguiente fórmula para obtener la fórmula que permite calcular el tamaño de la muestra

$$e = Z\sqrt{\frac{\pi(1-\pi)}{n}}$$

6) Siguiendo el razonamiento del ejercicio anterior, escriba el proceso para despejar n de la fórmula del error de muestreo que permita obtener la siguiente fórmula:

$$n = \frac{N\pi(1-\pi)Z^2}{(N-1)e^2 + \pi(1-\pi)Z^2}$$

7) Un departamento de la dirección aérea civil del Ecuador define que el vuelo de una línea llega “a tiempo” si aterriza menos de 15 minutos después del tiempo programado en el Sistema de Reservación Computarizado. Los vuelos cancelados y desviados se consideran con retraso. Un estudio de las aerolíneas encontró que Icaro tiene la menor proporción de retrasos. Suponga que se le pide realizar un estudio de seguimiento para Icaro cuyo objetivo sea actualizar la proporción estimada de retrasos. ¿Qué tamaño de muestra deberá emplear para estimar la proporción poblacional con un 95% de nivel de confianza y con un error de ± 0,06 ?. Calcule con lectura en la tabla, Excel y Winstats

267

8) Con lectura en la tabla, Excel y Winstats calcular el tamaño de muestra con un error de muestreo de ±0,05 para N =500, si se desea tener un nivel de confianza en la estimación de la media poblacional del:

8.1) 95%

218

8.2) 99 %

285

9) Cree y resuelva con Excel y Winstats y ejercicio similar al anterior.

10) Consulte en la biblioteca o en el internet sobre la importancia del cálculo del tamaño de la muestra, y presente la consulta a través de un organizador gráfico.

CAPÍTULO IV
PRUEBA DE HIPÓTESIS

RESULTADOS DE APRENDIZAJE DEL CAPÍTULO

Al finalizar el presente capítulo el lector podrá evidenciar que:

- ✓ Interpreta las definiciones, características, propiedades y aplicaciones de las hipótesis para toma de decisiones.
- ✓ Aplica algoritmos para comprobar hipótesis según sus distintas opciones de manera manual, empleando Excel, Winstats y GeoGebra.
- ✓ Plantea y resuelve ejercicios y problemas de aplicación sobre prueba de hipótesis de manera manual, utilizando Excel, Winstats y GeoGebra.

CONTENIDOS

- ✓ Prueba de hipótesis para medias: De una y dos muestras
- ✓ Análisis de varianza: Estimación interna, intermediante y la razón F de Fisher
- ✓ Prueba de hipótesis para proporciones: De una, dos, k muestras (χ^2) y bondad de Ajuste χ^2

4.1) PRUEBAS DE HIPÓTESIS PARA MEDIAS

En vez de estimar el valor de un parámetro, a veces se debe decidir si una afirmación relativa a un parámetro es verdadera o falsa. Es decir, *probar una hipótesis* relativa a un parámetro. Se realiza una prueba de hipótesis cuando se desea probar una afirmación realizada acerca de un parámetro o parámetros de una población.

Una *hipótesis* es un enunciado acerca del valor de un parámetro (media, proporción, etc.).

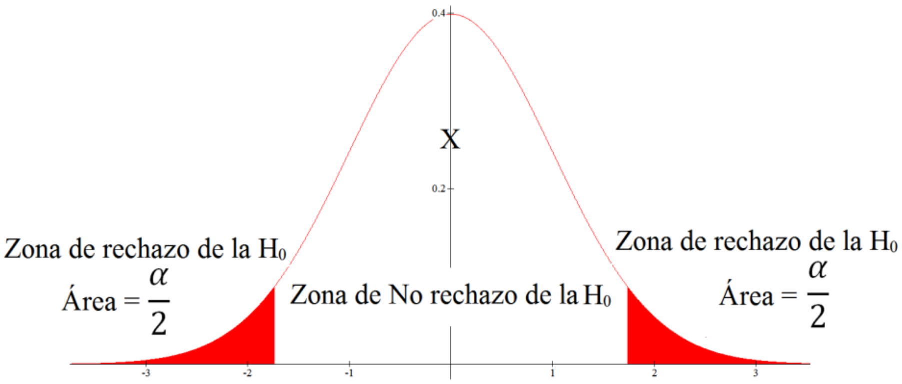
Prueba de Hipótesis es un procedimiento basado en evidencia muestral (estadístico) y en la teoría de probabilidad (distribución muestral del estadístico) para determinar si una hipótesis es razonable y no debe rechazarse, o si es irrazonable y debe ser rechazada.

La hipótesis de que el parámetro de la población es igual a un valor determinado se conoce como *hipótesis nula*. Una hipótesis nula es siempre una de status quo o de no diferencia. Se simboliza con el símbolo H_0 .Y cuando se desarrolla la prueba se asume que la hipótesis nula es verdadera y este supuesto será rechazado solo si se encuentran suficientes evidencias en base a la información muestral.

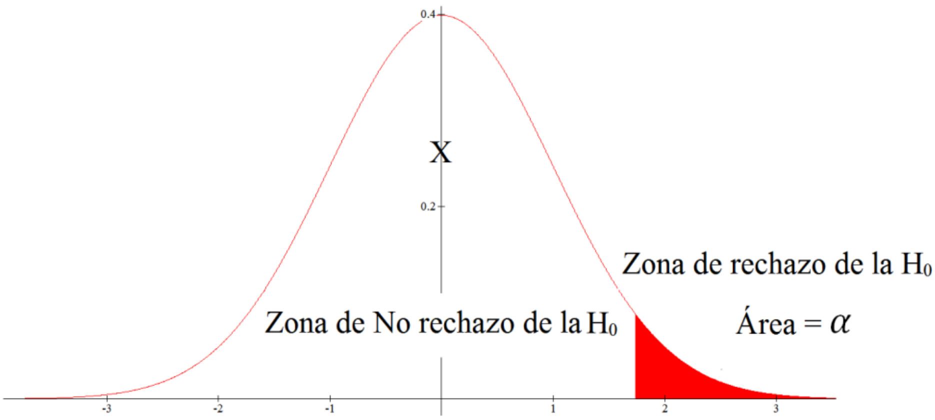
Siempre que se especifica una hipótesis nula, también se debe especificar una *hipótesis alternativa*, o una que debe ser verdadera si se encuentra que la hipótesis nula es falsa. La hipótesis alternativa se simboliza H_1. La hipótesis alternativa representa la conclusión a la que se llegaría si hubiera suficiente evidencia de la información de la muestra para decidir que es improbable que la hipótesis nula sea verdadera, y por tanto rechazarla. Es siempre opuesta a la Hipótesis Nula.

En toda prueba de hipótesis se presentan 3 casos de *zonas críticas* o llamadas también *zonas de rechazo de la hipótesis nula*, estos casos son los siguientes:

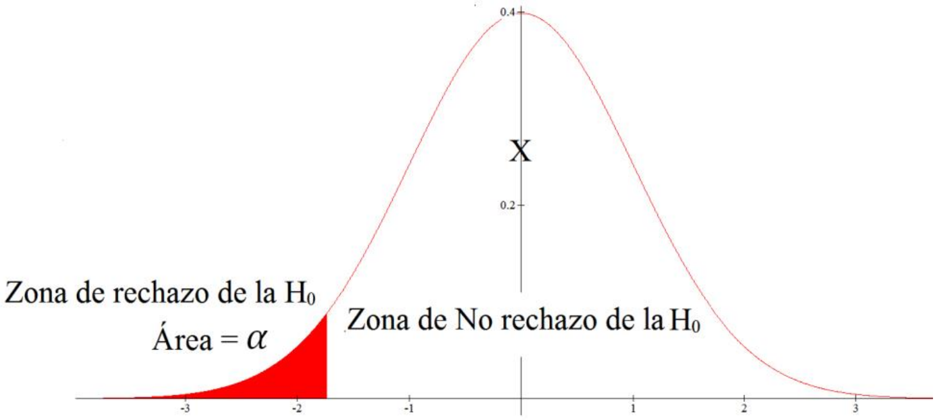
1) Prueba Bilaterial o a dos colas: $H_0: \mu = X; H_1 \neq X$



2) Prueba Unilateral con cola hacia la derecha: $H_0: \mu \leq X; H_1 > X$



3) Prueba Unilateral con cola hacia la izquierda: $H_0: \mu \geq X; H_1 < X$



En toda prueba de hipótesis se pueden cometer 2 *tipos de errores*:

1) Error tipo I: se comete error tipo I, cuando se rechaza la H_0, siendo esta realmente verdadera. A la probabilidad de cometer error tipo I, se le conoce como nivel de significación y se le denota como α

2) Error tipo II: se comete error tipo II, cuando no se rechaza la H_0, siendo esta realmente falsa. A la probabilidad de cometer error tipo II, se le denota como β

El complemento de la probabilidad de cometer error tipo II, se le llama potencia de la prueba y se denota como $1 - \beta$

Como resumen se da la siguiente tabla:

	Se Acepta H_0	Se Rechaza H_0
H_0 es Verdadera	Decisión Correcta	Error de Tipo I
H_0 es Falsa	Error de Tipo II	Decisión Correcta

TAREA DE INTERAPRENDIZAJE N° 17

Conteste

1) ¿Qué es probar una hipótesis relativa a un parámetro?

2) ¿Cuándo se realiza una prueba de hipótesis?

3) ¿Qué es una hipótesis?

4) ¿Qué es prueba de hipótesis?

5) ¿Qué es hipótesis nula?

6) ¿Qué es hipótesis alternativa o alterna?

7) ¿Cuáles son los casos de zonas de rechazo de la hipótesis nula?

8) ¿Cuáles son los tipos de errores que se pueden cometer en toda prueba de hipótesis?

9) Consulte en la biblioteca o en el internet sobre la importancia de la prueba de hipótesis. Presente la consulta mediante un organizador gráfico.

10) Consulte en la biblioteca o en el internet un ejercicio resuelto sobre prueba de hipótesis de medias de una muestra. Presente el ejercicio resuelto empleando Excel.

A) PRUEBA MEDIAS DE UNA MUESTRA

Se utiliza una prueba de una muestra para probar una afirmación con respecto a una media de una población única.

Si se conoce la desviación estándar de la población (σ), la distribución de muestreo adecuada es la distribución normal. Si la población que se muestra es normal, la distribución de muestreo será normal en el caso de todos los tamaños de la muestra, y el valor estadístico de prueba a utilizar es:

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Si la población no es normal, o si se desconoce su forma, se emplea la ecuación anterior solamente para tamaños de muestra iguales o mayores 30, es decir, para $n \geq 30$

Si no se conoce la desviación estándar de la población (σ), el valor estadístico de prueba es:

$$t_{prueba} = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

Nota: Se considera práctico utilizar la distribución t solamente cuando se requiera que el tamaño de la muestra sea menor de 30, ya que para muestras más grandes los valores t y z son aproximadamente iguales, y es posible emplear la distribución normal en lugar de la distribución t.

Las anteriores ecuaciones se aplican para poblaciones infinitas, pero cuando la población es finita y el tamaño de la muestra n constituye más del 5% del tamaño de la población N, es decir:

$$\frac{n}{N} \cdot 100\% > 5\%$$

En este caso se debe usar el factor finito de corrección para modificar las desviaciones estándar, por lo tanto se aplican las siguientes ecuaciones para (σ) conocida y desconocida, respectivamente.

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}} \qquad t_{prueba} = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$$

Ejemplos ilustrativos:

1) La duración media de una muestra de 300 focos producidos por una compañía resulta ser de 1620 horas. Se conoce que desviación típica de la población es 150 horas. Comprobar la hipótesis $\mu = 1600$ contra la hipótesis alternativa $\mu \neq 1600$ horas con un nivel de significación de 0,05 si la muestra fue tomada de 5000 focos.

Solución:

Los datos son:
$n = 300; \bar{x} = 1620; \sigma = 150; \alpha = 0{,}05; N = 5000$

Las hipótesis son:
$H_0: \mu = 1600$; $H_1: \mu \neq 1600$

Al observar $H_1: \mu \neq 1600$ se trata de una prueba a dos colas, por lo que se tiene que calcular:

$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

Como se conoce la desviación estándar de la población σ se debe utilizar la distribución normal. Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = \pm 1{,}96$. Se toma en cuenta el valor positivo y el negativo porque se trata de una prueba de hipótesis a dos colas.

Como se tiene como dato el tamaño de la población se tiene que verificar si cumple con la condición para utilizar el factor finito de corrección.

$$\frac{n}{N} \cdot 100\% > 5\% \Longrightarrow \frac{300}{5000} \cdot 100\% = 6\%$$

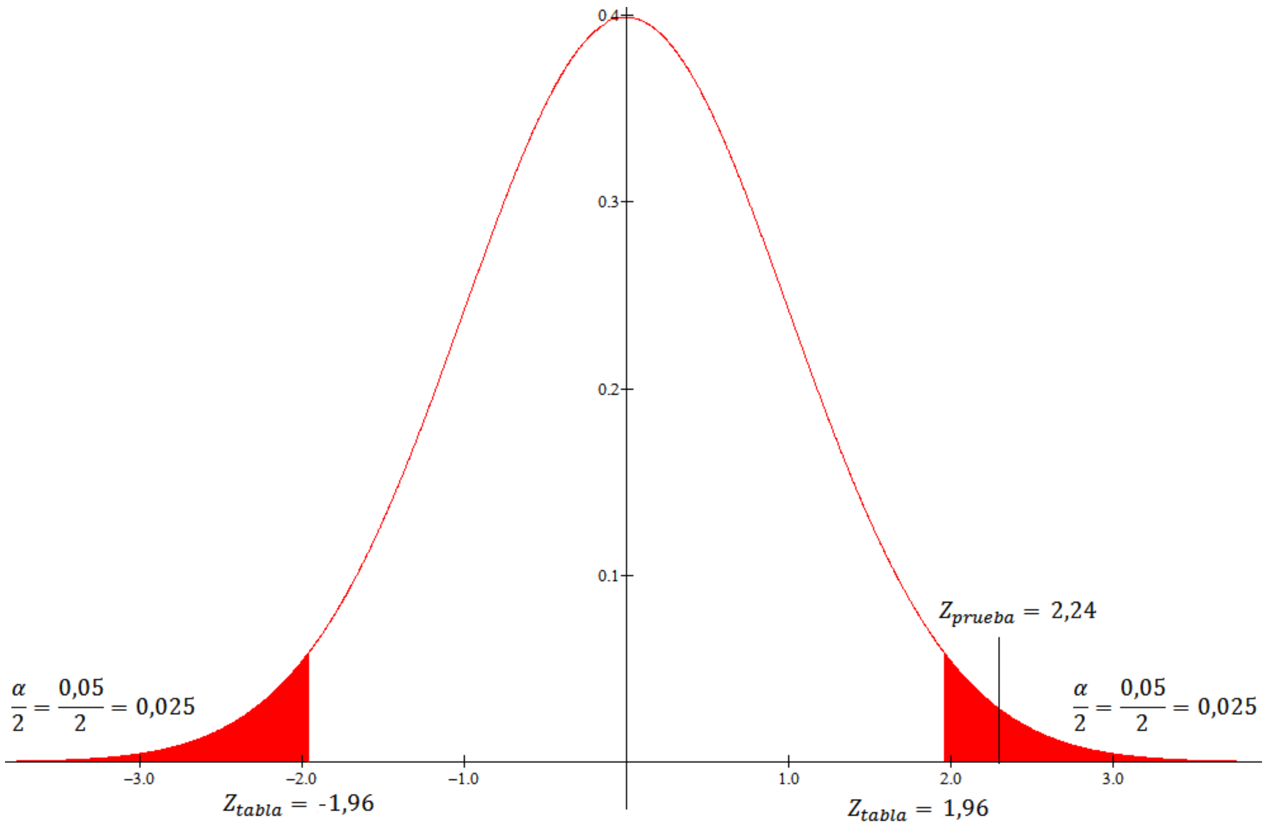
Entonces para calcular el valor de Z_{prueba} se emplea la siguiente ecuación:

$$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}} \Longrightarrow Z_{prueba} = \frac{1620 - 1600}{\frac{150}{\sqrt{300}} \cdot \sqrt{\frac{5000-300}{5000-1}}} = 2{,}24$$

Los cálculos en Excel se muestran en la siguiente imagen:

	A	B	C	D	E	F
1	N	5000				
2	n	300				
3	$\bar{x}$	1620				
4	σ	150				
5	μ	1600				
6	$H_0: \mu = 1600$					
7	$H_1: \mu \neq 1600$					
8	α	0,05				
9	$\frac{\alpha}{2}$	0,025	=B8/2			
10						
11	$\frac{n}{N} \cdot 100\% > 5\%$	6,00	=(B2/B1)*100			
12						
13	Z_{tabla}	-1,96	=INV.NORM.ESTAND(B9)			
14		1,96	=B13*-1			
15	$Z_{prueba} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$	2,24	=(B3-B5)/(B4/RAIZ(B2))*RAIZ((B1-B2)/(B1-1))			
16						
17						

El gráfico elaborado con Winstats y Paint se muestra en la siguiente imagen:



Decisión: Dado que $Z_{prueba} 2{,}24 > Z_{tabla} \pm 1{,}96$ se rechaza la H_0, y por lo tanto se acepta H_1

2) La duración media de lámparas producidas por una compañía han sido en el pasado de 1120 horas. Una muestra de 8 lámparas de la producción actual dio una duración media de 1070 horas con una desviación típica de 125 horas. Comprobar la hipótesis $\mu = 1120$ horas contra la hipótesis alternativa $\mu < 1200$ horas mediante un error tipo I de 0,05.

Solución:

Los datos son:
$\mu = 1120; n = 8; \bar{x} = 1070; S = 125; \alpha = 0{,}05$

Las hipótesis son:
$H_0: \mu = 1120; H_1: \mu < 1120$

Como se conoce la desviación estándar de la muestra S se debe utilizar la distribución t de Student. Con lectura en la tabla para un área de 0,05 y con $n - 1 = 8 - 7$ grados de libertad le corresponde un valor $t_{tabla} = -1{,}8946$. Se toma en cuenta el valor negativo porque se trata de una prueba de hipótesis a cola izquierda como se puede observar en la H_1.

Entonces para calcular el valor de t_{prueba} se emplea la siguiente ecuación:

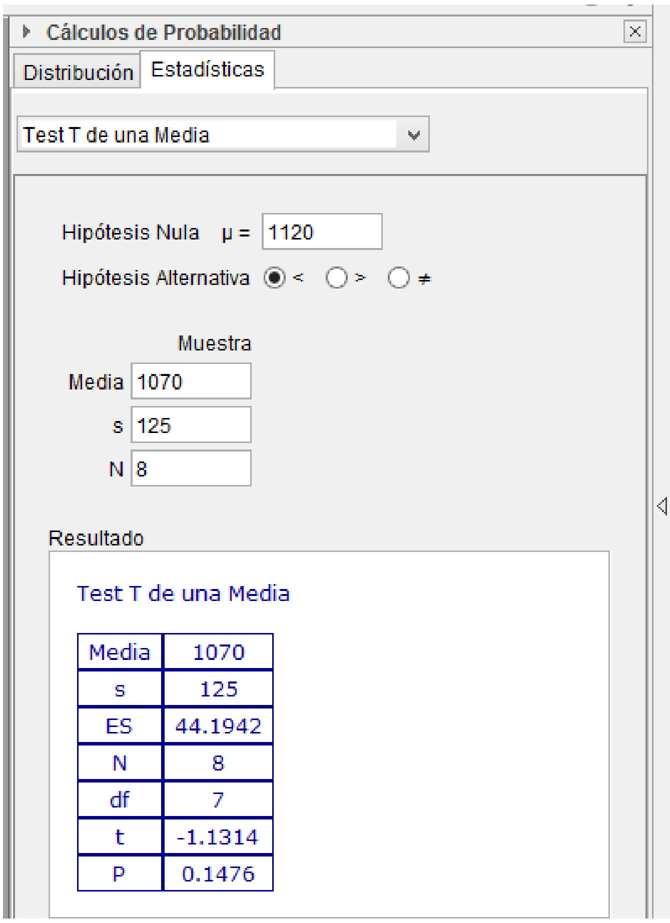
$$t_{prueba} = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$t_{prueba} = \frac{1070 - 1120}{\frac{125}{\sqrt{8}}} = -1{,}131$$

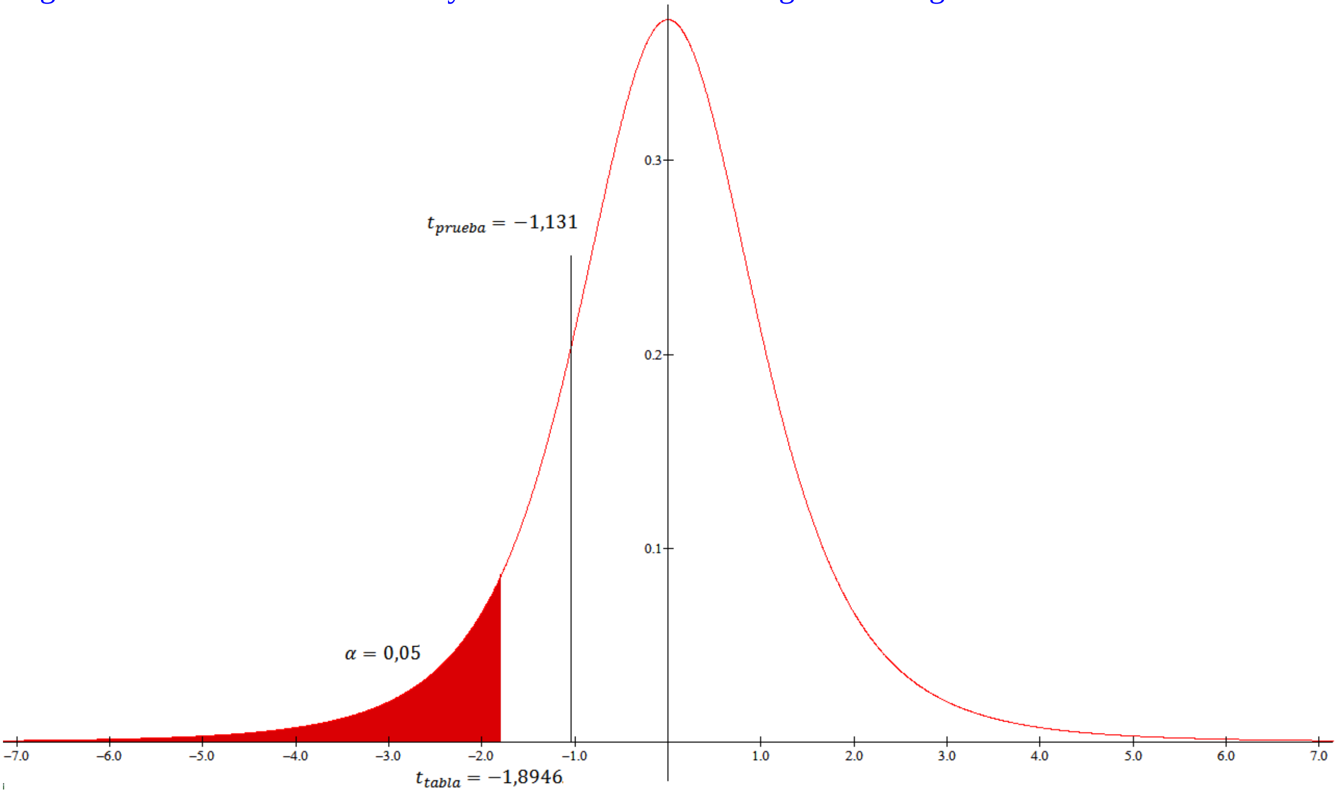
Los cálculos en Excel se muestran en la siguiente imagen:

	A	B	C	D
1	n	8		
2	$\bar{x}$	1070		
3	S	125		
4	μ	1120		
5	$H_0: \mu = 1120$			
6	$H_1: \mu < 1120$			
7	α	0,05		
8	n-1	7	=B1-1	
9	t_{tabla}	-1,895	=INV.T(B7;B8)	
10	$t_{prueba} = \dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$	-1,131	=(B2-B4)/(B3/RAIZ(B1))	
11				
12				

Los cálculos en GeoGebra se muestran en la siguiente imagen:



El gráfico elaborado con Winstats y Paint se muestra en la siguiente imagen:



Decisión: Dado que $t_{prueba} - 1{,}131 > t_{tabla} - 1{,}8946$ se Acepta la H_0

TAREA DE INTERAPRENDIZAJE N° 18

1) Contestar

1.1) ¿En qué caso se emplea la distribución normal para la pruebas de significación de medias?
1.2) ¿En qué caso se emplea la distribución t para la pruebas de significación de medias?
1.3) ¿En qué caso se emplea el factor finito de corrección para modificar las desviaciones estándar?

Los siguientes problemas resolver con lectura en la tabla, Excel y GeoGebra (para los cálculos) y el programa Winstats (para las gráficas)

2) La duración media de una muestra de 200 tubos fluorescentes producidos por una compañía resulta ser de 1620 horas. Se sabe que desviación típica de la población es de 100 horas. Comprobar la hipótesis $\mu = 1600$ contra la hipótesis alternativa $\mu \neq 1600$ horas con un nivel de significación de 0,05.

Dado que Z_{prueba} 2,83 $> Z_{tabla} \pm 1{,}96$ se rechaza la H_0

3) La duración media de las bombillas producidas por una compañía han sido en el pasado de 1000 horas. Una muestra de 28 bombillas de la producción actual dio una duración media de 1050 horas con una desviación típica de 120 horas. Comprobar la hipótesis $\mu = 1000$ horas contra la hipótesis alternativa $\mu < 1000$ horas mediante un nivel de significancia de 0,05.

Dado que t_{prueba} 2,205 $> t_{tabla} - 1{,}703$ se Acepta la H_0

4) Plantear y resolver un problema de prueba de hipótesis conociendo la desviación estándar de la población.

5) Plantear y resolver un problema de prueba de hipótesis conociendo la desviación estándar de la muestra y el tamaño de la población.

B) PRUEBA MEDIAS DE DOS MUESTRAS

Las pruebas de dos muestras se utilizan para decidir si las medias de dos poblaciones son iguales. Se requieren dos muestras independientes, una de cada una de las dos poblaciones. Considérese, por ejemplo, una compañía investigadora que experimentan con dos diferentes mezclas de pintura, para ver si se puede modificar el tiempo de secado de una pintura para uso doméstico. Cada mezcla es probada un determinado número de veces, y comparados posteriormente los tiempos medios de secado de las dos muestras. Una parece ser superior, ya que su tiempo medio de secado (muestra) es 30 minutos menor que el de la otra muestra.

Pero, ¿son realmente diferentes los tiempos medios de secado de las dos pinturas, o esta diferencia muestral es nada más la variación aleatoria que se espera, aun cuando las dos fórmulas presentan idénticos tiempos medios de secado? Una vez más, las diferencias casuales se deben distinguir de las diferencias reales.

Con frecuencia se utilizan pruebas de dos muestras para comparar dos métodos de enseñanza, dos marcas, dos ciudades, dos distritos escolares y otras cosas semejantes.

La hipótesis nula puede establecer que las dos poblaciones tienen medias iguales:

$$H_0: \mu_1 = \mu_2$$

Las alternativas pueden ser alguna de las siguientes:

$$H_1: \mu_1 \neq \mu_2 \qquad H_1: \mu_1 > \mu_2 \qquad H_1: \mu_1 < \mu_2$$

Cuando se conocen las desviaciones estándar de la población σ_1 y σ_1, el valor estadístico de prueba es el siguiente:

$$Z_{prueba} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Cabe suponer que el valor real de Z, cuando H_0 es verdadera, está distribuido normalmente con una media de 0 y una desviación estándar de 1 (es decir, la distribución normal estandarizada) para casos en los que la suma $n_1 + n_2$ es igual o mayor de 30. Para tamaños más pequeños de muestra, Z estará distribuida normalmente sólo si las dos poblaciones que se muestrean también lo están.

Cuando no se conocen las desviaciones estándar de la población, y $n_1 + n_2$ es menor a 30, el valor estadístico de prueba es como el que se presenta a continuación.

$$t_{prueba} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Cuando los tamaños de las dos muestras no son iguales, y su suma es menor de 30, la fórmula para el valor estadístico de prueba se convierte en:

$$t_{prueba} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}\right]\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

El valor de t cuando H_0 es verdadera, tiene una distribución t con $n_1 + n_2 - 2$ grados de libertad, sí se puede suponer que ambas poblaciones son aproximadamente normales.

Ejemplo ilustrativo

La media de las calificaciones de dos muestras de 15 estudiantes de primer semestre en la asignatura de Estadística de la universidad UTN resulta ser de 7 y 8,5. Se sabe que la desviación típica de las calificaciones en esta asignatura fue en el pasado de 1,5. Comprobar la hipótesis $\mu_1 = \mu_2$ contra la hipótesis alternativa $\mu_1 < \mu_2$ con un nivel de significación de 0,025.

Solución:

Los datos son:
$n_1 = n_2 = 15$
$\bar{x}_1 = 7$
$\bar{x}_2 = 8{,}5$
$\sigma_1 = \sigma_2 = 1{,}5$
$\alpha = 0{,}025$

Las hipótesis son:
$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 < \mu_2$

Como se conoce la desviación estándar de la población σ se debe utilizar la distribución normal. Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = -1{,}96$. Se toma en cuenta el valor negativo porque se trata de una prueba de hipótesis a cola izquierda.

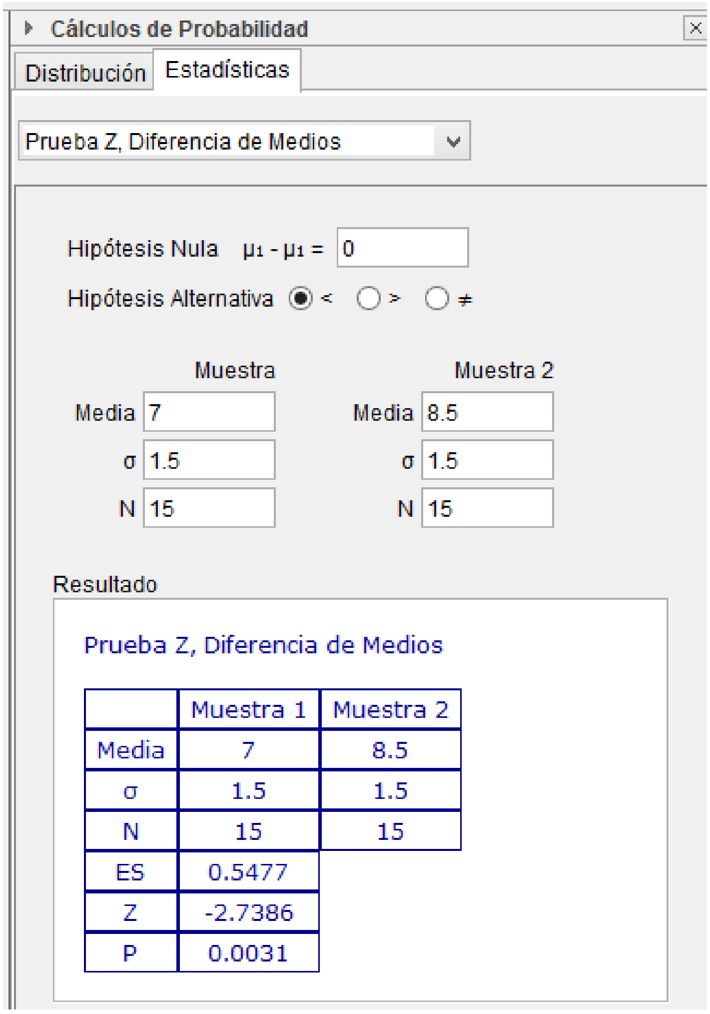
Se remplaza valores en la siguiente fórmula porque se conoce la desviación estándar de la población σ y el tamaño de las muestras son iguales y $n_1 + n_2$ es igual o mayor de 30:

$$Z_{prueba} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \Rightarrow Z_{prueba} = \frac{7 - 8{,}5}{\sqrt{\frac{1{,}5^2}{15} + \frac{1{,}5^2}{15}}} = -2{,}74$$

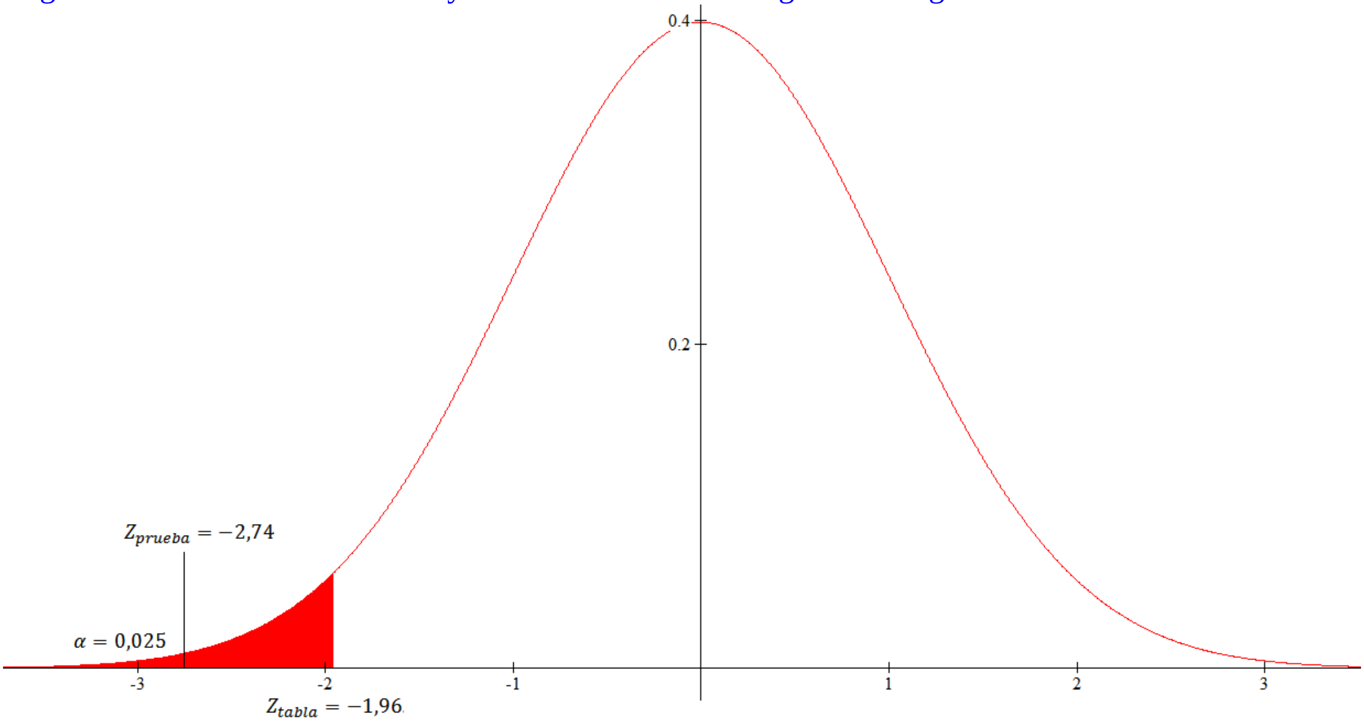
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	n_1	15			
2	n_2	15			
3	$\bar{x}_1$	7			
4	$\bar{x}_2$	8,5			
5	σ_1	1,5			
6	σ_2	1,5			
7	α	0,025			
8	$H_0: \mu_1 = \mu_2$				
9	$H_1: \mu_1 < \mu_2$				
10	Z_{tabla}	-1,96	=INV.NORM.ESTAND(B7)		
11	$Z_{prueba} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	-2,74	=(B3-B4)/RAIZ(B5^2/B1+B6^2/B2)		
12					
13					
14					

Los cálculos en GeoGebra se muestran en la siguiente figura:



El gráfico elaborado con Winstats y Paint se muestra en la siguiente imagen:



Decisión: Se rechaza H_0, ya que $Z_{prueba} < Z_{tabla}$

TAREA DE INTERAPRENDIZAJE N° 19

1) Responder

1.1) ¿Para qué se utilizan las pruebas de dos muestras?

1.2) ¿Qué son muestras independientes?. Explique con un ejemplo

1.3) ¿En qué se concentra la prueba de hipótesis de dos muestras?

1.4) ¿En qué caso se emplea la prueba z?, y ¿Qué elementos están presentes en la fórmula de la prueba z?.

1.5) ¿Qué cabe suponer para el valor real de z cuando H_0 es verdadera?

1.6) ¿En qué caso se emplea la prueba t?, y ¿Qué elementos están presentes en la fórmula de la prueba t?.

1.7) ¿En qué caso el valor de t, suponiendo que H_0 es verdadera se puede aproximar mediante z?

1.8) ¿Qué fórmula para el valor estadístico de prueba de hipótesis se emplearía cuando los tamaños de las dos muestras no son iguales, y su suma es mayor de 30?.

Realizar la prueba de hipótesis con lectura en la tabla, Excel y GeoGebra (para los cálculos) y el programa Winstats (para las gráficas)

2) H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 > \mu_2$, $\alpha = 0{,}05$, $\bar{x}_1 = 20$, $\bar{x}_2 = 18$, $\sigma_1=\sigma_2=3$, $n_1= n_2= 36$

Se rechaza H_0, ya que $z_{prueba} > z_{tabla}$

2) H_0: $\mu_1 = \mu_2$, H_1: $\mu_1 \neq \mu_2$, $\alpha = 0{,}05$, $\bar{x}_1 = 5{,}4$, $\bar{x}_2 = 5$, $S_1=1{,}1$, $S_2=1{,}2$, $n_1= n_2= 14$

Se aprueba H_0, ya que t_{prueba} se encuentra dentro del intervalo de aceptación

3) La media de las edades de dos muestras de 36 estudiantes de primer semestre de una universidad resulta ser de 20 años y 18 años. Se sabe que la desviación típica de todos los estudiantes fue en el pasado de 3 años. Comprobar la hipótesis $\mu_1 = \mu_2$ contra la hipótesis alternativa $\mu_1 > \mu_2$ con un nivel de significación de 0,05

Se rechaza H_0, ya que $z_{prueba} > z_{tabla}$

4) La duración media de dos muestras de 10 y 14 pantalones producidos por dos empresas resulta ser de 5,4 años y 5 años, con una desviación típica de 1,1 años y 1,2 años, respectivamente. Comprobar la hipótesis $\mu_1 = \mu_2$ contra la hipótesis alternativa $\mu_1 \neq \mu_2$ con un nivel de significación de 0,05

Se aprueba H_0, ya que t_{prueba} se encuentra dentro del intervalo de aceptación

5) Seleccione dos cursos de diferente paralelo de su institución educativa. Seleccione dos muestras iguales y menores a 30 de cada curso. Investigue las calificaciones de cada curso en la asignatura que usted crea conveniente y que imparta clases el mismo maestro. Con esas calificaciones calcule la media y la desviación típica de cada muestra. Comprobar la hipótesis $\mu_1 = \mu_2$ contra la hipótesis alternativa $\mu_1 \neq \mu_2$ con un nivel de significación de 0,05.

6) Plantee y resuelva un problema de prueba de hipótesis similar al anterior de cualquier tema de su preferencia.

4.2) ANÁLISIS DE VARIANZA

El análisis de varianza es una técnica que se puede utilizar para decidir si las medias de dos o más poblaciones son iguales. La prueba se basa en una muestra única, obtenida a partir de cada población. El análisis de varianza puede servir para determinar si las diferencias entre las medias muestrales revelan las verdaderas diferencias entre los valores medios de cada una de las poblaciones, o si las diferencias entre los valores medios de la muestra son más indicativas de una variabilidad de muestreo.

Si el valor estadístico de prueba (análisis de varianza) nos impulsa a aceptar la hipótesis nula, se concluiría que las diferencias observadas entre las medias de las muestras se deben a la variación casual en el muestreo (y por tanto, que los valores medios de población son iguales). Si se rechaza la hipótesis nula, se concluiría que las diferencias entre los valores medios de la muestra son demasiado grandes como para deberse únicamente a la casualidad (y por ello, no todas las medias de población son iguales).

Los datos para el análisis de varianza se obtienen tomando una muestra de cada población y calculando la media muestral y la variancia en el caso de cada muestra.

Existen tres supuestos básicos que se deben satisfacer antes de que se pueda utilizar el análisis de variancia.

1) Las muestras deben ser de tipo aleatorio independiente.
2) Las muestras deben ser obtenidas a partir de poblaciones normales.
3) Las poblaciones deben tener variancias iguales (es decir, $\sigma_1^2 = \sigma_2^2 = \ldots\ldots \sigma_k^2$)

El análisis de varianza, como su nombre lo indica, comprende el cálculo de varianzas. La varianza de una muestra es el promedio de las desviaciones elevadas al cuadrado de la media del grupo. Simbólicamente, esto se representa de la siguiente manera:

$$varianza\ de\ la\ muestra = s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

Cabe observar que se debe utilizar n - 1, ya que se está trabajando con datos muestrales. De ahí que, para obtener la varianza muestral, el procedimiento sea el siguiente:

1) Calcular la media muestral
2) Restar la media de cada valor de la muestra.
3) Elevar al cuadrado cada una de las diferencias.
4) Sumar las diferencias elevadas al cuadrado.
5) Dividir entre n - 1

A) ESTIMACIÓN INTERNA DE VARIANZA (WITHIN ESTÍMATE) s_w^2

Aunque parezca extraño un examen de las varianzas puede revelar si todas las medias de la población son iguales o no. El análisis de varianza utiliza dos métodos un poco diferentes para estimar las varianzas de la población (iguales). Si las dos estimaciones son aproximadamente iguales, esto tiende a confirmar H_0; si una de las dos estimaciones es mucho mayor que la otra, esto tiende a confirmar H_1. Si la hipótesis nula es verdadera, entonces las muestras se habrán obtenido de poblaciones con medias iguales. Y como se supone que todas las poblaciones son normales y poseen variancias iguales, cuando H_0 es verdadera se presenta una situación conceptualmente idéntica a otra en la que todas las muestras hayan sido tomadas realmente a partir de una población única. Si H_0 es falsa, entonces las muestras provendrán de poblaciones que no presentan todas la misma media, sin embargo, cabe observar que, aún en ese caso, se debe suponer que las poblaciones son normales y tienen variancias iguales.

Una forma de calcular la varianza poblacional es sacar el promedio de las varianzas de las muestras. Es evidente que se podrá utilizar cualquiera de las varianzas muestrales, pero el promedio de todas ellas por lo general proporcionará la mejor estimación debido al mayor número de observaciones que representa. Como cada varianza muestral sólo refleja la variación dentro de una muestra en particular, la estimación de la varianza basada en el promedio de las varianzas muestrales se llama *estimación interna de variancia*. La estimación interna de variancia se calcula de la siguiente manera:

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \ldots \ldots \ldots \ldots . s_k^2}{k}$$

Donde:
s_1^2 = variancia de la muestra 1 ; s_2^2 = variancia de la muestra 2; s_3^2 = variancia de la muestra 3
s_k^2 = variancia de la muestra k ; k = número de muestras

B) ESTIMACIÓN INTERMEDIANTE DE VARIANZA (BETWEEN ESTÍMATE) s_x^2

Como se supone que las varianzas de la población son iguales, independientemente de si las medias lo son o no, la estimación interna de varianza no se altera por la verdad o falsedad de H_0. Por tanto, *no se puede utilizar por sí misma para determinar si las medias de la población podrían ser iguales*. No obstante, sirve como una norma de comparación respecto a la cual puede evaluarse una segunda estimación llamada estimación intermediante de varianza. *Esta segunda estimación es sensible a diferencias entre las medias de población.*

La estimación interna de varianza sirve como una norma respecto a la cual se puede comparar la estimación intermediante de varianza.

La estimación de varianza entre muestras determina una estimación de las varianzas iguales de la población de una forma indirecta a través de una distribución de muestreo de medias. Recuérdese que si H_0, es verdadera, esto equivale a tomar todas las muestras de la misma población normal. Además, por el Teorema del Límite Central, se sabe que la distribución de muestreo de medias, obtenida de una población normal, estará distribuida normalmente, y que la desviación estándar de la distribución de muestreo (raíz cuadrada de su varianza) está directamente relacionada con el tamaño de la desviación estándar de la población (raíz cuadrada de la varianza de la población). Es decir,

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

Donde:
$s_{\bar{x}}$ = desviación estándar del muestreo de distribución de medias
s_x = desviación estándar de la población
n = tamaño de la muestra
Elevando al cuadrado ambos miembros de la ecuación, se obtiene la relación en términos de variancias:

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$$

Ahora, si se conoce la variancia de la distribución de muestreo, únicamente multiplicándola por el tamaño de la muestra, se obtendrá exactamente el valor de σ_x^2 decir,

$$\sigma_x^2 = n\sigma_{\bar{x}}^2$$

Pero como por lo general no se conoce σ_x^2 se emplea

$$s_x^2 = ns_{\bar{x}}^2$$

Donde

$$s_{\bar{x}}^2 = varianza\ de\ las\ medidas\ aritméticas = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1}$$

Ejemplos ilustrativos

1) Calcular la varianza muestral de 16, 19, 17, 16, 20, 19, 20

Solución:
Calculando la media aritmética se obtiene:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{16+19+17+16+20+19+20}{7} = \frac{127}{7} = 18{,}143$$

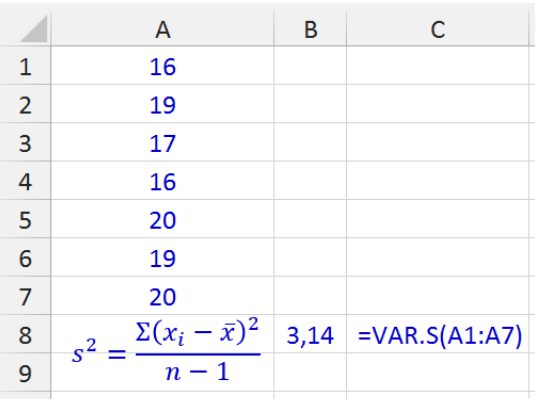
Llenando la tabla para obtener datos para reemplazar valores de la fórmula de la varianza se obtiene:

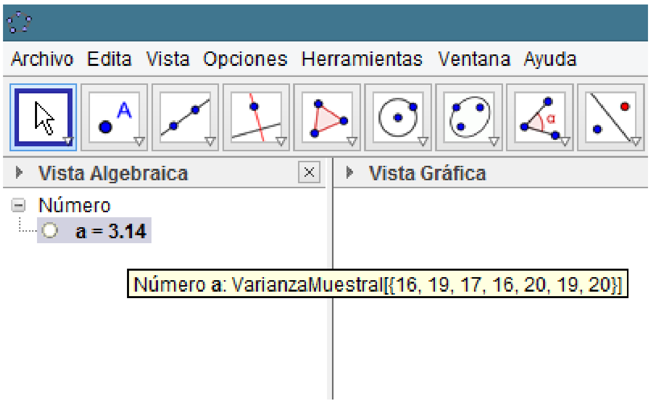
x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
16	-2,143	4,592
19	0,857	0,734
17	-1,143	1,306
16	-2,143	4,592
20	1,857	3,448
19	0,857	0,734
20	1,857	3,448
Total		18,854

Remplazando valores en la fórmula y realizando las operaciones respectivas se tiene:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{18{,}854}{7-1} = \frac{18{,}854}{6} = 3{,}14$$

Los cálculos en Excel y en GeoGebra se muestran en la siguientes figuras:





2) Dado la siguiente tabla con datos acerca del peso en kg por 1,7 m de estatura

Observación	Muestra			
	1	2	3	4
1	70	74	68	75
2	75	77	70	70
3	74	70	65	73
4	72	80	60	72
5	68	72	72	71
6	59	76	73	72

a) Calcular la estimación interna de variancia
b) Calcular la estimación intermediante de variancia

Solución:
Calculando las medias aritméticas se obtiene:

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{x}_1 = \frac{70 + 75 + 74 + 72 + 68 + 59}{6} = \frac{418}{6} = 69{,}667$$

$$\bar{x}_2 = \frac{74 + 77 + 70 + 80 + 72 + 76}{6} = \frac{449}{6} = 74{,}833$$

$$\bar{x}_3 = \frac{68 + 70 + 65 + 60 + 72 + 73}{6} = \frac{408}{6} = 68$$

$$\bar{x}_4 = \frac{75 + 70 + 73 + 72 + 71 + 72}{6} = \frac{433}{6} = 72{,}167$$

Se llena la siguiente tabla:

Observación	Muestra				$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$	$(x_3 - \bar{x}_3)^2$	$(x_4 - \bar{x}_4)^2$
	1	2	3	4				
1	70	74	68	75	0,111	0,694	0	8,026
2	75	77	70	70	28,441	4,696	4	4,696
3	74	70	65	73	18,775	23,358	9	0,694
4	72	80	60	72	5,443	26,698	64	0,028
5	68	72	72	71	2,779	8,026	16	1,361
6	59	76	73	72	113,785	1,362	25	0,028
Total	418	449	408	433	169,334	64,834	118	14,833

Remplazando los datos en la fórmula de la varianza se obtienen las varianzas de las 4 muestras.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Varianza de la muestra 1

$$s_1^2 = \frac{169{,}334}{5} = 33{,}867$$

Varianza de la muestra 2

$$s_2^2 = \frac{64{,}834}{5} = 12{,}967$$

Varianza de la muestra 3

$$s_3^2 = \frac{118}{5} = 23{,}6$$

Varianza de la muestra 4

$$s_4^2 = \frac{14{,}833}{5} = 2{,}967$$

a) Calculando la estimación interna de varianza se obtiene:

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k} \Longrightarrow s_w^2 = \frac{33{,}867 + 12{,}967 + 23{,}6 + 2{,}967}{4} = \frac{73{,}401}{4} = 18{,}35$$

Nota: La estimación interna de varianza es la media aritmética de la varianzas.

b) Para calcular la estimación intermediante de varianza primero se calcular la varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k - 1}$$

Para calcular la varianza de las medias aritméticas se calcula la media aritmética de las medias aritméticas, la cual es:

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{69{,}667 + 74{,}833 + 68 + 72{,}167}{4} = \frac{284{,}667}{4} = 71{,}167$$

Se llena la siguiente tabla:

$\bar{x}$	$(\bar{x} - \bar{\bar{x}})^2$
69,667	2,25
74,833	13,44
68	10,03
72,167	1
Total	26,72

Se remplaza los datos de la tabla para calcular varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k - 1} = \frac{26{,}72}{3} = 8{,}907$$

Finalmente se calcula la estimación intermediante de varianza, la cual queda:

$$s_{\bar{x}}^2 = n \cdot s_{\bar{x}}^2 = 6 \cdot 8{,}907 = 53{,}44$$

Los cálculos en Excel se muestra la siguiente figura:

	A	B	C	D	E	F	G
1		Muestra					
2	Observación	1	2	3	4	$\bar{x}$	
3	1	70	74	68	75	69,667	=PROMEDIO(B3:B8)
4	2	75	77	70	70	74,833	=PROMEDIO(C3:C8)
5	3	74	70	65	73	68,000	=PROMEDIO(D3:D8)
6	4	72	80	60	72	72,167	=PROMEDIO(E3:E8)
7	5	68	72	72	71		
8	6	59	76	73	72		
9		33,87	12,97	23,60	2,97		
10	s^2	=VAR.S(B3:B8)	=VAR.S(C3:C8)	=VAR.S(D3:D8)	=VAR.S(E3:E8)		
11							
12	Estimación interna		18,350	=PROMEDIO(B9:E9)		$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$	
13	Varianza de la medias aritméticas		8,907	=VAR.S(F3:F6)		$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k - 1}$	
14	n		6	=CONTAR(A3:A8)			
15	Estimación intermediante		53,44	=C14*C13		$s_{\bar{x}}^2 = n \cdot s_{\bar{x}}^2$	

TAREA DE INTERAPRENDIZAJE N° 20

1) Realizar un organizador gráfico sobre el análisis de varianza

2) Obtenga la varianza de los siguientes datos muestrales empleando la fórmula, con Excel y GeoGebra:
2.1) 5, 10, 10, 15, 20, 20 y 25

$s^2 = 50$

2.2) 16, 19, 17,16, 20, 19, 20

$s^2 = 3{,}14$

3) Dado la siguiente tabla con datos acerca del peso en kg por 1,7 m de estatura

Observación	Muestra			
	1	2	3	4
1	70	74	68	75
2	75	77	70	70
3	74	70	65	73
4	72	80	60	72
5	68	72	72	71
6	59	76	73	72

3.1) Realizar una estimación interna de varianza de manera manual y empleando Excel

$s_w^2 = 18{,}35$

3.2) Calcular la estimación intermediante de varianza de manera manual y empleando Excel

$s_{\bar{x}}^2 = 53{,}44$

4) Crear y resolver un ejercicio similar al anterior

C) LA RAZÓN F

A diferencia de otras pruebas de medias que se basan en la diferencia existente entre dos valores, el análisis de varianza emplea la razón de las estimaciones, dividiendo la estimación intermediante entre la estimación interna

$$Razón\ F = \frac{s_{\bar{x}}^2}{s_w^2} = \frac{ns_{\bar{x}}^2}{(s_1^2 + s_2^2 + s_3^2 + \cdots s_k^2)/k}$$

Esta razón F fue creada por el matemático británico Ronald Fisher (1890-1962). El valor estadístico de prueba resultante se debe comparar con un valor tabular de F, que indicará el valor máximo del valor estadístico de prueba que ocurría si H_0 fuera verdadera, a un nivel de significación seleccionado. Antes de proceder a efectuar este cálculo, se debe considerar las características de la distribución F

i) Características de la distribución F

- Existe una distribución F diferente para cada combinación de tamaño de muestra y número de muestras. En el caso de la distribución F, los valores críticos para los niveles 0,05 y 0,01 generalmente se proporcionan para determinadas combinaciones de tamaños de muestra y número de muestras.

- La distribución es continua respecto al intervalo de 0 a + ∞. La razón más pequeña es 0. La razón no puede ser negativa, ya que ambos términos de la razón F están elevados al cuadrado.

- La forma de cada distribución de muestreo teórico F depende del número de grados de libertad que estén asociados a ella. Tanto el numerador como el denominador tienen grados de libertad relacionados.

ii) Determinación de los grados de libertad

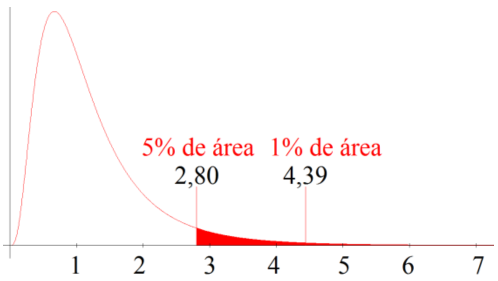
Los grados de libertad para el numerador y el denominador de la razón F se basan en los cálculos necesarios para derivar cada estimación de la variancia de la población. La *estimación intermediante* de variancia (numerador) comprende la división de la suma de las diferencias elevadas al cuadrado entre el número de medias (muestras) menos uno, o bien, k - 1. Así, **k - 1** *es el número de grados de libertad para el numerador.*

En forma semejante, el calcular cada variancia muestral, la suma de las diferencias elevadas al cuadrado entre el valor medio de la muestra y cada valor de la misma se divide entre el número de observaciones de la muestra menos uno, o bien, n - 1. Por tanto, el promedio de las variancias muestrales se determina dividiendo la suma de las variancias de la muestra entre el número de muestras, o k. *Los grados de libertad para el denominador son entonce*s, **k(n -l).**

iii) Uso de la tabla de F del análisis de variancia (ANOVA)

En la tabla 5 se ilustra la estructura de una tabla de F para un nivel de significación de 0,01 o 1% y 0,05 o 5%. Se obtiene el valor tabular, localizando los grados de libertad del numerador n_1 (que se listan en la parte superior de la tabla), así como los del denominador n_2 (que se listan en una de las columnas laterales de la tabla) que corresponden a una situación dada. Utilizando el nivel de significación de 0,05 para $n_1 = 7\ y\ n_2 = 3$ grados de libertad, el valor de F es 8,89

TABLA Nº 5
DISTRIBUCIÓN F DE FISHER



Ejemplos:

$Para\ n_1 = 9\ ;\ n_2 = 12\ grados\ de\ libertad$

$P(F > 2{,}80) = 0{,}05 = 5\%$

$\boldsymbol{P(F > 4{,}39) = 0{,}01 = 1\%}$

5% (normal) **y 1% (negritas)**

$n_1 =$ grados de libertad del numerador

$n_2 =$ grados de libertad del denominador

n_1 / n_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**15**	**20**	**25**	**50**	**100**
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	245,95	248,01	249,26	251,77	253,04
	4052,2	**4999,5**	**5403,4**	**5624,6**	**5763,6**	**5859,0**	**5928,4**	**5981,1**	**6022,5**	**6055,8**	**6157,3**	**6208,7**	**6239,8**	**6302,5**	**6334,1**
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43	19,45	19,46	19,48	19,49
	98,50	**99,00**	**99,17**	**99,25**	**99,30**	**99,33**	**99,36**	**99,37**	**99,39**	**99,40**	**99,43**	**99,45**	**99,46**	**99,48**	**99,49**
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66	8,63	8,58	8,55
	34,12	**30,82**	**29,46**	**28,71**	**28,24**	**27,91**	**27,67**	**27,49**	**27,35**	**27,23**	**26,87**	**26,69**	**26,58**	**26,35**	**26,24**

iv) Cálculo de la razón F a partir de datos muestrales

$$F_{prueba} = \frac{estimación\ intermediante\ de\ variancia}{estimación\ interna\ de\ variancia} \Longrightarrow F_{prueba} = \frac{s_{\bar{x}}^2}{s_w^2} = \frac{ns_{\bar{x}}^2}{(s_1^2 + s_2^2 + s_3^2 + \cdots s_k^2)/k}$$

v) Hipótesis Nula y Alternativa

H_0: Todas las proporciones de la población son iguales.

H_1: No todas las proporciones de la población son iguales.

Ejemplo ilustrativo

Los pesos en kg por 1,7 m de estatura se ilustran en la siguiente tabla. La finalidad es determinar si existen diferencias reales entre las cuatro muestras. Emplear un nivel de significación de 0,05

Observación	Muestra			
	1	2	3	4
1	70	74	68	75
2	75	77	70	70
3	74	70	65	73
4	72	80	60	72
5	68	72	72	71
6	59	76	73	72

Solución:

Las hipótesis Nula y Alternativa son:

H_0: Todas las proporciones de la población son iguales.

H_1: No todas las proporciones de la población son iguales.

Calculando los grados de libertad de numerador se tiene:

$k-1=4-1=3$

Calculando los grados de libertad del denominador se tiene:

$k(n-1)=4(6-1)=20$

Con 3 grados de libertad en el numerador, 20 grados de libertad en el denominador y con un nivel de significación $\alpha=0,05$ con lectura la tabla se obtiene $F_{tabla}=3,10$

Para calcular F_{prueba} se procede de la siguiente manera:

Calculando las medias aritméticas se obtiene:

$$\bar{x}=\frac{\sum x_i}{n}$$

$$\bar{x}_1=\frac{70+75+74+72+68+59}{6}=\frac{418}{6}=69,667$$

$$\bar{x}_2=\frac{74+77+70+80+72+76}{6}=\frac{449}{6}=74,833$$

$$\bar{x}_3=\frac{68+70+65+60+72+73}{6}=\frac{408}{6}=68$$

$$\bar{x}_4=\frac{75+70+73+72+71+72}{6}=\frac{433}{6}=72,167$$

Se llena la siguiente tabla para calcular las varianzas muestrales:

Observación	Muestra				$(x_1-\bar{x}_1)^2$	$(x_2-\bar{x}_2)^2$	$(x_3-\bar{x}_3)^2$	$(x_4-\bar{x}_4)^2$
	1	2	3	4				
1	70	74	68	75	0,111	0,694	0	8,026
2	75	77	70	70	28,441	4,696	4	4,696
3	74	70	65	73	18,775	23,358	9	0,694
4	72	80	60	72	5,443	26,698	64	0,028
5	68	72	72	71	2,779	8,026	16	1,361
6	59	76	73	72	113,785	1,362	25	0,028
Total	418	449	408	433	169,334	64,834	118	14,833

Remplazando los datos en la fórmula de la varianza se obtienen las varianzas de las 4 muestras.

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$
$$s_1^2 = \frac{169{,}334}{5} = 33{,}867$$
$$s_2^2 = \frac{64{,}834}{5} = 12{,}967$$
$$s_3^2 = \frac{118}{5} = 23{,}6$$
$$s_4^2 = \frac{14{,}833}{5} = 2{,}967$$

Calculando la estimación interna de varianza se obtiene:

$$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$$

$$s_w^2 = \frac{33{,}867 + 12{,}967 + 23{,}6 + 2{,}967}{4} = \frac{73{,}401}{4} = 18{,}35$$

Para calcular la estimación intermediante de varianza primero se calcular la varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1}$$

Para calcular la varianza de las medias aritméticas se calcula la media aritmética de las medias aritméticas, la cual es:

$$\bar{\bar{x}} = \frac{\sum \bar{x}_i}{k} = \frac{69{,}667 + 74{,}833 + 68 + 72{,}167}{4} = \frac{284{,}667}{4} = 71{,}167$$

Se llena la siguiente tabla:

$\bar{x}$	$(\bar{x} - \bar{\bar{x}})^2$
69,667	2,25
74,833	13,44
68	10,03
72,167	1
Total	26,72

Se remplaza los datos de la tabla para calcular varianza de las medias aritméticas

$$s_{\bar{x}}^2 = \frac{\sum(\bar{x} - \bar{\bar{x}})^2}{k-1} = \frac{26{,}72}{3} = 8{,}907$$

Calculando la estimación intermediante de varianza se obtiene:

$$s_x^2 = n \cdot s_{\bar{x}}^2 = 6 \cdot 8{,}907 = 53{,}44$$

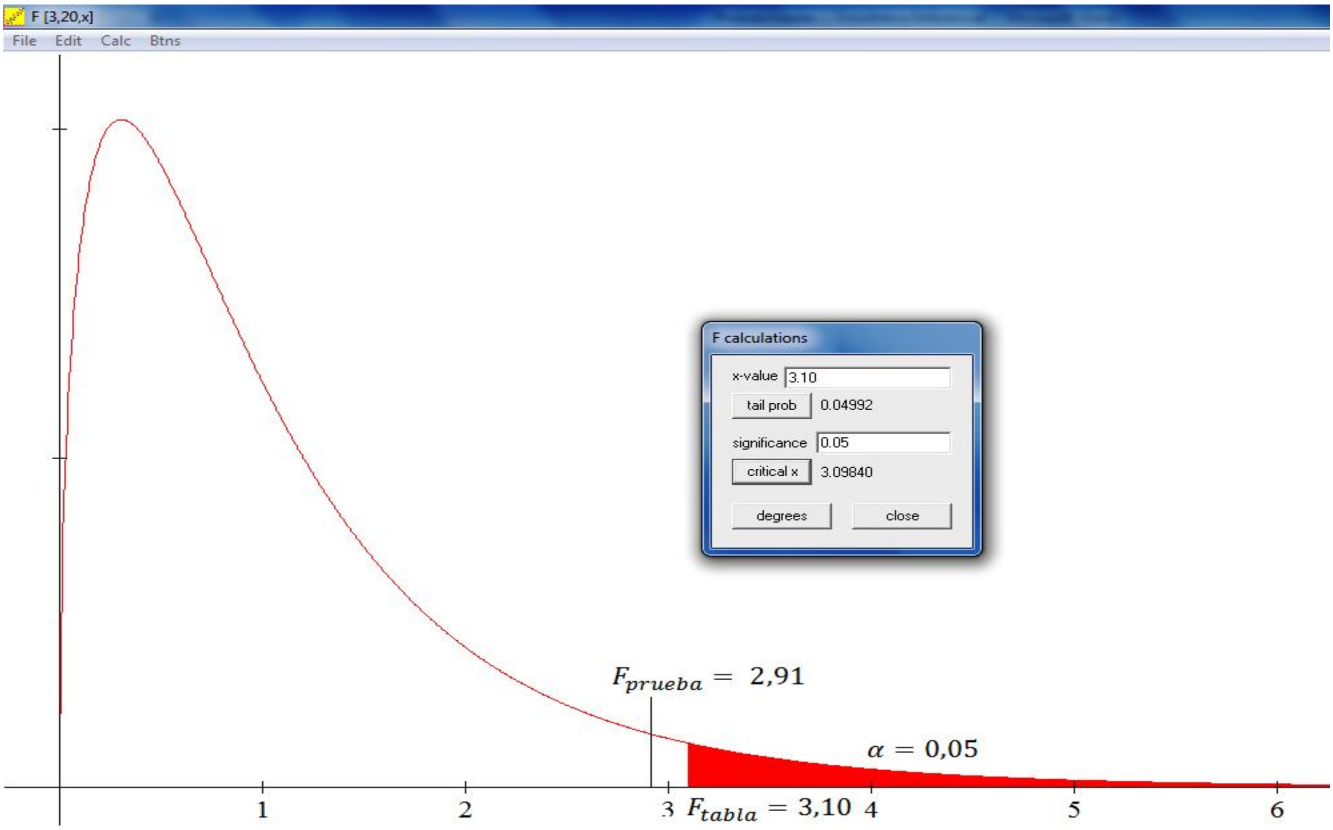
Finalmente calculando F_{prueba} se tiene:

$$F_{prueba} = \frac{s_x^2}{s_w^2} = \frac{53{,}44}{18{,}35} = 2{,}91$$

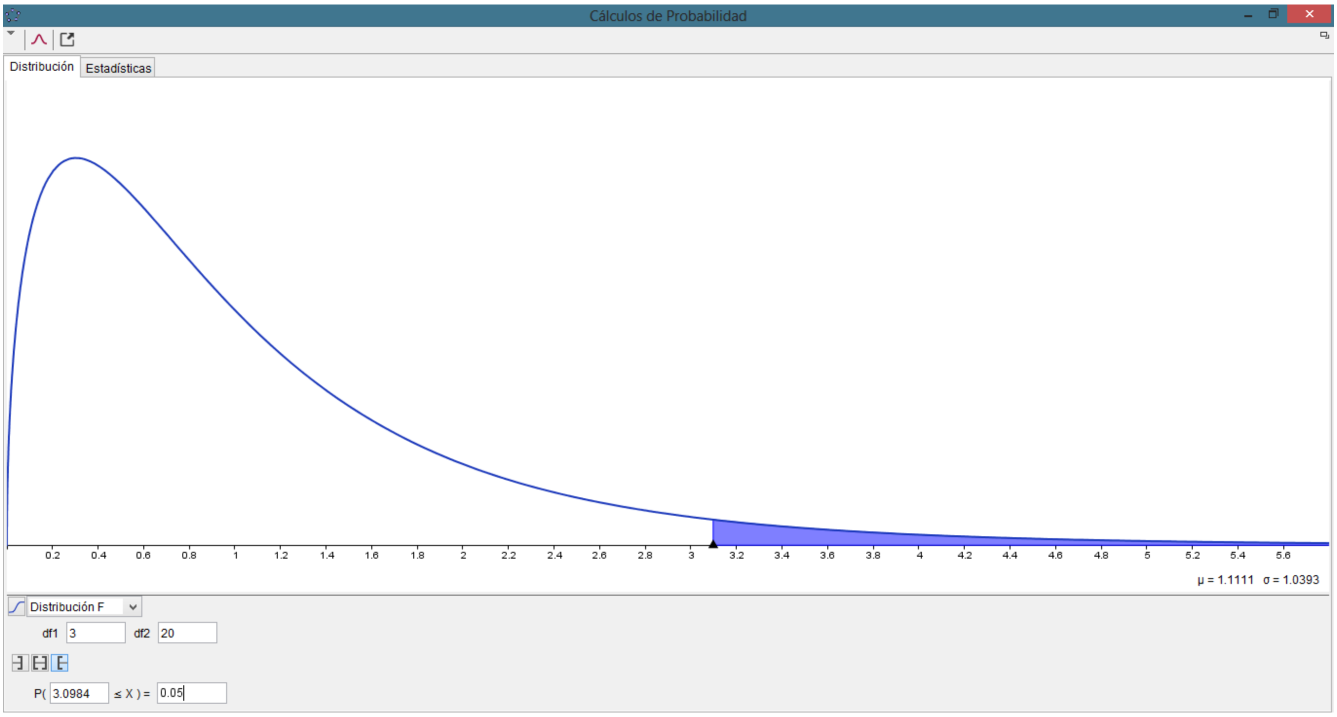
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H
1		Muestra k						
2	Observación n	1	2	3	4	$\bar{x}$		
3	1	10	12	15	18	12,833	=PROMEDIO(B3:B8)	
4	2	11	14	16	15	13,667	=PROMEDIO(C3:C8)	
5	3	13	15	14	14	16,167	=PROMEDIO(D3:D8)	
6	4	14	15	20	15	16,167	=PROMEDIO(E3:E8)	
7	5	15	10	16	20			
8	6	14	16	16	15			
9	s^2	3,77	5,07	4,17	5,37			
10		=VAR.S(B3:B8)	=VAR.S(C3:C8)	=VAR.S(D3:D8)	=VAR.S(E3:E8)			
11								
12	Estimación interna		4,59	=PROMEDIO(B9:E9)		$s_w^2 = \frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_k^2}{k}$		
13	Varianza de las medias aritméticas		2,95	=VAR.S(F3:F6)		$s_{\bar{x}}^2 = \frac{\Sigma(\bar{x} - \bar{\bar{x}})^2}{k-1}$		
14	n		6	=CONTAR(A3:A8)				
15	Estimación intermediante		17,71	=C14*C13		$s_x^2 = n \cdot s_{\bar{x}}^2$		
16	F_{prueba}		3,86	=C15/C12		$F_{prueba} = \frac{s_x^2}{s_w^2}$		
17								
18	k		4	=CONTAR(B2:E2)				
19	$k-1$		3	=C18-1				
20	$k(n-1)$		20	=C18*(C14-1)				
21	α		0,05					
22	F_{tabla}		3,10	=INV.F.CD(C21;C19;C20)				

La gráfica elaborada en Winstats y Paint se muestra en la siguiente figura:



La gráfica elaborada en GeoGebra se muestra en la siguiente figura:



Decisión: Como F_{prueba} es menor que F_{tabla}, H_0 se aprueba, por lo tanto no existen diferencias reales en los pesos de las 4 muestras, es decir, todas las proporciones de la población son iguales**.**

TAREA DE INTERAPRENDIZAJE N° 21

1) ¿Qué es razón F?

2) Describa brevemente las características de la distribución F

3) ¿En qué se asemeja la distribución F a la distribución t?. y ¿En qué se diferencian?

4) Las calificaciones promedio de 4 asignaturas de un curso de cierta universidad se ilustran en la siguiente tabla. La finalidad es determinar si existen diferencias reales entre las calificaciones de las cuatro asignaturas. Emplear un nivel de significación de 0,05. Resuelva de forma manual, con Excel, Winstats y GeoGebra.

Observación	Muestra			
	1	2	3	4
1	6	8	9	9
2	9	7	8	7
3	9	10	8	10
4	6	8	9	10
5	5	9	10	9
Totales	35	42	44	45

Como F_{prueba} es menor que F_{tabla}, H_0 se aprueba, por lo tanto no existen diferencias reales en las calificaciones de las 4 asignaturas.

5) Determinar si existen diferencias reales entre 4 materias que usted recibe, empleando 10 observaciones y un nivel de significación de 0,05. Resuelva de forma manual, con Excel y Winstats

4.3) PRUEBAS DE HIPÓTESIS PARA PROPORCIONES

Las pruebas de proporciones son adecuadas cuando los datos que se están analizando constan de cuentas o frecuencias de elementos de dos o más clases. El objetivo de estas pruebas es evaluar las afirmaciones con respecto a una proporción (o Porcentaje) de población. Las pruebas se basan en la premisa de que una proporción muestral (es decir, x ocurrencias en n observaciones, o x/n) será igual a la proporción verdadera de la población si se toman márgenes o tolerancias para la variabilidad muestral. Las pruebas suelen enfocarse en la diferencia entre un número esperado de ocurrencias, suponiendo que una afirmación es verdadera, y el número observado realmente. La diferencia se compara con la variabilidad prescrita mediante una distribución de muestreo que tiene como base el supuesto de que H_0 es realmente verdadera.

En muchos aspectos, las pruebas de proporciones se parecen a las pruebas de medias, excepto que, en el caso de las primeras, los datos muestrales se consideran como cuentas en lugar de como mediciones. Por ejemplo, las pruebas para medias y proporciones se pueden utilizar para evaluar afirmaciones con respecto a:

1) Un parámetro de población único (prueba de una muestra)

2) La igualdad de parámetros de dos poblaciones (prueba de dos muestras), y

3) La igualdad de parámetros de más de dos poblaciones (prueba de k muestras). Además, para tamaños grandes de muestras, la distribución de muestreo adecuada para pruebas de proporciones de una y dos muestras es aproximadamente normal, justo como sucede en el caso de pruebas de medias de una y dos muestras.

A) PRUEBA DE PROPORCIONES DE UNA MUESTRA

Cuando el objetivo del muestreo es evaluar la validez de una afirmación con respecto a la proporción de una población, es adecuado utilizar una prueba de una muestra. La metodología de prueba depende de si el número de observaciones de la muestra es grande o pequeño.

Como se habrá observado anteriormente, las pruebas de grandes muestras de medias y proporciones son bastante semejantes. De este modo, los valores estadísticos de prueba miden la desviación de un valor estadístico de muestra a partir de un valor propuesto. Y ambas pruebas se basan en la distribución normal estándar para valores críticos. Quizá la única diferencia real entre las ambas radica en la forma corno se obtiene la desviación estándar de la distribución de muestreo.

Esta prueba comprende el cálculo del valor estadístico de prueba Z

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Donde:

$x = ocurrencias$

$n = observaciones$

$\frac{x}{n} = proporción\ de\ la\ muestra$

$p_0 = proporción\ propuesta$

$\sqrt{\frac{p_0(1-p_0)}{n}} = desviación\ estándar\ de\ la\ proporción$

Si se muestra a partir de una población finita

$$\frac{n}{N} \cdot 100\% > 5\%$$

Se debe utilizar el factor finito de corrección

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}}$$

Posteriormente este valor es comparado con el valor de Z, obtenido a partir de una tabla normal a un nivel de significación seleccionado.

Como ocurrió con la prueba de medias de una muestra, las pruebas de proporciones pueden ser de una o dos colas. El tipo de prueba refleja H_1. Por ejemplo, hay tres posibilidades para H_1:

$$H_1\text{: } p > p_0 \qquad H_1\text{: } p < p_0 \qquad H_1\text{: } p \neq p_0$$

La hipótesis nula es: H_0: $p = p_0$

La primera alternativa establece una prueba de cola derecha, la segunda, izquierda y la tercera, una prueba de dos colas.

Ejemplo ilustrativo

En un estudio se afirma que 3 de 10 estudiantes universitarios trabajan. Pruebe esta aseveración, a un nivel de significación de 0,025, respecto a la alternativa de que la proporción real de los estudiantes universitarios trabajan es mayor de lo que se afirma, si una muestra aleatoria de 600 estudiantes universitarios revela que 200 de ellos trabajan. La muestra fue tomada de 10000 estudiantes**.**

Los datos son:

$p_0 = \frac{3}{10} = 0{,}3; \alpha = 0{,}025;\ n = 600;\ X = 200;\ N = 10000$

Las hipótesis son:

$H_0 \colon p = p_0$

$H_1 \colon p > p_0$

Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = 1{,}96$. Se toma en cuenta el valor positivo porque se trata de una prueba de hipótesis a cola derecha.

Como en los datos aparece el tamaño de la población, se debe verificar si el tamaño de la nuestra es mayor que el 5%. Se remplaza valores en la siguiente fórmula:

$$\frac{n}{N} \cdot 100\% > 5\% \Longrightarrow \frac{600}{10000} \cdot 100\% = 6\%$$

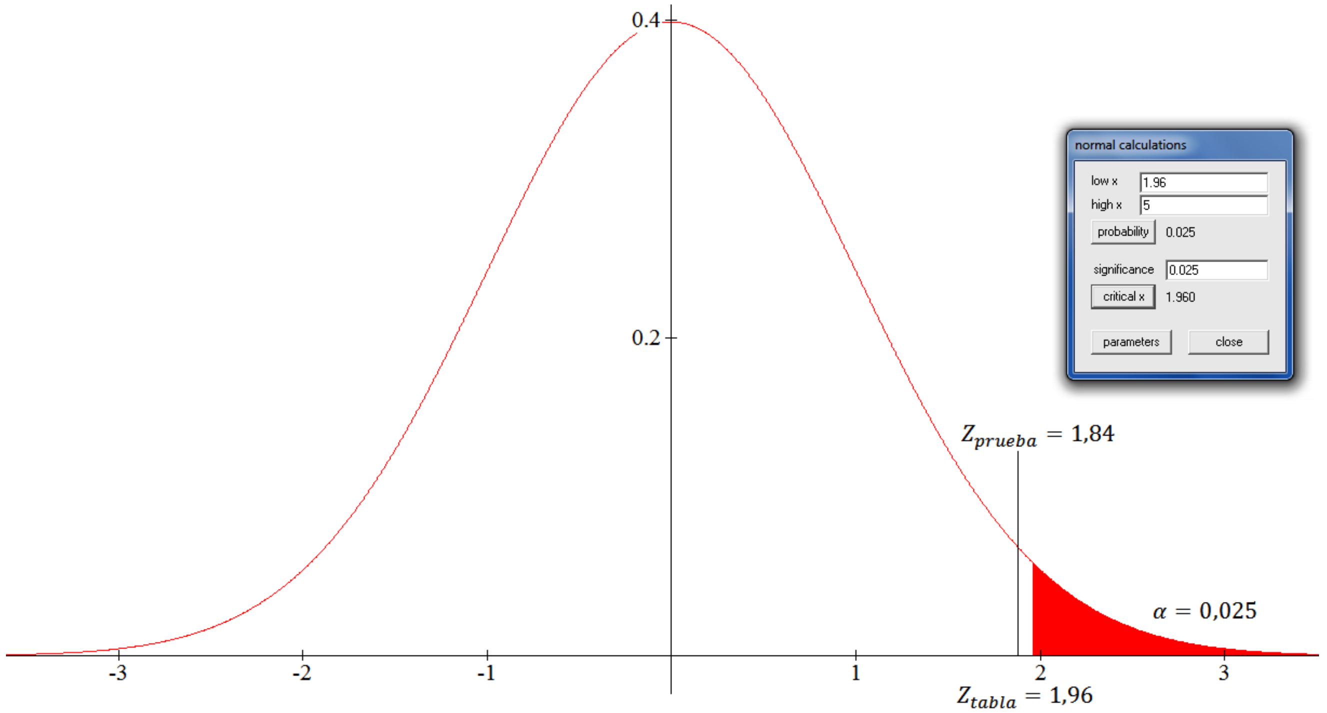
Por lo tanto se debe utilizar la fórmula con el factor finito de corrección.

$$Z_{prueba} = \frac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}} \cdot \sqrt{\frac{N-n}{N-1}}} = \frac{\frac{200}{600} - 0{,}3}{\sqrt{\frac{0{,}3(1-0{,}3)}{600}} \cdot \sqrt{\frac{10000-600}{10000-1}}} = 1{,}84$$

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1	p_0	0,3	=3/10						
2	n	600							
3	x	200							
4	α	0,025							
5	N	10000							
6	$\frac{n}{N}\cdot 100\% > 5\%$	6	=(B2/B5)*100						
7	$H_0: p = p_0$								
8	$H_1: p > p_0$								
9	Z_{tabla}	-1,96	=DISTR.NORM.ESTAND.INV(B4)						
10		1,96	=B8*-1						
11									
12	$z_{prueba} = \dfrac{\frac{x}{n} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}\sqrt{\frac{N-n}{N-1}}}$			1,84	=(B3/B2-B1)/(RAIZ(B1*(1-B1)/B2)*RAIZ((B5-B2)/(B5-1)))				
13									
14									
15									

El gráfico elaborado en Winstats y Paint se muestra a continuación:



Decisión: H_0 es aceptada, ya que $Z_{prueba} = 1{,}84$ es menor que $Z_{tabla} = 1{,}96$, por lo tanto es cierto que 3 de 10 estudiantes universitarios trabajan.

TAREA DE INTERAPRENDIZAJE N° 22

Los siguientes problemas resolver en forma manual (con lectura en la tabla), con Excel y Winstats

1) Un fabricante asegura que un embarque de ciertos sacos contiene menos del 0,5% de piezas defectuosas. Una muestra aleatoria de 100 de ellos presenta 2 defectuosos. Se desea evitar que acepte una remesa con más del 0,5% de sacos defectuosos. Realiza una prueba al nivel 0,02

H_0 es rechazada, ya que z_{prueba} (2,13) es mayor que z_{tabla} (2,05)

2) En un estudio se afirma que 6 de 10 estudiantes universitarios trabajan. Pruebe esta aseveración, a un nivel de significación de 0,025, respecto a la alternativa de que la proporción real de los estudiantes universitarios trabajan es menor de lo que se afirma, si una muestra aleatoria de 200 estudiantes universitarios revela que 110 de ellos trabajan.

H_0 es aceptada, ya que z_{prueba} (-1,44) es mayor que z_{tabla} (-1,96)

3) En una ciudad se afirma que aproximadamente el 90% de los jóvenes en su área terminan la educación secundaria. Ponga a prueba esta aseveración contra la alternativa de que el porcentaje verdadero no es el 90%, y utilice una probabilidad del 5% de que se comete un error de tipo I. Una muestra de 100 jóvenes indica que el 80% de ellos terminan la educación secundaria.

H_0 es rechazada, ya que z_{prueba} (-3,3) está en la zona de rechazo de z_{tabla} (±1,96)

4) Una empresa fabricante de zapatos asegura que menos del 2% de las cajas de zapatos están con defectos. Una comprobación aleatoria revela que 2 de 50 cajas están con defectos. La muestra fue tomada de una remesa de 200 cajas de zapatos. Se quiere evitar que haya demasiadas cajas con defectos. Realice la prueba con un 0,0075 nivel de significación.

H_0 es aceptada, ya que z_{prueba} (1,16) está en la zona de aceptación de z_{tabla} (2,43)

5) Consulte en la biblioteca o en el internet un ejercicio resuelto sobre prueba de hipótesis para proporciones. Presente el ejercicio resuelto empleando Excel y Winstats.

B) PRUEBA DE PROPORCIONES DE DOS MUESTRAS

El objetivo de una prueba de dos muestras es determinar si las dos muestras independientes fueron tomadas de dos poblaciones, las cuales presentan la misma proporción de elementos con determinada característica. La prueba se concentra en la diferencia relativa (diferencia dividida entre la desviación estándar de la distribución de muestreo) entre las dos proporciones muestrales. Diferencias pequeñas denotan únicamente la variación casual producto del muestreo (se acepta H_0), en tanto que grandes diferencias significan lo contrario (se rechaza H_0). El valor estadístico de prueba (diferencia relativa) es comparado con un valor tabular de la distribución normal, a fin de decidir si H_0 es aceptada o rechazada. Una vez más, esta prueba se asemeja considerablemente a la prueba de medias de dos muestras.

La hipótesis nula en una prueba de dos muestras es

$$H_0: p_1 = p_2$$

Las hipótesis alternativas posibles son

$$H_1: p_1 \neq p_2 \quad H_1: p_1 > p_2 \quad H_1: p_1 < p_2$$

La estimación combinada de p se puede calcular de la siguiente manera:

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

Donde:

p = proporción muestral; x_1 = número de aciertos en la muestra 1

x_2 = número de aciertos en la muestra 2 ; n_1 = número de observaciones de la muestra 1

n_2 = número de observaciones de la muestra 2

Este valor de p se utiliza para calcular el valor estadístico de prueba

$$z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Ejemplo ilustrativo

Se ponen a prueba la enseñanza de la Estadística empleando Excel y Winstats. Para determinar si los estudiantes difieren en términos de estar a favor de la nueva enseñanza se toma una muestra de 20 estudiantes de dos paralelos. Del paralelo A 18 están a favor, en tanto que del paralelo B están a favor 14. ¿Es posible concluir con un nivel de significación de 0,05 que los estudiantes que están a favor de la nueva enseñanza de la Estadística es la misma en los dos paralelos?.

Los datos son:
$n_1 = 20\,;\ n_2 = 20;\ x_1 = 18; x_2 = 14; x_2 = 14$

Las hipótesis son:
$H_0: p_1 = p_2; H_1: p_1 \neq p_2$

Como se trata de una prueba de hipótesis a dos colas se debe calcular

$$\frac{\alpha}{2} = \frac{0{,}05}{2} = 0{,}025$$

Con lectura en la tabla para un área de 0,025 le corresponde un valor $Z_{tabla} = \pm 1{,}96$.

Calculando la proporción muestral se obtiene:

$$p = \frac{x_1 + x_2}{n_1 + n_2} = \frac{18 + 14}{20 + 20} = 0{,}8$$

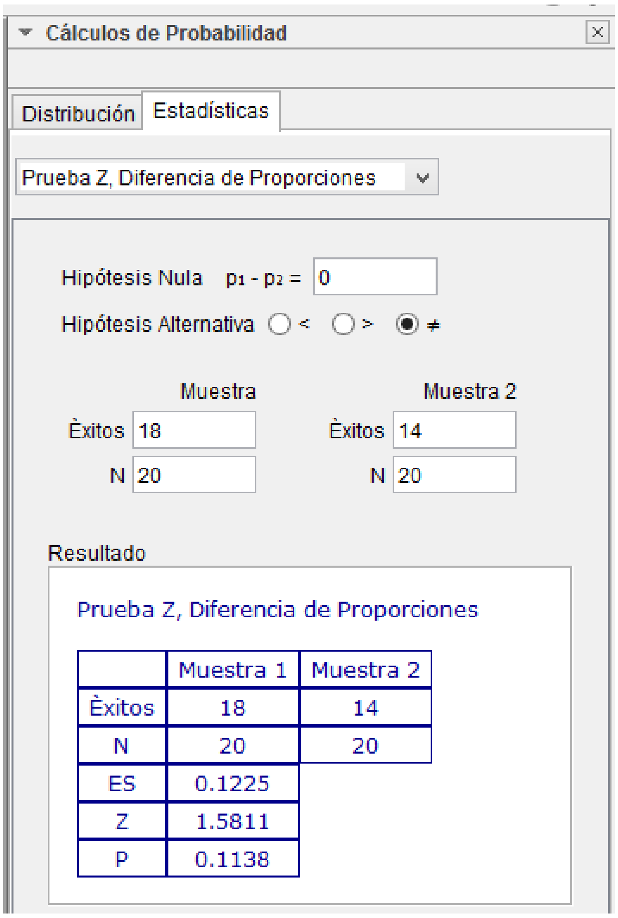
Calculando Z_{prueba} se obtiene:

$$Z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\frac{18}{20} - \frac{14}{20}}{\sqrt{0{,}8(1-0{,}8)\left(\frac{1}{20} + \frac{1}{20}\right)}} = 1{,}58$$

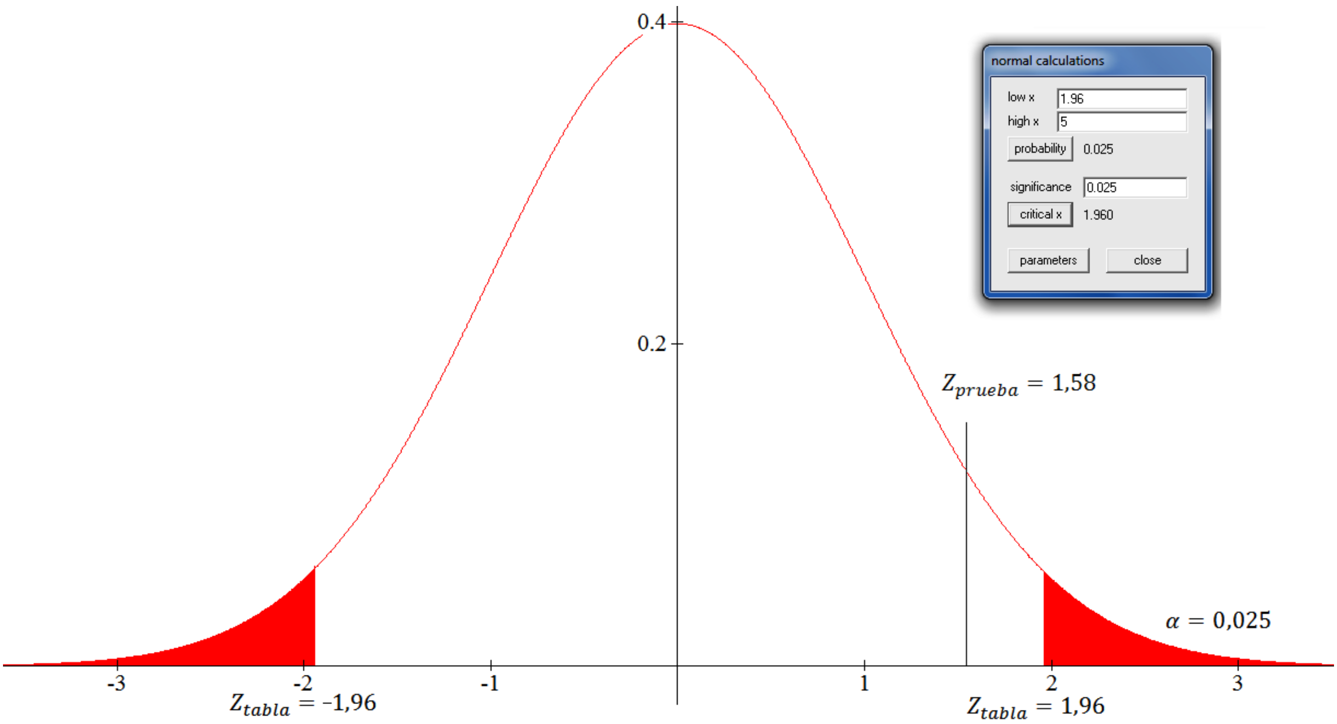
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F
1	n_1	20				
2	n_2	20				
3	x_1	18				
4	x_2	14				
5	α	0,05				
6	$H_0: p_1 = p_2$					
7	$H_1: p_1 \neq p_2$					
8	$\frac{\alpha}{2}$	0,025				
9						
10	Z_{tabla}	-1,96	=DISTR.NORM.ESTAND.INV(B8)			
11		1,96	=B10*-1			
12	$p = \frac{x_1 + x_2}{n_1 + n_2}$	0,8	=(B3+B4)/(B1+B2)			
13						
14						
15	$Z_{prueba} = \frac{\frac{x_1}{n_1} - \frac{x_2}{n_2}}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	1,58	=(B3/B1-B4/B2)/RAIZ(B12*(1-B12)*(1/B1+1/B2))			
16						
17						

Los cálculos en GeoGebra se muestran en la siguiente figura:



El gráfico elaborado en Winstats y Paint se muestra a continuación:



Decisión: H_0 es aceptada, ya que $Z_{prueba} = 1{,}58$ está en la zona de aceptación $Z_{tabla} = \pm 1{,}96$, entonces, la proporción de los estudiantes que están a favor de la nueva enseñanza de la Estadística es la misma en los dos paralelos.

TAREA DE INTERAPRENDIZAJE N° 23

1) ¿Cuál es el objetivo de una prueba de proporciones de dos muestras?

2) ¿En qué caso se acepta o rechaza H_0?

Los siguientes problemas resolver en forma manual (con lectura en la tabla), con Excel, GeoGebra y Winstats

3) A los votantes de dos ciudades se les pregunta si están a favor o en contra de una ley que está actualmente en estudio en la Asamblea Constituyente. Para determinar si los votantes de las dos ciudades difieren en términos del porcentaje que está a favor, se toma una muestra de 100 votantes de cada ciudad. Treinta de los muestreados de una ciudad están a favor, en tanto que, en la otra, lo están veinte. ¿Se podría decir que la proporción de votantes que están a favor de la aprobación de la ley es la misma en estas ciudades?. Realice la prueba con un nivel de significación de 0,05

H_0 es aceptada, ya que z_{prueba} (1,63) está en la zona de aceptación de z_{tabla} (±1,96)

4) Se ponen a prueba dos métodos posibles para tapar botellas. En un lote de 1000, la máquina A genera 30 rechazos, en tanto que la máquina B solamente genera 20. ¿Es posible concluir con un nivel de significación de 0,05 que las dos máquinas son diferentes?

H_0 es aceptada, ya que z_{prueba} (1,43) está en la zona de aceptación de z_{tabla} (±1,96)

5) Consulte en la biblioteca o en internet un problema similar a cualquiera de los anteriores y presente resuelto empleando Excel y GeoGebra.

C) PRUEBA DE PROPORCIONES DE k MUESTRAS

La finalidad de una prueba de k muestras es evaluar la aseveración que establece que todas las k muestras independientes provienen de poblaciones que presentan la misma proporción de algún elemento. De acuerdo con esto, las hipótesis nula y alternativa son:

H_0:Todas las proporciones de la población son iguales.
H_1: No todas las proporciones de la población son iguales.

La estimación combinada de la proporción muestral "p" se calcula de la siguiente manera:

$$p=\frac{x_1+x_2+x_3+\cdots x_n}{n_1+n_2+n_3+\cdots n_n}$$

En una muestra se puede dar un conjunto de sucesos, los cuales ocurren con frecuencias observadas "o" (las que se observa directamente) y frecuencias esperadas o teóricas "e" (las que se calculan de acuerdo a las leyes de probabilidad).

La frecuencia esperada "e" se calcula así: $e=p\cdot o_{total}$
Donde:
p = proporción muestral
o_{total} = frecuencia total observada

El estadístico de prueba es

$$\chi^2_{prueba}=\frac{(o_1-e_1)^2}{e_1}+\frac{(o_2-e_2)^2}{e_2}+\frac{(o_3-e_3)^2}{e_3}+\cdots\frac{(o_n-e_n)^2}{e_n}\Rightarrow\chi^2_{prueba}=\sum\frac{(o_i-e_i)^2}{e_i}$$

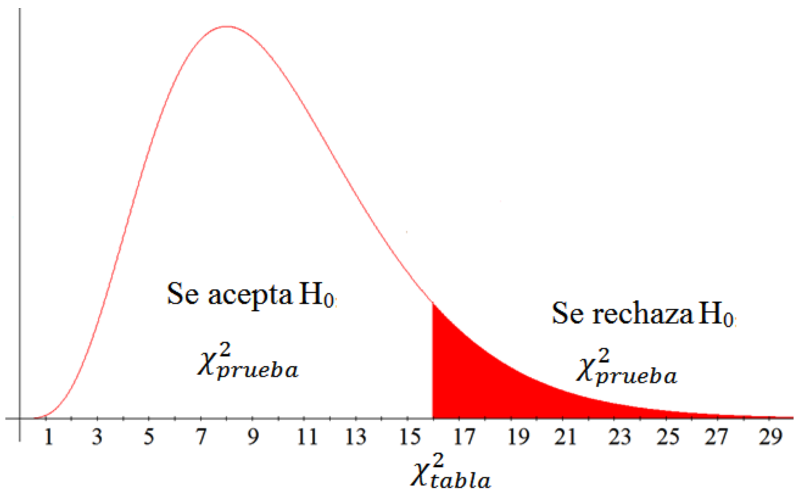
Donde:
χ es la letra griega ji ; χ^2 se lee ji cuadrado

Por lo tanto el valor estadístico de prueba para este caso es la prueba *ji cuadrado* o conocida también como *chi cuadrado*
Como sucede con las distribuciones t y F, la distribución ji cuadrado tiene una forma que depende del número de grados de libertad asociados a un determinado problema.

Para obtener un valor crítico (valor que deja un determinado porcentaje de área en la cola) a partir de una tabla de ji cuadrado, se debe seleccionar un nivel de significación y determinar los grados de libertad para el problema que se esté resolviendo.

Los grados de libertad son una función del número de casillas en una tabla de $2 \cdot k$. Es decir, los grados de libertad reflejan el tamaño de la tabla. Los grados de libertad de la columna son el número de filas (categorías) menos 1, o bien, $r-1$ .Los grados de libertad de cada fila es igual al número de columnas (muestras) menos 1, o bien, $k-1$. El efecto neto es que el número de grados de libertad para la tabla es el producto de (número de filas -1) por (número de columnas -1), o bien, $(r-1)(k-1)$.Por lo tanto con 2 filas y 4 columnas, los grados de libertad son $(2-1)(4-1)=3$

La prueba ji cuadrado requiere la comparación del χ^2_{prueba} con el χ^2_{tabla}. Si el valor estadístico de prueba es menor que el valor tabular, la hipótesis nula es aceptada, caso contrario, H_0 es rechazada.



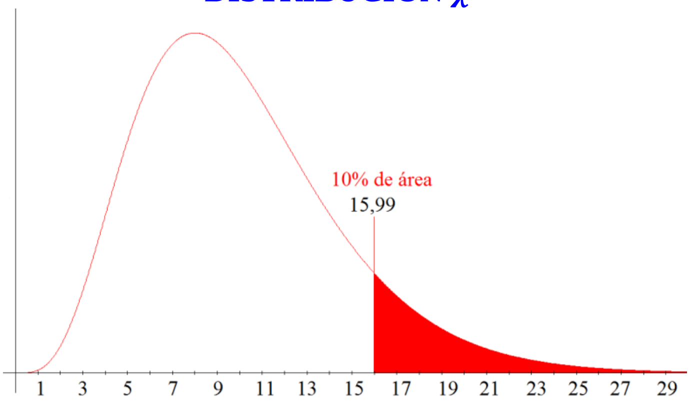
Nota: Un valor estadístico de χ^2_{prueba}menor que el valor crítico χ^2_{tabla} o igual a él se considera como prueba de la variación casual en donde H_0 es aceptada.

Ejemplos ilustrativos:

1) El siguiente valor $4 \cdot 5$ representa el tamaño de una tabla $r \cdot k$. Determine el número de grados de libertad y obtenga el valores crítico en el niveles 0,05 se significación.

Solución:
Los grados de libertad se calculan aplicando la fórmula:
$Grados\ de\ libertad = (r-1)(k-1)$
$Grados\ de\ libertad = (4-1)(5-1) = 12$

TABLA Nº 6
DISTRIBUCIÓN χ^2



Ejemplo:
Para 10 grados de libertad
$P(\chi^2 > 15,99) = 0,10 = 10\%$

	0,995	0,990	0,975	0,950	0,900	0,750	0,500	0,250	0,100	**0,050**	0,025	0,010	0,005
1	0,000	0,000	0,001	0,004	0,016	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340	14,845	18,549	**21,026**	23,337	26,217	28,300

Con lectura en la tabla con 12 grados de libertad y 0,05 de área se obtiene $\chi^2_{tabla} = 21,026$
Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E
1	r	4			
2	k	5			
3					
4	(r-1)(k-1)	12	=(B1-1)*(B2-1)		
5					
6	α	0,05			
7	χ^2_{tabla}	21,0261	=INV.CHICUAD.CD(B6;B4)		
8					
9	α	0,01			
10	χ^2_{tabla}	26,2170	=INV.CHICUAD.CD(B9;B4)		

2) La siguiente tabla muestra las frecuencias observadas y las frecuencias esperadas al lanzar un dado 60 veces. Contrastar la hipótesis de que el dado es bueno, con un nivel de significación de 0,01.

Cara del dado	1	2	3	4	5	6
Frecuencia observada	6	8	9	15	14	8
Frecuencia esperada	10	10	10	10	10	10

Solución:
$r = 2$
$k = 6$
$\alpha = 0,01$

Las hipótesis son:
H_0:Todas las proporciones de la población son iguales.
H_1: No todas las proporciones de la población son iguales.

Los grados de libertad se calculan aplicando la fórmula:
$Grados\ de\ libertad = (2-1)(6-1)$
$Grados\ de\ libertad = 5$

Con lectura en la tabla con 5 grados de libertad y 0,01 de área se obtiene $\chi^2_{tabla} = 15{,}086$

Calculando χ^2_{prueba} se obtiene:

$$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$$

$$\chi^2_{prueba} = \frac{(6-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(15-10)^2}{10} + \frac{(14-10)^2}{10} + \frac{(8-10)^2}{10}$$

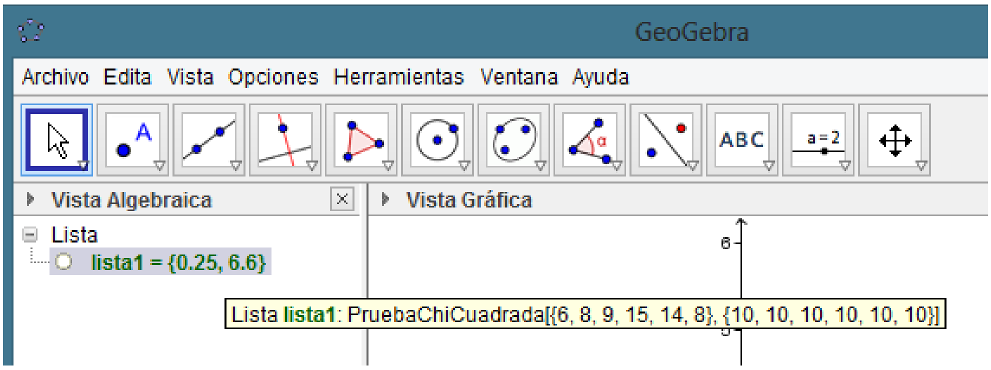
$$\chi^2_{prueba} = 1{,}6 + 0{,}4 + 0{,}1 + 2{,}5 + 1{,}6 + 0{,}4$$

$$\chi^2_{prueba} = 6{,}6$$

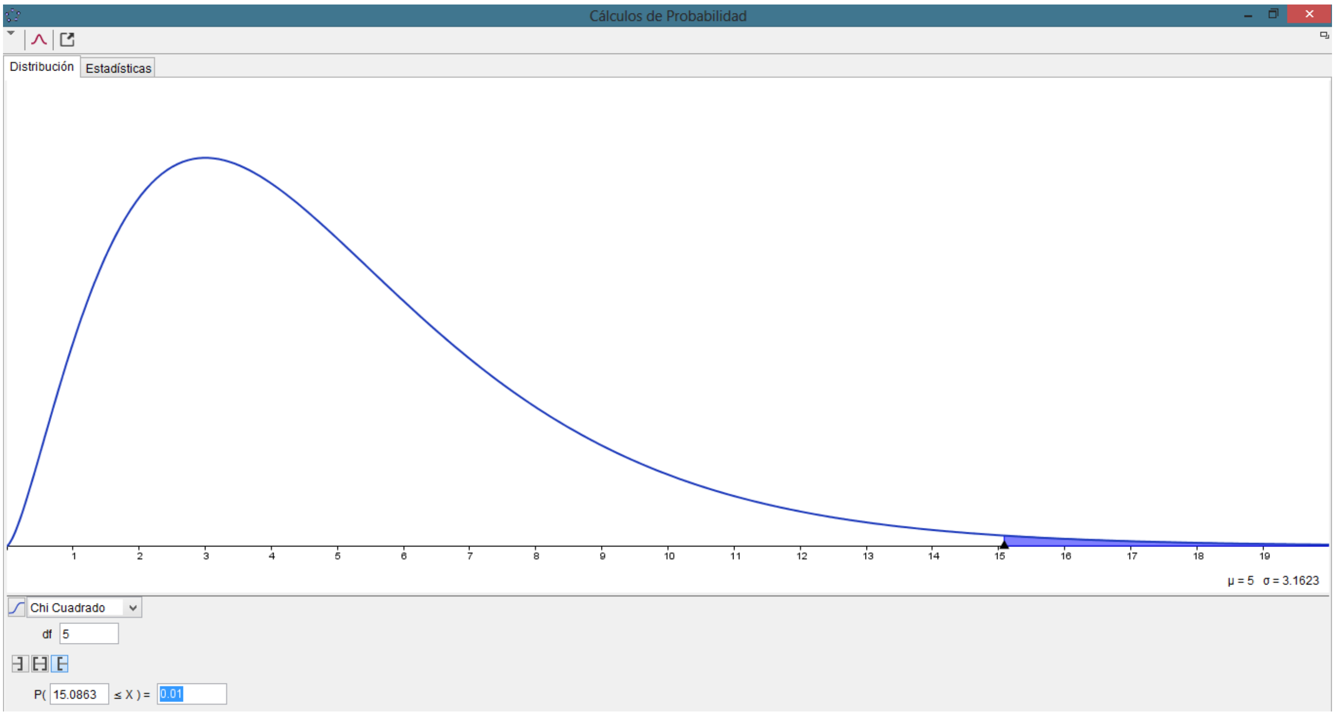
Los cálculos en Excel se muestran en la siguiente figura:

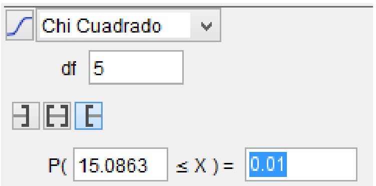
	A	B	C	D	E	F	G
1	Cara del dado	1	2	3	4	5	6
2	Frecuencia observada	6	8	9	15	14	8
3	Frecuencia esperada	10	10	10	10	10	10
4							
5	α	0,01					
6	r	2	=CONTAR(B2:B3)				
7	k	6	=CONTAR(B2:G2)				
8	(r-1)(k-1)	5	=(B6-1)*(B7-1)				
9	χ^2_{tabla}	15,086	=INV.CHICUAD.CD(B5;B8)				
10							
11	Probabilidad de χ^2_{prueba}	0,2521	=PRUEBA.CHICUAD(B2:G2;B3:G3)				
12	$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$	6,6	=INV.CHICUAD.CD(B11;B8)				
13							

Los cálculos en GeoGebra se muestran en la siguiente figura:

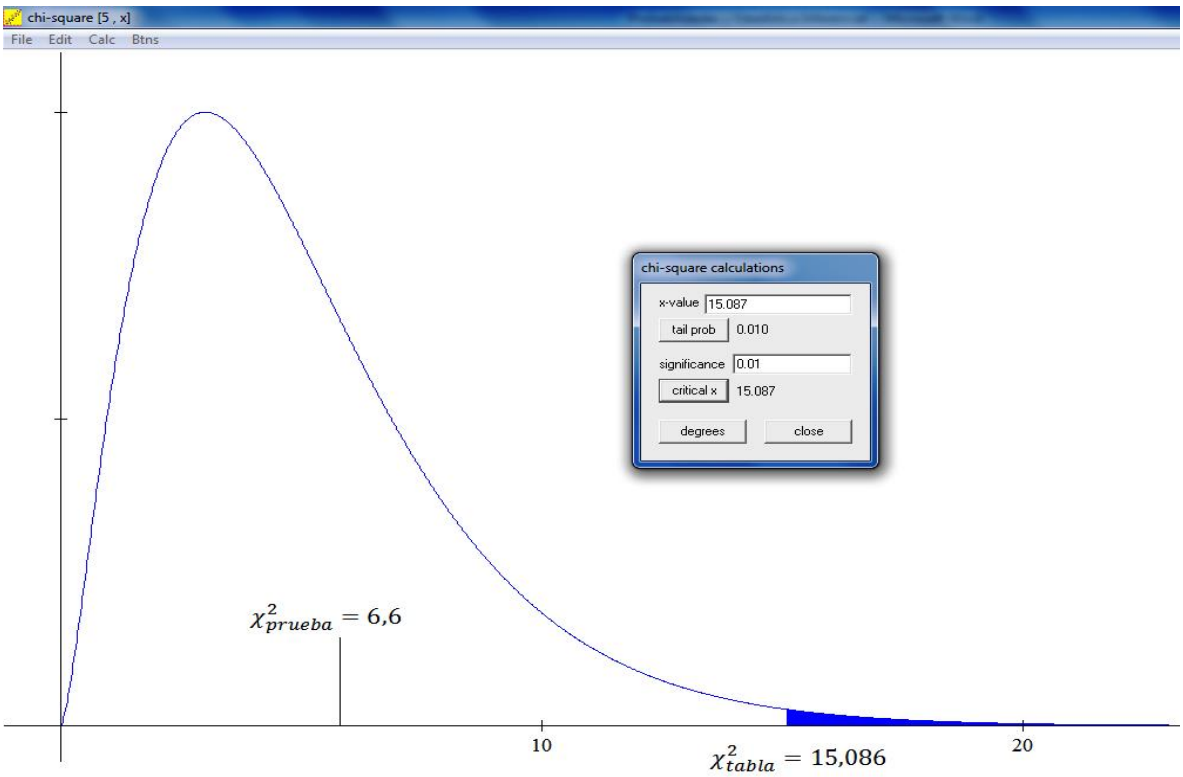


El gráfico elaborado con GeoGebra se muestra a continuación:





El gráfico elaborado en Winstats y Paint se muestra a continuación:



Decisión: H_0 es aceptada, ya que χ^2_{prueba} (6,6) es menor que χ^2_{tabla}(15,086), por lo tanto, se concluye que todas las proporciones de la población son iguales, es decir, el dado es bueno.

TAREA DE INTERAPRENDIZAJE N° 24

1) ¿Cuál es la finalidad de una prueba de k muestras?

2) ¿Qué es valor crítico?

3) Elabore un organizador gráfico sobre Prueba de significación de proporciones de k muestras

Resuelva los ejercicios y problemas de forma manual, empleando Excel, GeoGebra y el Winstats

4) Cada uno de los siguientes valores representa el tamaño de una tabla $r \cdot k$. Determine el número de grados de libertad y obtenga los valores críticos en los niveles 0,05 y 0,01

4.1) $3 \cdot 4$

6; 12,59; 16,81

4.2) $4 \cdot 3$

6; 12,59; 16,81

5) La siguiente tabla muestra las frecuencias observadas y las frecuencias esperadas al lanzar un dado 60 veces. Contrastar la hipótesis de que el dado es bueno, con un nivel de significación de 0,05.

Cara del dado	1	2	3	4	5	6
Frecuencia observada	13	8	7	12	12	8
Frecuencia esperada	10	10	10	10	10	10

H_0 es aceptada, ya que χ^2_{prueba} (3,4) es menor que χ^2_{tabla}(11,07), por lo tanto, se concluye que todas las proporciones de la población son iguales, es decir, el dado es bueno.

6) Un fabricante de helados desea determinar cuál de los sabores es el más popular entre sus clientes. Una muestra de 200 clientes reveló lo siguiente:

Helados	Chocolate	Mora	Fresa	Coco
Frecuencia observada	62	55	45	38
Frecuencia esperada	50	50	50	50

¿Es razonable concluir que hay una diferencia en la proporción de clientes que gustan cada sabor? Use el 0,05 nivel de significación.

H_0 es aceptada, ya que χ^2_{prueba} (6,76) es menor que χ^2_{tabla}(7,81), por lo tanto, se concluye que no existe una diferencia en la proporción de clientes que gustan cada sabor.

7) En un estudio para determinar la preferencia por determinados sabores de helados en diferentes regiones del país, se recopilaron los siguientes datos:

Sabor del helado	Frecuencias observadas por región			
	Costa	Sierra	Oriente	
Vainilla	86	44	70	
Chocolate	45	30	50	
Fresa	34	6	10	
Otros	85	20	20	
Total	250	100	150	

7.1) Calcule proporción muestral "p" de cada sabor del helado

Respuesta: $p_v = 0,4$; $p_{ch} = 0,25$; $p_f = 0,10$; $p_o = 0,25$

7.2) Calcule las frecuencias esperadas de cada sabor del helado en cada región
Respuesta:

Sabor del helado	Frecuencias esperadas por región		
	Costa	Sierra	Oriente
Vainilla	100	40	60
Chocolate	62,5	25	37,5
Fresa	25	10	15
Otros	62,5	25	37,5

7.3) Determine si la preferencia por cierto sabor es independiente de la región (es la misma en cada región), utilizando el nivel de significación 0,05

H_0 se rechaza, ya que χ^2_{prueba} (37,87) es mayor que χ^2_{tabla}(12,592), por lo tanto, se concluye que la preferencia por cierto sabor depende de la región.

8) Consulte en la biblioteca o en el internet un ejercicio resuelto de prueba de hipótesis con Ji Cuadrado. Presente el ejercicio resuelto con Excel y GeoGebra.

D) BONDAD DE AJUSTE DE LA PRUEBA JI CUADRADO

La prueba χ^2 de bondad de ajuste es una variante de la prueba χ^2. El cálculo del valor estadístico de prueba y su evaluación son muy semejantes en ambos casos, aunque existen unas cuantas excepciones. Entre las principales excepciones se encuentran la forma como se plantea H_0 y H_1, cómo se calculan las frecuencias esperadas y cómo se determinan los grados de libertad. Se emplea para determinar la calidad del ajuste mediante distribuciones teóricas (como la distribución normal o la binomial) de distribuciones empíricas (o sea las obtenidas de los datos de la muestra)

En realidad, una prueba de bondad de ajuste es una prueba de una muestra, pero en la que la población se ha dividido en k proporciones. Así pues, difiere de una prueba de proporciones de una muestra, estudiada anteriormente, la cual incluye sólo dos categorías (éxito y fracaso) en la población.

Los grados de libertad para una prueba de bondad de ajuste son (k-1) – c, en donde
k = número de categorías o clases
c = número de valores estadísticos o parámetros utilizados de la muestra que se utilizan para determinar frecuencias esperadas (número de decimales de la frecuencia esperada)

Ejemplo ilustrativo
La siguiente tabla muestra las frecuencias observadas al lanzar un dado 100 veces. Contrastar la hipótesis de que el dado es bueno empleando la Bondad de Ajuste de la Prueba Ji Cuadrado, con un nivel de significación de 0,01.

Cara del dado	1	2	3	4	5	6
Frecuencia observada	18	13	17	22	12	18

Solución:
Las hipótesis son:
H_0:Todas las proporciones de la población son iguales.
H_1: No todas las proporciones de la población son iguales.
El nivel de significación es: $\alpha = 0{,}01$

Se calcula la frecuencia esperada e
$e = p \cdot o_{total}$

$p = proporción\ muestral = probabilidad = \frac{1}{6}$
$o_{total} = frecuencia\ total\ observada = 100$

Cada número del dado (categoría) tiene la misma probabilidad o frecuencia esperada

$e = p \cdot o_{total} = \frac{1}{6} \cdot 100 = 16{,}67$

Número del dado	1	2	3	4	5	6
Frecuencia observada	18	13	17	22	12	18
Frecuencia esperada	16,67	16,67	16,67	16,67	16,67	16,67

Observando la tabla se tiene que:
$k = 6; c = 2$

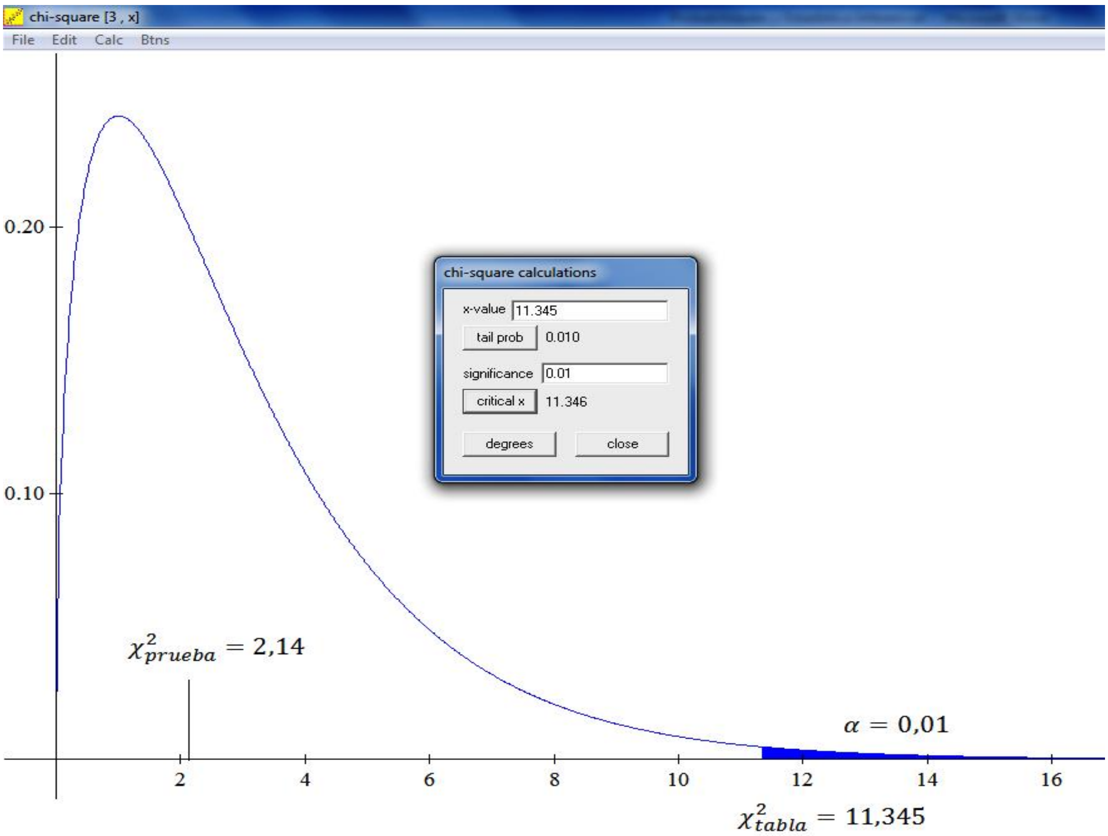
Calculando los grados de libertad se tiene:
$Grados\ de\ libertad = (k-1) - c = (6-1) - 2 = 3$

Con lectura en la tabla con 3 grados de libertad y 0,01 de área se obtiene $\chi^2_{tabla} = 11{,}345$
Calculando χ^2_{prueba} se obtiene:

$$\chi^2_{prueba} = \frac{(18-16{,}67)^2}{16{,}67} + \frac{(13-16{,}67)^2}{16{,}67} + \frac{(17-16{,}67)^2}{16{,}67} + \frac{(22-16{,}67)^2}{16{,}67} + \frac{(12-16{,}67)^2}{16{,}67} + \frac{(18-16{,}67)^2}{16{,}67}$$

$\chi^2_{prueba} = 2{,}14$

El gráfico elaborado en Winstats y Paint se muestra en la siguiente figura:



Decisión: H_0 es aceptada, ya que χ^2_{prueba} (2,14) es menor que χ^2_{tabla}(11,345), por lo tanto, se concluye que todas las proporciones de la población son iguales, es decir, el dado es bueno.

TAREA DE INTERAPRENDIZAJE N° 25

1) Elabore un organizador gráfico sobre Bondad de Ajuste de la Prueba Ji Cuadrado

2) Se pone a prueba un dado para verificar si está cargado o no, es decir, se quiere probar que H_0: El dado no está cargado (es decir, las diferencias se deben únicamente a la variación casual en el muestreo); H_1: El dado está cargado (es decir, que las categorías no son igualmente probables). Cabe esperar que la frecuencia de ocurrencias de cada uno de los seis posibles resultados (categorías) sean igualmente probables. Se lanza el dado 180 veces, con los siguientes resultados

Cara del dado	1	2	3	4	5	6
Frecuencia observada	20	35	25	35	32	33

2.1) Calcular las frecuencias esperadas

30

2.2) Calcular los grados de libertad

5

2.3) Calcular el valor estadístico de χ^2_{tabla} con un nivel de significación de 0,05 empleando la tabla y con Excel.

11,07

2.4) Calcular el valor estadístico de χ^2_{prueba} con la fórmula y con Excel

6,27

2.5) Elabore un gráfico empleando Winstats

2.6) Realizar la comprobación de hipótesis

Se acepta H_0, ya que χ^2_{prueba} (6,27) es menor que χ^2_{tabla}(11,07)

3) Cree y resuelva un problema similar al anterior de forma manual, empleando Excel (para los cálculos) y Winstats (para la gráfica)

4) Consulte en la biblioteca o en el internet sobre las aplicaciones de la Bondad de Ajuste de la Prueba Ji Cuadrado. Presente la consulte a través de un organizador gráfico.

CAPÍTULO V
APLICACIONES DE GRÁFICAS ESTADÍSTICAS EN EL CONTROL DE LA CALIDAD

RESULTADOS DE APRENDIZAJE DEL CAPÍTULO

Al finalizar el presente capítulo el lector podrá evidenciar que:

- ✓ Interpreta las definiciones, características, propiedades y aplicaciones de las gráficas de control de la calidad.
- ✓ Aplica algoritmos relativos a las gráficas de control de la calidad en la resolución de ejercicios y problemas prácticos de manera manual, empleando Excel y Graph.
- ✓ Plantea y resuelve ejercicios de aplicación sobre las gráficas de control de la calidad de manera manual, utilizando Excel y Graph.

CONTENIDOS

- ✓ Introducción
- ✓ Gráficas de control para variables: Gráficas R y $\bar{X}$
- ✓ Gráficas de control para atributos: Gráficas p y c

5.1) INTRODUCCIÓN

Las gráficas de control permiten monitorear la variación en una característica del producto o servicio a lo largo del tiempo. Las gráficas de control se utilizan para estudiar el desempeño pasado, para evaluar las condiciones presentes, o para predecir los resultados futuros. La información obtenida al analizar una gráfica de control constituye la base para el proceso de mejoramiento. Los diferentes tipos de gráficas de control nos permiten analizar diferentes tipos de variables críticas para la calidad (CPC): variables categóricas como la proporción de habitaciones de hotel no aceptables en términos de disponibilidad de comodidades y el correcto funcionamiento de todos los electrodomésticos en la habitación; variables discretas como el número de huéspedes que registraron alguna queja durante la semana; y variables continuas como el tiempo requerido para entregar el equipaje en la habitación. Además de proporcionar una exhibición visual de los datos que representan un proceso, la gráfica de control hace énfasis principalmente en separar las causas especiales de las causas comunes de la variación.

Las causas especiales de variación representan grandes fluctuaciones o patrones en los datos que no son inherentes al proceso. Estas fluctuaciones a menudo son causadas por cambios en el proceso que representan problemas para corregir u oportunidades para aprovechar. Algunas organizaciones se refieren a las causas especiales de variación como causas asignables de variación.

Las causas comunes de variación representan la variabilidad inherente que existe en un proceso. Estas fluctuaciones consisten en numerosas pequeñas causas de variabilidad que operan aleatoriamente o por casualidad, Algunas organizaciones se refieren a las causas comunes de variabilidad como causas aleatorias de variación.

La distinción entre las dos causas de variación es crucial porque las causas especiales de variación no forman parte de un proceso y son corregibles o explotables sin cambiar el sistema. Sin embargo, las causas comunes de variación se reducen tan sólo cambiando el sistema. Estos cambios sistémicos son responsabilidad de la administración.

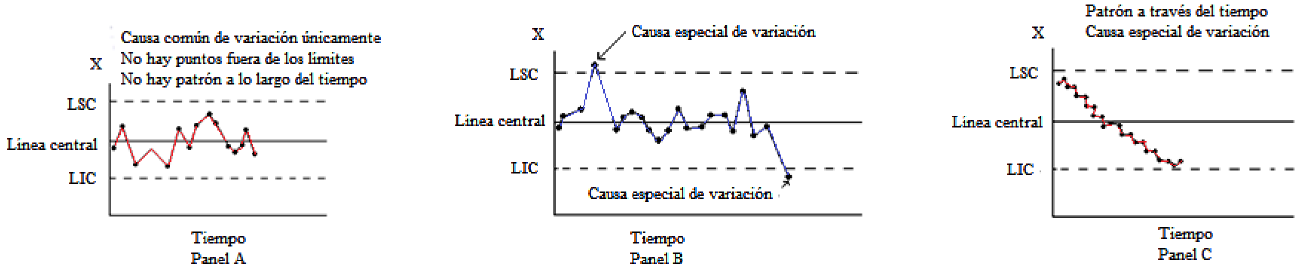
Las gráficas de control nos permiten monitorear un proceso e identificar la presencia o ausencia de causas especiales. Al hacerlo así, las gráficas de control nos ayudan a prevenir dos tipos de errores. El primer tipo de error implica la creencia de que un valor observado representa una causa especial de variación cuando en realidad se debe a una causa común de variación del sistema. Tratar una causa común de variación como si fuera una causa especial de variación a menudo tiene como consecuencia el sobreajuste de un proceso. Este sobreajuste, conocido como manipulación, incrementa la variación del proceso. El segundo tipo de error implica tratar una causa especial de variación como si fuera una causa común de variación. Este error es el resultado de no tomar una acción correctiva inmediata cuando es necesario. Aunque ambos tipos de errores pueden ocurrir aun cuando usemos una gráfica de control, es menos probable que suceda.

Para construir una gráfica de control, se recolectan muestras de las salidas de un proceso a lo largo del tiempo. Las muestras utilizadas para construir gráficas de control se conocen como subgrupos. Para cada subgrupo (es decir, muestra), se calcula el valor de un estadístico asociado con una variable CPC. Los estadísticos utilizados comúnmente incluyen la fracción disconforme y la media y el rango de una variable numérica. Entonces se grafican los valores contra el tiempo y se agregan los límites de control a la gráfica. La forma más común de gráfica de control establece límites de control que están dentro de ± 3 desviaciones estándar de la medida estadística de interés. La ecuación media del proceso $\pm$ 3 desviaciones estándar define, en general, los límites de control superior e inferior para las gráficas de control.

Construcción de límites de control

Media del proceso ±3 desviaciones estándar .Por lo que el Límite de control superior (LCS) = media del proceso +3 desviaciones estándar
Límite de control inferior (LIC) = media del proceso -3 desviaciones estándar

Cuando se establecen estos límites de control, se evalúa la gráfica de control tratando de encontrar un patrón que pudiera existir en los valores a lo largo del tiempo y determinando si algunos puntos caen fuera de los límites de control. Las siguientes figuras ilustran tres diferentes situaciones.



En el panel A, no existe un patrón aparente de los valores a lo largo del tiempo y no hay puntos que caigan fuera del límite de control de 3 desviaciones estándar. El proceso parece estable y contiene sólo causas comunes de variación. El panel B, por el contrario, contiene dos puntos que caen fuera de los límites de control de las 3 desviaciones estándar. Se deben investigar estos puntos para tratar de determinar las causas especiales que llevan a su ocurrencia. Aunque el panel C no tiene ningún punto fuera de los límites de control, tiene una serie de puntos consecutivos por arriba del valor promedio (la línea central), así como una serie de puntos consecutivos por debajo del valor promedio. Además, se observa claramente una tendencia global descendente. Se debe investigar esta situación para tratar de determinar qué podría haber causado ese patrón.

Detectar una tendencia no es siempre tan obvio. Hay otras dos reglas simples que nos permiten detectar un cambio en el nivel medio de un proceso:
- Ocho o más puntos consecutivos que caen por arriba de la línea central u ocho o más puntos consecutivos que caen por debajo de la línea central.
- Ocho o más puntos consecutivos se mueven hacia arriba en valor u ocho o más puntos consecutivos se mueven hacia abajo en valor.

Se dice que un proceso cuya gráfica de control indica una condición fuera de control (un punto fuera de los límites de control o la exhibición de una tendencia) está fuera de control. Un proceso fuera de control contiene tanto causas comunes de variación como causas especiales de variación. Puesto que las causas especiales de variación no forman parte del diseño del proceso, un proceso fuera de control es impredecible. Una vez que se determina que un proceso está fuera de control, se deben identificar las causas especiales de variación que están provocando las condiciones fuera de control. Si las causas especiales actúan en detrimento de la calidad del producto o servicio, se requiere elaborar planes para eliminar esta fuente de variación. Cuando una causa especial incrementa la calidad, se debería cambiar el proceso para que la causa especial se incorpore dentro del diseño del proceso. Por lo tanto, esta causa especial benéfica se vuelve una causa común fuente de variación y el proceso se mejora.

Se dice que un proceso cuya gráfica de control no indica condiciones fuera de control está bajo control. Un proceso bajo control contiene únicamente causas comunes de variación. En ocasiones se dice que los procesos bajo control están en un estado de control estadístico. Cuando un proceso se encuentra bajo control, usted debe determinar si la cantidad de causa común de variación en el proceso es lo suficientemente pequeña como para satisfacer a los usuarios de los productos o servicios. Si la causa común de variación es lo suficientemente pequeña como para satisfacer al cliente, entonces se utiliza la gráfica de control para monitorear el proceso sobre una base continua para asegurarse de que el proceso permanece bajo control. Si la causa común de variación es demasiado grande, se requiere alterar el proceso en sí mismo.

TAREA DE INTERAPRENDIZAJE N° 26

1) Consulte en la biblioteca o en el internet sobre la administración o control de la calidad total y realice un organizador gráfico de la misma.
2) Consulte en la biblioteca o en el internet sobre la administración seis sigma y realice un organizador gráfico de la misma.
3) ¿Piensa que el control de la calidad tiene impacto en nuestro trabajo diario y en nuestras vidas personales?
4) Escriba una semejanza y una diferencia entre causas especiales de variación y causas comunes de variación
5) ¿Cuáles son los dos tipos de errores que las gráficas de control nos ayudan a prevenir?
6) ¿Qué es un límite de control?
7) ¿En qué situación se dice que un proceso está fuera de control?. Ilustre su respuesta con una gráfica de control
8) ¿En qué situación se dice que un proceso está bajo control?. Ilustre su respuesta con una gráfica de control

5.2) GRÁFICAS DE CONTROL PARA VARIABLES

Estas gráficas de control ayudan a la detección de la variación de causa asignable (variación en el producto o proceso de producción que señala que el proceso está fuera de control y que se requieren medidas correctivas)

A) La Gráfica R

Mide la variación en el rango de las muestras. Aunque la desviación estándar es una medida que depende de la dispersión, las técnicas de control de calidad generalmente confían en el rango como un indicio de la variabilidad del proceso.
Límite superior de control para el rango

$$LSC_R = \bar{R} + 3s_R$$

Límite inferior de control para el rango

$$LIC_R = \bar{R} - 3s_R$$

Donde s_R es la desviación estándar en los rangos muestrales. Sin embrago, en la práctica, es más simple de utilizar

Límite superior de control para el rango

$$LSC_R = D_4\bar{R}$$

Límite inferior de control para el rango

$$LIC_R = D_3\bar{R}$$

Los valores D_4 y D_3 se toman de la tabla de factores críticos de las gráficas o cartas de control de acuerdo al tamaño n de la muestra y el rango promedio de los rangos muestrales $\bar{R} = \frac{\Sigma \mathrm{R}}{\mathrm{k}}$, siendo k = número de muestras

B) La Gráfica $\bar{X}$

Se diseña para medir la variación en las medias muestrales alrededor de algún nivel generalmente aceptado.
Límite superior de control para el rango

$$LSC_{\bar{X}} = \bar{\bar{X}} + 3\sigma_{\bar{X}}$$

Límite inferior de control para el rango

$$LIC_{\bar{X}} = \bar{\bar{X}} - 3\sigma_{\bar{X}}$$

Donde $\sigma_{\bar{X}}$ es la desviación estándar para las medias. Sin embrago, en la práctica, se estima $3\sigma_{\bar{X}}$ como $A_2\bar{R}$, en donde $\bar{R}$ es el rango promedio de los rangos muestrales, y A_2 es una constante basada en el tamaño de la muestra. Los valores de A_2 se hallan en la tabla de factores críticos de las gráficas o cartas de control. Utilizando $A_2\bar{R}$ en lugar de $3\sigma_{\bar{X}}$ produce resultados similares y es considerablemente más fácil de calcular. Se tiene entonces:

Límite superior de control para las medias

$$LSC_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}$$

Límite inferior de control para las medias

$$LIC_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R}$$

Donde:

$$\bar{\bar{X}} = \frac{\Sigma\bar{X}}{k}$$

Siendo k = número de muestras

TABLA
Factores críticos de las gráficas o cartas de control

	Gráfica para medias	Gráfica para rangos			
n	Factor para el límite de control $A_2 = 3/(d_2\sqrt{n})$	Factor para la recta central d_2	Factores de los límites de control $D_3 = 1-3(d_3/d_2)$	$D_4 = 1+3(d_3/d_2)$	d_3
2	1,881	1,128	-1,267=0	3,267	0,8525
3	1,023	1,693	-0,574=0	2,574	0,8884
4	0,729	2,059	-0,282=0	2,282	0,8798
5	0,577	2,326	-0,114=0	2,114	0,8641
6	0,483	2,534	-0,004=0	2,004	0,8480
7	0,419	2,704	0,076	1,924	0,8330
8	0,373	2,847	0,136	1,864	0,8200
9	0,337	2,970	0,184	1,816	0,8080
10	0,308	3,078	0,223	1,777	0,7970
11	0,285	3,173	0,256	1,744	0,7870
12	0,266	3,258	0,284	1,716	0,7780
13	0,249	3,336	0,308	1,692	0,7700
14	0,235	3,407	0,329	1,671	0,7620
15	0,223	3,472	0,348	1,652	0,7550
16	0,212	3,532	0,364	1,636	0,7490
17	0,203	3,588	0,379	1,621	0,7430
18	0,194	3,640	0,392	1,608	0,7380
19	0,187	3,689	0,404	1,596	0,7330
20	0,180	3,735	0,414	1,586	0,7290
21	0,173	3,778	0,425	1,575	0,7240
22	0,167	3,819	0,434	1,566	0,7200
23	0,162	3,858	0,443	1,557	0,7160
24	0,157	3,895	0,452	1,548	0,7120
25	0,153	3,931	0,459	1,541	0,7090

Fuente: WEBSTER, Allen, (2000), *Estadística Aplicada a los Negocios y a la Economía*, Ed. McGraw Hill.

Ejemplo ilustrativo

Una fábrica elabora planchas de madera para tapas de mesas, las cuales deben cumplir ciertas especificaciones de tamaño. Para garantizar que se cumplan estos estándares de calidad, se recolecta K= 24 muestras (subgrupos) de tamaño n = 6, y mide su largo. Los resultados aparecen en la siguiente tabla:

Nº de muestra	Medias muestrales					
1	14,5	15,9	15,7	16,3	14,5	16,2
2	15,4	15,2	15,9	15,2	14,5	14,5
3	16,5	15,9	14,8	16,2	16,5	16,2
4	14,8	16,8	15,5	15,2	15,2	14,2
5	15,7	14,5	16,9	14,2	14,5	15,2
6	15,9	15,4	17,1	14,8	16,8	14,8
7	15,2	14,2	18,5	15,8	15,9	15,7
8	14,5	14,8	17,2	16,2	15,0	16,8
9	15,6	15,7	19,2	16,1	16,8	15,9
10	16,5	16,8	18,4	14,8	18,9	16,1
11	14,5	15,8	14,2	14,5	18,7	16,3
12	17,1	15,8	16,2	15,4	15,7	16,2
13	18,5	15,9	17,2	14,2	15,9	14,7
14	17,2	15,7	16,8	14,8	14,8	14,9
15	19,2	15,7	15,9	15,7	15,5	14,8
16	18,4	16,8	15,0	16,8	16,9	14,7
17	14,2	16,9	16,8	15,8	17,1	15,4
18	16,2	17,2	18,9	15,8	18,5	18,9
19	17,2	17,6	18,7	15,9	17,2	16,0
20	16,8	14,5	19,8	15,7	18,2	18,7
21	15,9	17,9	18,7	15,7	18,4	17,5
22	15,0	18,0	18,2	16,8	14,2	17,8
23	16,8	18,9	20,0	16,9	16,2	18,5
24	18,9	17,9	17,4	17,5	17,2	16,5

a) Calcular el rango promedio
b) Calcular el límite superior de control para el rango
c) Calcular el límite inferior de control para el rango
d) Elaborar la gráfica R.
e) Calcular $\bar{\bar{X}}$
f) Calcular el límite superior de control para las medias
g) Calcular el límite inferior de control para las medias
h) Elaborar la gráfica $\bar{X}$.

Solución:

Calculando manualmente el rango se obtiene:
Recuerde que el rango es igual al valor mayor menes el valor menor, es decir:

$$R = X_{máxima} - X_{minima}$$

Nº de muestra	Medias muestrales						R
1	14,5	15,9	15,7	16,3	14,5	16,2	1,8
2	15,4	15,2	15,9	15,2	14,5	14,5	1,4
3	16,5	15,9	14,8	16,2	16,5	16,2	1,7
4	14,8	16,8	15,5	15,2	15,2	14,2	2,6
5	15,7	14,5	16,9	14,2	14,5	15,2	2,7
6	15,9	15,4	17,1	14,8	16,8	14,8	2,3
7	15,2	14,2	18,5	15,8	15,9	15,7	4,3
8	14,5	14,8	17,2	16,2	15,0	16,8	2,7
9	15,6	15,7	19,2	16,1	16,8	15,9	3,6
10	16,5	16,8	18,4	14,8	18,9	16,1	4,1
11	14,5	15,8	14,2	14,5	18,7	16,3	4,5
12	17,1	15,8	16,2	15,4	15,7	16,2	1,7
13	18,5	15,9	17,2	14,2	15,9	14,7	4,3
14	17,2	15,7	16,8	14,8	14,8	14,9	2,4
15	19,2	15,7	15,9	15,7	15,5	14,8	4,4
16	18,4	16,8	15,0	16,8	16,9	14,7	3,7
17	14,2	16,9	16,8	15,8	17,1	15,4	2,9
18	16,2	17,2	18,9	15,8	18,5	18,9	3,1
19	17,2	17,6	18,7	15,9	17,2	16,0	2,8
20	16,8	14,5	19,8	15,7	18,2	18,7	5,3
21	15,9	17,9	18,7	15,7	18,4	17,5	3,0
22	15,0	18,0	18,2	16,8	14,2	17,8	4,0
23	16,8	18,9	20,0	16,9	16,2	18,5	3,8
24	18,9	17,9	17,4	17,5	17,2	16,5	2,4
Total							**75,5**

a) Calculando el rango promedio se tiene:

$$\bar{R} = \frac{\Sigma \mathrm{R}}{\mathrm{k}} = \frac{75{,}5}{24} = 3{,}146$$

b) Calcular el límite superior de control para el rango
Con lectura en la tabla para n = 6 se obtiene $D_4 = 2{,}004$

Calculando el límite superior se obtiene:

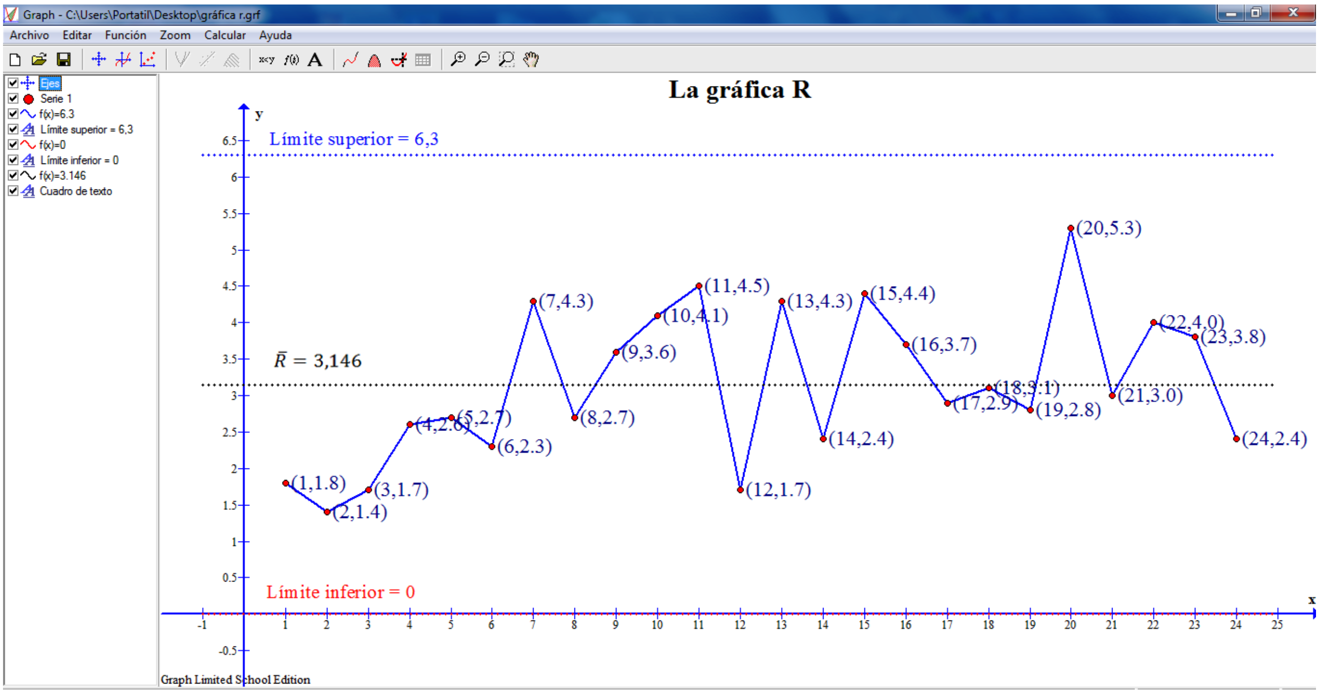
$$LSC_R = D_4\bar{R} = 2{,}004 \cdot 3{,}146 = 6{,}3$$

c) Calcular el límite inferior de control para el rango
Con lectura en la tabla para n = 6 se obtiene $D_3 = 0$

Calculando el límite inferior se obtiene:

$$LIC_R = D_3\bar{R} = 0 \cdot 3{,}146 = 0$$

d) Elaborando la gráfica R en Graph se obtiene:



Interpretación: Observando la gráfica se concluye que la misma está bajo control, ya que no existen variaciones de causa asignable, es decir, no existe ningún punto que se salga de los límites de control.

e) Calculando $\bar{x}$ se obtiene:

Nº de muestra	Medias muestrales						$\bar{X}$
1	14,5	15,9	15,7	16,3	14,5	16,2	15,52
2	15,4	15,2	15,9	15,2	14,5	14,5	15,12
3	16,5	15,9	14,8	16,2	16,5	16,2	16,02
4	14,8	16,8	15,5	15,2	15,2	14,2	15,28
5	15,7	14,5	16,9	14,2	14,5	15,2	15,17
6	15,9	15,4	17,1	14,8	16,8	14,8	15,80
7	15,2	14,2	18,5	15,8	15,9	15,7	15,88
8	14,5	14,8	17,2	16,2	15,0	16,8	15,75
9	15,6	15,7	19,2	16,1	16,8	15,9	16,55
10	16,5	16,8	18,4	14,8	18,9	16,1	16,92
11	14,5	15,8	14,2	14,5	18,7	16,3	15,67
12	17,1	15,8	16,2	15,4	15,7	16,2	16,07
13	18,5	15,9	17,2	14,2	15,9	14,7	16,07
14	17,2	15,7	16,8	14,8	14,8	14,9	15,70
15	19,2	15,7	15,9	15,7	15,5	14,8	16,13
16	18,4	16,8	15,0	16,8	16,9	14,7	16,43
17	14,2	16,9	16,8	15,8	17,1	15,4	16,03
18	16,2	17,2	18,9	15,8	18,5	18,9	17,58
19	17,2	17,6	18,7	15,9	17,2	16,0	17,10
20	16,8	14,5	19,8	15,7	18,2	18,7	17,28
21	15,9	17,9	18,7	15,7	18,4	17,5	17,35
22	15,0	18,0	18,2	16,8	14,2	17,8	16,67
23	16,8	18,9	20,0	16,9	16,2	18,5	17,88
24	18,9	17,9	17,4	17,5	17,2	16,5	17,57
Total							**391,53**

Calculando $\bar{\bar{X}}$ se obtiene:

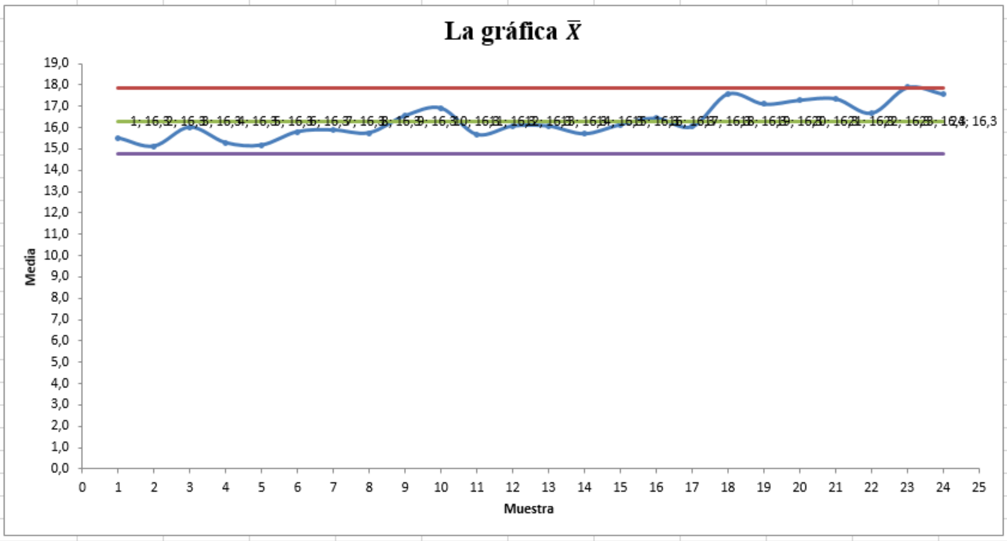
$$\bar{\bar{X}} = \frac{\Sigma \bar{X}}{k} = \frac{391{,}53}{24} = 16{,}314$$

f) Con lectura en la tabla para n = 6 se obtiene $A_2 = 0{,}483$
Calculando el límite superior se obtiene: $LSC_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R} = 16{,}314 + 0{,}483 \cdot 3{,}146 = 17{,}83$

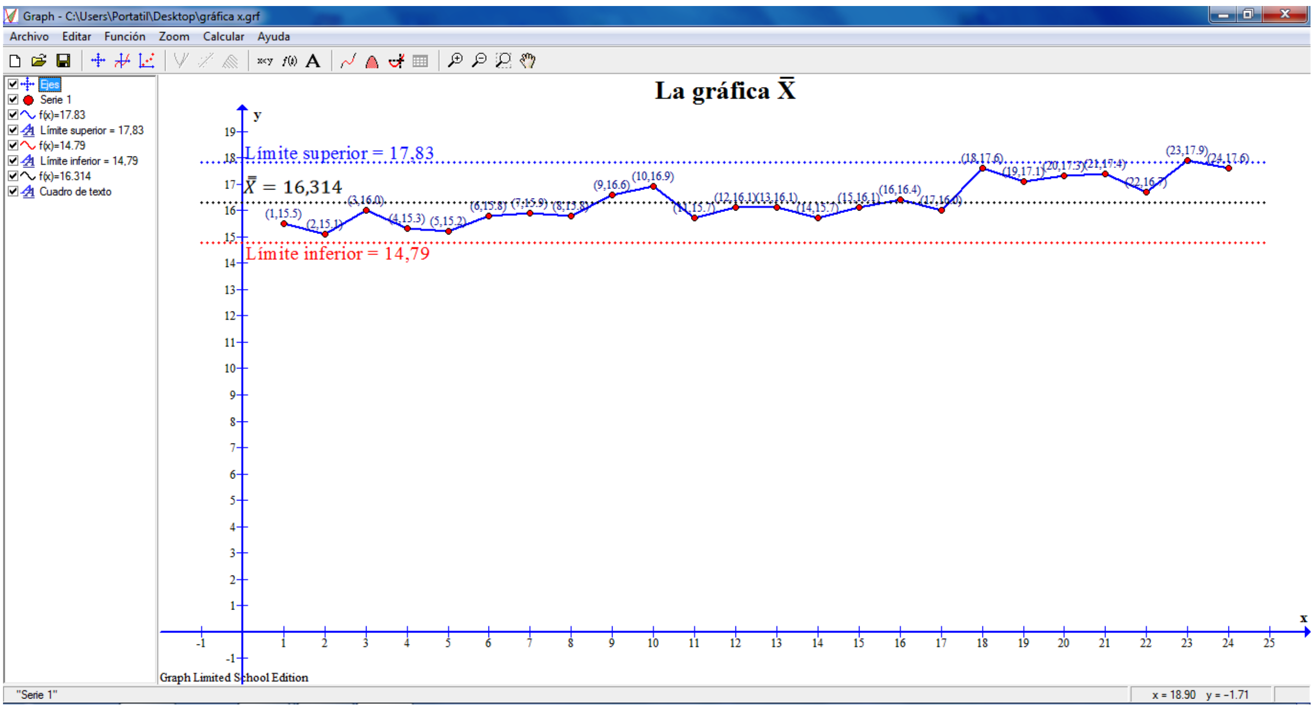
g) Calculando el límite inferior se obtiene: $LIC_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} = 16{,}314 - 0{,}483 \cdot 3{,}146 = 14{,}79$

h) Elaborando la gráfica $\bar{X}$ en Excel se obtiene:

	A	B	C	D	E
1	1	15,5	17,83	16,3	14,79
2	2	15,1	17,83	16,3	14,79
3	3	16,0	17,83	16,3	14,79
4	4	15,3	17,83	16,3	14,79
5	5	15,2	17,83	16,3	14,79
6	6	15,8	17,83	16,3	14,79
7	7	15,9	17,83	16,3	14,79
8	8	15,8	17,83	16,3	14,79
9	9	16,6	17,83	16,3	14,79
10	10	16,9	17,83	16,3	14,79
11	11	15,7	17,83	16,3	14,79
12	12	16,1	17,83	16,3	14,79
13	13	16,1	17,83	16,3	14,79
14	14	15,7	17,83	16,3	14,79
15	15	16,1	17,83	16,3	14,79
16	16	16,4	17,83	16,3	14,79
17	17	16,0	17,83	16,3	14,79
18	18	17,6	17,83	16,3	14,79
19	19	17,1	17,83	16,3	14,79
20	20	17,3	17,83	16,3	14,79
21	21	17,4	17,83	16,3	14,79
22	22	16,7	17,83	16,3	14,79
23	23	17,9	17,83	16,3	14,79
24	24	17,6	17,83	16,3	14,79



Elaborando la gráfica $\bar{X}$ en Graph se obtiene:



Interpretación: Observando la gráfica se concluye que la misma está fuera de control, ya que, la muestra 23 representa una variación de causa asignable, es decir, la muestra 23 se sale del límite superior de control.

Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Nº Muestra:k	Medias muestrales:n						R		$\bar{X}$				
2	1	14,5	15,9	15,7	16,3	14,5	16,2	1,8	=MAX(B2:G2)-MIN(B2:G2)	15,5	=PROMEDIO(B2:G2)			
3	2	15,4	15,2	15,9	15,2	14,5	14,5	1,4	=MAX(B3:G3)-MIN(B3:G3)	15,1	=PROMEDIO(B3:G3)			
4	3	16,5	15,9	14,8	16,2	16,5	16,2	1,7	=MAX(B4:G4)-MIN(B4:G4)	16,0	=PROMEDIO(B4:G4)			
5	4	14,8	16,8	15,5	15,2	15,2	14,2	2,6	=MAX(B5:G5)-MIN(B5:G5)	15,3	=PROMEDIO(B5:G5)			
6	5	15,7	14,5	16,9	14,2	14,5	15,2	2,7	=MAX(B6:G6)-MIN(B6:G6)	15,2	=PROMEDIO(B6:G6)			
7	6	15,9	15,4	17,1	14,8	16,8	14,8	2,3	=MAX(B7:G7)-MIN(B7:G7)	15,8	=PROMEDIO(B7:G7)			
8	7	15,2	14,2	18,5	15,8	15,9	15,7	4,3	=MAX(B8:G8)-MIN(B8:G8)	15,9	=PROMEDIO(B8:G8)			
9	8	14,5	14,8	17,2	16,2	15,0	16,8	2,7	=MAX(B9:G9)-MIN(B9:G9)	15,8	=PROMEDIO(B9:G9)			
10	9	15,6	15,7	19,2	16,1	16,8	15,9	3,6	=MAX(B10:G10)-MIN(B10:G10)	16,6	=PROMEDIO(B10:G10)			
11	10	16,5	16,8	18,4	14,8	18,9	16,1	4,1	=MAX(B11:G11)-MIN(B11:G11)	16,9	=PROMEDIO(B11:G11)			
12	11	14,5	15,8	14,2	14,5	18,7	16,3	4,5	=MAX(B12:G12)-MIN(B12:G12)	15,7	=PROMEDIO(B12:G12)			
13	12	17,1	15,8	16,2	15,4	15,7	16,2	1,7	=MAX(B13:G13)-MIN(B13:G13)	16,1	=PROMEDIO(B13:G13)			
14	13	18,5	15,9	17,2	14,2	15,9	14,7	4,3	=MAX(B14:G14)-MIN(B14:G14)	16,1	=PROMEDIO(B14:G14)			
15	14	17,2	15,7	16,8	14,8	14,8	14,9	2,4	=MAX(B15:G15)-MIN(B15:G15)	15,7	=PROMEDIO(B15:G15)			
16	15	19,2	15,7	15,9	15,7	15,5	14,8	4,4	=MAX(B16:G16)-MIN(B16:G16)	16,1	=PROMEDIO(B16:G16)			
17	16	18,4	16,8	15,0	16,8	16,9	14,7	3,7	=MAX(B17:G17)-MIN(B17:G17)	16,4	=PROMEDIO(B17:G17)			
18	17	14,2	16,9	16,8	15,8	17,1	15,4	2,9	=MAX(B18:G18)-MIN(B18:G18)	16,0	=PROMEDIO(B18:G18)			
19	18	16,2	17,2	18,9	15,8	18,5	18,9	3,1	=MAX(B19:G19)-MIN(B19:G19)	17,6	=PROMEDIO(B19:G19)			
20	19	17,2	17,6	18,7	15,9	17,2	16,0	2,8	=MAX(B20:G20)-MIN(B20:G20)	17,1	=PROMEDIO(B20:G20)			
21	20	16,8	14,5	19,8	15,7	18,2	18,7	5,3	=MAX(B21:G21)-MIN(B21:G21)	17,3	=PROMEDIO(B21:G21)			
22	21	15,9	17,9	18,7	15,7	18,4	17,5	3,0	=MAX(B22:G22)-MIN(B22:G22)	17,4	=PROMEDIO(B22:G22)			
23	22	15,0	18,0	18,2	16,8	14,2	17,8	4,0	=MAX(B23:G23)-MIN(B23:G23)	16,7	=PROMEDIO(B23:G23)			
24	23	16,8	18,9	20,0	16,9	16,2	18,5	3,8	=MAX(B24:G24)-MIN(B24:G24)	17,9	=PROMEDIO(B24:G24)			
25	24	18,9	17,9	17,4	17,5	17,2	16,5	2,4	=MAX(B25:G25)-MIN(B25:G25)	17,6	=PROMEDIO(B25:G25)			
26							$\bar{R}$	3,146	=PROMEDIO(H2:H25) $\bar{\bar{X}}$	16,314	=PROMEDIO(J2:J25)			
27							n	6	=CONTAR(B2:G2)					
28							D_3	0	A_2	0,483				
29							D_4	2,004						
30						Limite superior		6,30425	=H29*H26	Limite Superior		17,83	=J26+J28*H26	
31						Limite inferior		0	=H28*H26	Limite inferior		14,79	=J26-J28*I t	

TAREA DE INTERAPRENDIZAJE Nº 27

1) ¿Qué es variación de causa asignable?

2) Realice un organizador gráfico de las Gráficas para la Media

3) Realice un organizador gráfico de las Gráficas para el Rango

Resuelva el siguiente problema de forma manual y empleando el programa Excel y Graph.

4) Una fábrica produce estructuras para computadores de mesa, los cuales deben cumplir ciertas especificaciones de tamaño. Para garantizar que se cumplan estos estándares de calidad, el gerente de la fábrica, recolecta K= 24 muestras (subgrupos) de tamaño n = 6, y mide su ancho. Los resultados aparecen en la siguiente tabla

Nº Muestra	Medias muestrales					
1	15,2	14,5	15,4	16,5	15,9	16,2
2	16,2	15,4	15,9	15,2	15,2	14,5
3	15,6	16,5	15,9	16,2	15,9	16,2
4	18,5	14,8	15,7	15,2	16,8	14,2
5	17,5	15,7	14,5	14,2	14,5	15,2
6	14,3	15,9	16,5	14,8	15,4	14,8
7	15,4	15,2	15,4	15,8	14,2	15,7
8	18,0	14,5	14,4	16,2	14,8	16,8
9	14,2	15,6	14,5	16,1	15,7	15,9
10	15,7	16,5	14,5	14,8	16,8	16,1
11	14,8	14,5	16,5	14,9	15,8	16,3
12	16,8	15,8	15,2	15,8	15,7	16,2
13	15,2	15,9	14,5	15,1	15,9	14,7
14	15,4	15,7	16,8	15,3	14,8	14,9
15	18,4	15,7	15,9	14,8	15,5	14,8

16	16,5	16,8	15,0	15,7	16,9	14,7
17	15,2	16,9	16,8	17,0	17,1	15,4
18	16,8	17,2	18,9	18,5	18,5	18,9
19	13,5	17,6	18,7	21,1	17,2	16,0
20	19,8	14,5	20,8	19,2	19,2	18,7
21	18,7	17,9	18,7	20,8	18,4	17,5
22	17,5	18,0	18,2	20,2	14,2	17,8
23	14,9	18,9	20,0	16,8	16,2	18,5
24	18,7	17,9	17,4	18,7	17,2	16,5

4.1) Calcular el rango promedio

2,9625

4.2) Calcular el límite superior de control para el rango

5,936

4.3) Calcular el límite inferior de control para el rango

0

4.4) Graficar manualmente la gráfica R.

4.5) ¿El proceso está bajo control?. ¿Por qué?

5) Empleando los conocimientos de las gráfica R, plantee y resuelva un ejercicio de aplicación con $15 \leq k \leq 20$ y $3 \leq n \leq 5$. Resuelva empleando Excel y Graph.

6) Empleando los datos de la fábrica que produce estructuras para computadores de mesa, problema presentado anteriormente.

6.1) Calcular $\bar{\bar{X}}$

16,3194

6.2) Calcular el límite superior de control para las medias

17,75

6.3) Calcular el límite inferior de control para las medias

14,89

6.4) Graficar manualmente la gráfica $\bar{X}$.

6.5) ¿El proceso está bajo control?. ¿Por qué?

6.6) Escriba una semejanza y una diferencia entre la gráfica R del problema Nº 4 y la gráfica $\bar{X}$ del presente problema.

7) Consulte en la biblioteca o en el internet un ejercicio de aplicación de gráfica $\bar{X}$. Presente el ejercicio empleando Excel y Graph.

5.3) GRÁFICAS DE CONTROL PARA ATRIBUTOS

A) Gráfica de control para la proporción de artículos disconformes: la gráfica p

Se utilizan diferentes tipos de gráficas de control para monitorear procesos y para determinar si se encuentra presente en el proceso alguna causa especial de variación. Las gráficas de atributos se utilizan para variables categóricas o discretas. En esta sección se estudiará la gráfica p, que se utiliza cuando los elementos que son muestreados se clasifican de acuerdo a si se conforman o no con los requerimientos definidos operacionalmente. Por lo tanto, la gráfica p nos ayuda a monitorear y analizar la proporción de elementos disconformes que están en las muestras repetidas (es decir, subgrupos) que se seleccionan de un proceso.

Para iniciar la explicación de las gráficas p, recuerde que la proporción de muestra se define como $p = \frac{X}{n}$, y la desviación estándar como $\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$

Los límites de control para la gráfica p son:

$$\bar{p} \pm 3 \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

$$LCS = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

$$LCI = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

Para igual n_i

$$\bar{n} = n_i \; y \; \bar{p} = \frac{\sum p_i}{k}$$

O en general,

$$\bar{n} = \frac{\sum n_i}{k} \; y \; \bar{p} = \frac{\sum X_i}{\sum n_i}$$

Donde:

$X_i = número\ de\ elementos\ disconformes\ en\ el\ subgrupo\ i$

$n_i = tamaño\ de\ muestra\ (o\ subgrupo) para\ el\ subgrupo\ i$

$p_i = \frac{X_i}{n_i} = proporción\ de\ elementos\ disconformes\ en\ el\ subgrupo\ i$

k = número de subgrupos seleccionados

$\bar{n}$ = tamaño promedio del subgrupo

$\bar{p}$ = proporción estimada de elementos disconformes

Cualquier valor negativo para el límite de control inferior significa que el límite de control inferior no existe.

Otros casos de Gráficas p

En la construcción de las gráficas p simplemente se toma nota de la proporción de artículos defectuosos en una muestra. Esta proporción, p es

$$p = \frac{Número\ de\ defectos\ en\ una\ muestra}{Tamaño\ de\ la\ muestra} = \frac{X}{n}$$

Si se toman varias muestras, produciendo varios valores para p. La proporción media de defectos para estas varias muestras, $\bar{p}$ se calcula de la siguiente manera

$$\bar{p} = \frac{Número\ total\ de\ defectos\ en\ todas\ las\ muestras}{Tamaño\ total\ de\ todos\ los\ artículos\ inspeccionados}$$

La desviación estándar para la proporción de defectos es:

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}}$$

La desviación estándar para la proporción de defectos cuando σ es desconocida es:

$$S_p = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

El límite superior de control para las proporciones es:

$$LSC = \bar{p} + 3S_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

El límite inferior de control para las proporciones es:

$$LIC = \bar{p} - 3S_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Ejemplo ilustrativo

Durante la fase de análisis del modelo Seis Sigma DMAIC, se recolectaron los datos de las disconformidades diariamente de una muestra de 200 habitantes de un hotel. La siguiente tabla lista el número y proporción de habitaciones disconformes para cada día durante un periodo de 4 semanas

Día	Habitaciones estudiadas (n)	Habitaciones no preparadas (X)	Proporción (X/n)
1	200	16	0,08
2	200	7	0,035
3	200	21	0,105
4	200	17	0,085
5	200	25	0,125
6	200	19	0,095
7	200	16	0,08
8	200	15	0,075
9	200	11	0,055
10	200	12	0,06
11	200	22	0,11
12	200	20	0,1
13	200	17	0,085
14	200	26	0,13
15	200	18	0,09

16	200	13	0,065
17	200	15	0,075
18	200	10	0,05
19	200	14	0,07
20	200	25	0,125
21	200	19	0,095
22	200	12	0,06
23	200	6	0,03
24	200	12	0,06
25	200	18	0,09
26	200	15	0,075
27	200	20	0,1
28	200	22	0,11
Total	5600	463	2,315

Para estos datos, $k = 28; \sum p_i = 2{,}315,;\ \bar{n} = n = 200;\ \ \bar{p} = \frac{\sum p_i}{k} = \frac{2{,}315}{28} = 0{,}0827$

Remplazando valores en

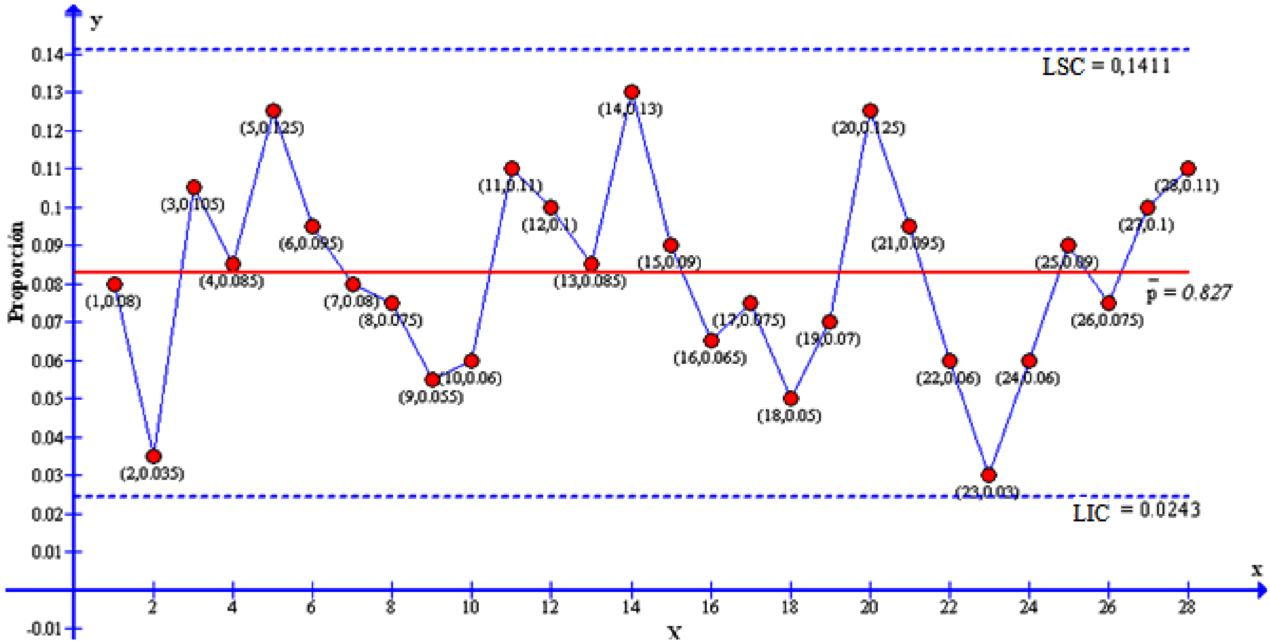
$$\bar{p} \pm 3 \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{\bar{n}}}$$

Se obtiene:

$$0{,}0827 \pm 3 \cdot \sqrt{\frac{0{,}0827(1-0{,}0827)}{200}}$$

Entonces
LSC = 0,0827+0,0584 = 0,1411
LIC = 0,0827-0,0584 = 0,0243

La siguiente figura realizada en el programa Graph representa la gráfica de Control p para cuartos que no están listos



Interpretación: Se observa que la proporción de disconformidades es mayor en el día Nº 14 y menor en el día Nº 23
No hay causas especiales de variación, ya que las proporciones están dentro de los límites de control

TAREA DE INTERAPRENDIZAJE N° 28

Resolver de forma manual, empleando Excel (para los cálculos) y Graph (para las gráficas).

1) Los siguientes datos fueron recolectados de disconformidades durante un periodo de 10 días

Día	Tamaño de la muestra	Disconformidades
1	100	10
2	100	11
3	100	13
4	100	14
5	100	15
6	100	17
7	100	9
8	100	14
9	100	13
10	100	12

1.1) ¿En qué día la proporción de disconformidades es mayor?. ¿Y menor?

Día 6, Día 7

1.2) ¿Cuáles son los límites de control inferior y superior?

0,0278 y 0,2282

1.3) ¿Hay algunas causas especiales de variación?

No, las proporciones están dentro de los límites de control

1.4) Realice la gráfica p en forma manual y con Graph

2) Los siguientes datos fueron recolectados de disconformidades durante un periodo de 10 días

Día	Tamaño de la muestra	Disconformidades
1	111	12
2	93	14
3	105	10
4	92	18
5	117	22
6	88	14
7	117	15
8	87	13
9	119	14
10	107	16

2.1) ¿En qué día la proporción de disconformidades es mayor?. ¿Y menor?

Día 4, Día 3

2.2) ¿Cuáles son los límites de control inferior y superior?

0,0397 y 0,2460

2.3) ¿Hay algunas causas especiales de variación?

No, las proporciones están dentro de los límites de control

2.4) Realice la gráfica p en forma manual y con Graph

3) Una fábrica de instrumentos musicales realiza un procedimiento de control de calidad para detectar los defectos en un modelo de guitarras, para lo cual selecciona 15 muestras de tamaño 40. El número de defectos en cada muestra se indica en la siguiente tabla

Muestra	Número de defectos
1	10
2	12
3	9
4	15
5	27
6	8
7	11
8	11
9	13
10	15
11	17
12	3
13	25
14	18
15	17

3.1) Calcular el número total de defectos en todas las muestras

211

3.2) Calcular el tamaño total de todos los artículos inspeccionados

600

3.3) Calcular $\bar{p}$

0,3517

3.4) Calcular el límite superior de control para proporciones

0,5782

3.5) Calcular el límite inferior de control para proporciones

0,1252

3.6) Elabore la gráfica p en forma manual y con Graph

4) Consulte en la biblioteca o en el internet un ejercicio de aplicación de la gráfica p. Presente el ejercicio empleando Excel y Graph.

B) Las Gráficas c

Estas gráficas están diseñadas para detectar el número de defectos en una sola unidad. Al desarrollar la gráficas p una unidad completa se consideraba defectuosa o no defectuosa. Sin embrago, en muchos casos, la presencia de uno o más defectos puede no producir necesariamente una unidad inaceptable. Un fabricante de muebles puede encontrar defectos menores en un sofá y sin embargo no considerarlo inaceptable. Si los defectos por cada 100 metros cuadrados de tapetes para el piso fueran pocas y menores, el fabricante puede decidir ofrecerlos en venta a pesar de estas imperfecciones. Una gráfica c se utiliza para analizar el número de imperfecciones por unidad de producción.

La gráfica c tiene que ver con el número de imperfecciones (defectos) por unidad (por sofá o por cada 100 metros cuadrados).

Los límites de control se establecen alrededor del número de defectos en la población, c. En el caso probable de que c sea desconocido, se estima mediante $\bar{c}$, el número promedio de defectos en las unidades (número de defectos dividido para el número de muestras)

Una unidad puede constar de un solo artículo como un sofá, o una pieza de tapete de 100 metros cuadrados, o por ejemplo puede contener, un envío de 50 páginas impresas en las cuales se detectaron errores tipográficos. La unidad debe ser consistente en tamaño, número o área. Anteriormente se definió la desviación estándar del número de ocurrencias como la raíz cuadrada del número promedio de defecto. Así:

Desviación estándar para el número de defectos

$$s_{\bar{c}} = \sqrt{\bar{c}}$$

Los límites de control están a tres desviaciones estándar por encima y por debajo de $\bar{c}$

Límite superior de control para el número de defectos

$$LSC_c = \bar{c} + 3s_{\bar{c}}$$

Límite inferior de control para el número de defectos

$$LIC_c = \bar{c} - 3s_{\bar{c}}$$

Nota: Si $LIC_c < 0$, se considera que $LIC_c = 0$

TAREA DE INTERAPRENDIZAJE N° 29

1) Elabore un organizador gráfico sobre las gráficas c

2) ¿Por qué, si $LIC_c < 0$, se considera que $LIC_c = 0$?

Resolver de forma manual, empleando Excel y Graph.

3) Una empresa dedicada a la elaboración de papel para computador inspeccionó 20 hojas de un nuevo tipo de papel para buscar defectos. Los resultados se observan en la siguiente tabla:

Hoja	Número de defectos
1	5
2	4
3	3
4	5
5	16
6	1
7	8
8	9
9	9
10	4
11	3
12	15
13	10
14	8
15	4
16	2
17	10
18	12
19	7
20	17

3.1) Calcular $\bar{c}$

7,6

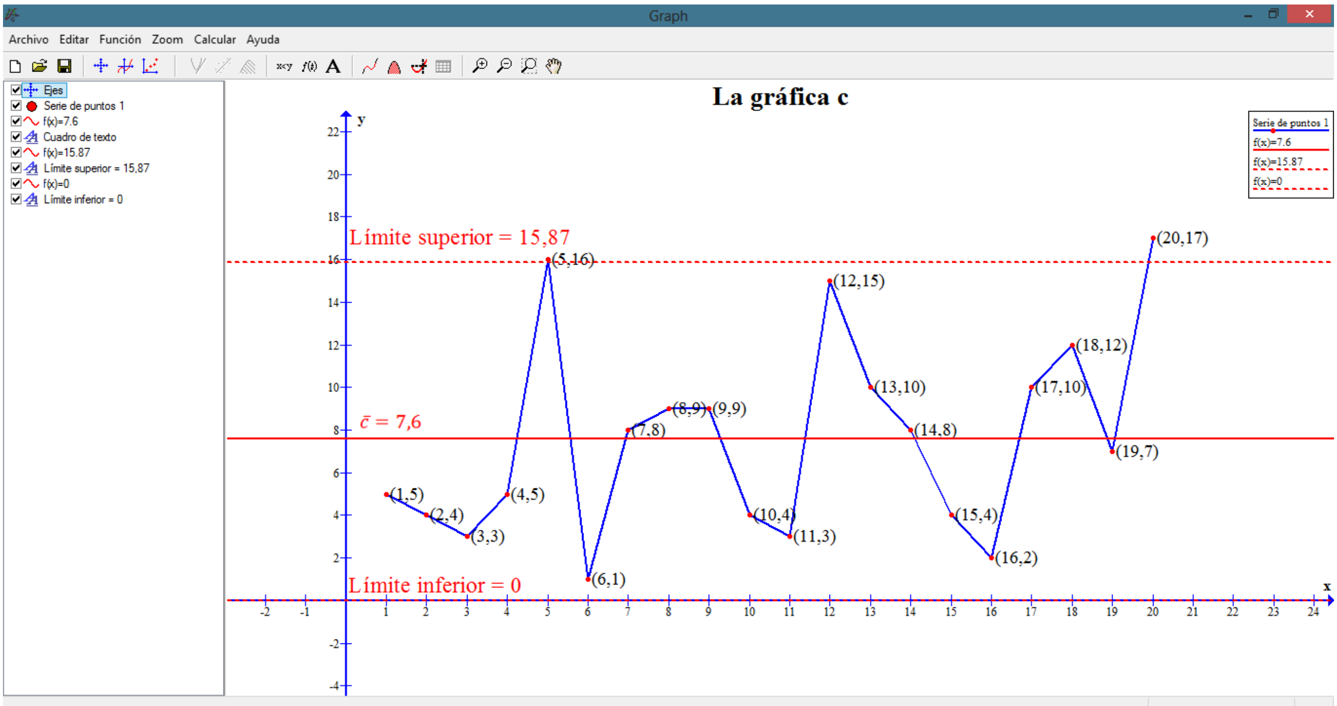
3.2) Calcular el límite superior de control para el número de defectos

15,87

3.3) Calcular el límite inferior de control para el número de defectos

0

3.4) Elabore la gráfica c en forma munual y con Graph



3.5) Realice una interpretación de la gráfica c

4) El jefe de personal de una empresa con 90 empleados ha introducido una estrategia para controlar el número de empleados ausentes al trabajo cada día. Para probar la efectividad del procedimiento, se seleccionan 20 días aleatoriamente y se registran los números de trabajadores ausentes, información que se indica en la siguiente tabla:

Día	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Número de ausencias	6	3	3	5	2	0	5	12	0	0	5	6	5	8	7	5	6	3	5	6

4.1) Calcular $\bar{c}$

1,022

4.2) Calcular el límite superior de control para el número de defectos

4,055

4.3) Calcular el límite inferior de control para el número de defectos

0

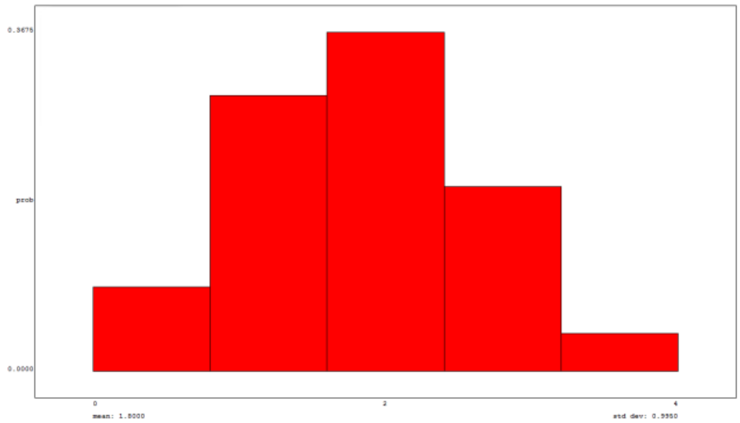
4.4) Elabore manualmente la gráfica c

4.5) Realice una interpretación de la gráfica c

5) Empleando los conocimientos de las gráficas c, plantee y resuelva un ejercicio de aplicación. Resuelva en forma manual, empleando Excel y Graph.

TABLAS ELABORADAS CON EXCEL

TABLA Nº 1
PROBABILIDADES BINOMIALES



$$Ejemplo: Para\ n = 4\ y\ p = 0,45 \ \Rightarrow P(X = 2) = 0,3675$$

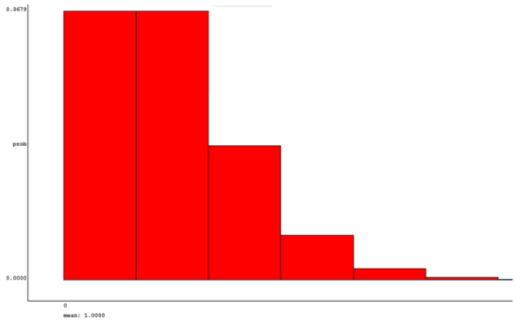
n	X	P									
		0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
1	0	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000	0,6500	0,6000	0,5500	0,5000
1	1	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000	0,3500	0,4000	0,4500	0,5000
n	X	P									
2	0	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900	0,4225	0,3600	0,3025	0,2500
2	1	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200	0,4550	0,4800	0,4950	0,5000
2	2	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900	0,1225	0,1600	0,2025	0,2500
n	X	P									
3	0	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430	0,2746	0,2160	0,1664	0,1250
3	1	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410	0,4436	0,4320	0,4084	0,3750
3	2	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890	0,2389	0,2880	0,3341	0,3750
3	3	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270	0,0429	0,0640	0,0911	0,1250
n	X	P									
4	0	0,8145	0,6561	0,5220	0,4096	0,3164	0,2401	0,1785	0,1296	0,0915	0,0625
4	1	0,1715	0,2916	0,3685	0,4096	0,4219	0,4116	0,3845	0,3456	0,2995	0,2500
4	2	0,0135	0,0486	0,0975	0,1536	0,2109	0,2646	0,3105	0,3456	0,3675	0,3750
4	3	0,0005	0,0036	0,0115	0,0256	0,0469	0,0756	0,1115	0,1536	0,2005	0,2500
4	4	0,0000	0,0001	0,0005	0,0016	0,0039	0,0081	0,0150	0,0256	0,0410	0,0625
n	X	P									
5	0	0,7738	0,5905	0,4437	0,3277	0,2373	0,1681	0,1160	0,0778	0,0503	0,0313
5	1	0,2036	0,3281	0,3915	0,4096	0,3955	0,3602	0,3124	0,2592	0,2059	0,1563
5	2	0,0214	0,0729	0,1382	0,2048	0,2637	0,3087	0,3364	0,3456	0,3369	0,3125
5	3	0,0011	0,0081	0,0244	0,0512	0,0879	0,1323	0,1811	0,2304	0,2757	0,3125
5	4	0,0000	0,0005	0,0022	0,0064	0,0146	0,0284	0,0488	0,0768	0,1128	0,1563
5	5	0,0000	0,0000	0,0001	0,0003	0,0010	0,0024	0,0053	0,0102	0,0185	0,0313
n	x	P									
6	0	0,7351	0,5314	0,3771	0,2621	0,1780	0,1176	0,0754	0,0467	0,0277	0,0156
6	1	0,2321	0,3543	0,3993	0,3932	0,3560	0,3025	0,2437	0,1866	0,1359	0,0938
6	2	0,0305	0,0984	0,1762	0,2458	0,2966	0,3241	0,3280	0,3110	0,2780	0,2344
6	3	0,0021	0,0146	0,0415	0,0819	0,1318	0,1852	0,2355	0,2765	0,3032	0,3125
6	4	0,0001	0,0012	0,0055	0,0154	0,0330	0,0595	0,0951	0,1382	0,1861	0,2344
6	5	0,0000	0,0001	0,0004	0,0015	0,0044	0,0102	0,0205	0,0369	0,0609	0,0938
6	6	0,0000	0,0000	0,0000	0,0001	0,0002	0,0007	0,0018	0,0041	0,0083	0,0156
n	X	P									
7	0	0,6983	0,4783	0,3206	0,2097	0,1335	0,0824	0,0490	0,0280	0,0152	0,0078
7	1	0,2573	0,3720	0,3960	0,3670	0,3115	0,2471	0,1848	0,1306	0,0872	0,0547
7	2	0,0406	0,1240	0,2097	0,2753	0,3115	0,3177	0,2985	0,2613	0,2140	0,1641
7	3	0,0036	0,0230	0,0617	0,1147	0,1730	0,2269	0,2679	0,2903	0,2918	0,2734
7	4	0,0002	0,0026	0,0109	0,0287	0,0577	0,0972	0,1442	0,1935	0,2388	0,2734
7	5	0,0000	0,0002	0,0012	0,0043	0,0115	0,0250	0,0466	0,0774	0,1172	0,1641
7	6	0,0000	0,0000	0,0001	0,0004	0,0013	0,0036	0,0084	0,0172	0,0320	0,0547
7	7	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0006	0,0016	0,0037	0,0078
n	X	P									
8	0	0,6634	0,4305	0,2725	0,1678	0,1001	0,0576	0,0319	0,0168	0,0084	0,0039
8	1	0,2793	0,3826	0,3847	0,3355	0,2670	0,1977	0,1373	0,0896	0,0548	0,0313
8	2	0,0515	0,1488	0,2376	0,2936	0,3115	0,2965	0,2587	0,2090	0,1569	0,1094
8	3	0,0054	0,0331	0,0839	0,1468	0,2076	0,2541	0,2786	0,2787	0,2568	0,2188
8	4	0,0004	0,0046	0,0185	0,0459	0,0865	0,1361	0,1875	0,2322	0,2627	0,2734
8	5	0,0000	0,0004	0,0026	0,0092	0,0231	0,0467	0,0808	0,1239	0,1719	0,2188
8	6	0,0000	0,0000	0,0002	0,0011	0,0038	0,0100	0,0217	0,0413	0,0703	0,1094

8	7	0,0000	0,0000	0,0000	0,0001	0,0004	0,0012	0,0033	0,0079	0,0164	0,0313
8	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0007	0,0017	0,0039
n	X	P									
9	0	0,6302	0,3874	0,2316	0,1342	0,0751	0,0404	0,0207	0,0101	0,0046	0,0020
9	1	0,2985	0,3874	0,3679	0,3020	0,2253	0,1556	0,1004	0,0605	0,0339	0,0176
9	2	0,0629	0,1722	0,2597	0,3020	0,3003	0,2668	0,2162	0,1612	0,1110	0,0703
9	3	0,0077	0,0446	0,1069	0,1762	0,2336	0,2668	0,2716	0,2508	0,2119	0,1641
9	4	0,0006	0,0074	0,0283	0,0661	0,1168	0,1715	0,2194	0,2508	0,2600	0,2461
9	5	0,0000	0,0008	0,0050	0,0165	0,0389	0,0735	0,1181	0,1672	0,2128	0,2461
9	6	0,0000	0,0001	0,0006	0,0028	0,0087	0,0210	0,0424	0,0743	0,1160	0,1641
9	7	0,0000	0,0000	0,0000	0,0003	0,0012	0,0039	0,0098	0,0212	0,0407	0,0703
9	8	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0013	0,0035	0,0083	0,0176
9	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0008	0,0020
n	X	P									
10	0	0,5987	0,3487	0,1969	0,1074	0,0563	0,0282	0,0135	0,0060	0,0025	0,0010
10	1	0,3151	0,3874	0,3474	0,2684	0,1877	0,1211	0,0725	0,0403	0,0207	0,0098
10	2	0,0746	0,1937	0,2759	0,3020	0,2816	0,2335	0,1757	0,1209	0,0763	0,0439
10	3	0,0105	0,0574	0,1298	0,2013	0,2503	0,2668	0,2522	0,2150	0,1665	0,1172
10	4	0,0010	0,0112	0,0401	0,0881	0,1460	0,2001	0,2377	0,2508	0,2384	0,2051
10	5	0,0001	0,0015	0,0085	0,0264	0,0584	0,1029	0,1536	0,2007	0,2340	0,2461
10	6	0,0000	0,0001	0,0012	0,0055	0,0162	0,0368	0,0689	0,1115	0,1596	0,2051
10	7	0,0000	0,0000	0,0001	0,0008	0,0031	0,0090	0,0212	0,0425	0,0746	0,1172
10	8	0,0000	0,0000	0,0000	0,0001	0,0004	0,0014	0,0043	0,0106	0,0229	0,0439
10	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0016	0,0042	0,0098
10	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010
n	X	P									
11	0	0,5688	0,3138	0,1673	0,0859	0,0422	0,0198	0,0088	0,0036	0,0014	0,0005
11	1	0,3293	0,3835	0,3248	0,2362	0,1549	0,0932	0,0518	0,0266	0,0125	0,0054
11	2	0,0867	0,2131	0,2866	0,2953	0,2581	0,1998	0,1395	0,0887	0,0513	0,0269
11	3	0,0137	0,0710	0,1517	0,2215	0,2581	0,2568	0,2254	0,1774	0,1259	0,0806
11	4	0,0014	0,0158	0,0536	0,1107	0,1721	0,2201	0,2428	0,2365	0,2060	0,1611
11	5	0,0001	0,0025	0,0132	0,0388	0,0803	0,1321	0,1830	0,2207	0,2360	0,2256
11	6	0,0000	0,0003	0,0023	0,0097	0,0268	0,0566	0,0985	0,1471	0,1931	0,2256
11	7	0,0000	0,0000	0,0003	0,0017	0,0064	0,0173	0,0379	0,0701	0,1128	0,1611
11	8	0,0000	0,0000	0,0000	0,0002	0,0011	0,0037	0,0102	0,0234	0,0462	0,0806
11	9	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0018	0,0052	0,0126	0,0269
11	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0007	0,0021	0,0054
11	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0005
n	X	P									
12	0	0,5404	0,2824	0,1422	0,0687	0,0317	0,0138	0,0057	0,0022	0,0008	0,0002
12	1	0,3413	0,3766	0,3012	0,2062	0,1267	0,0712	0,0368	0,0174	0,0075	0,0029
12	2	0,0988	0,2301	0,2924	0,2835	0,2323	0,1678	0,1088	0,0639	0,0339	0,0161
12	3	0,0173	0,0852	0,1720	0,2362	0,2581	0,2397	0,1954	0,1419	0,0923	0,0537
12	4	0,0021	0,0213	0,0683	0,1329	0,1936	0,2311	0,2367	0,2128	0,1700	0,1208
12	5	0,0002	0,0038	0,0193	0,0532	0,1032	0,1585	0,2039	0,2270	0,2225	0,1934
12	6	0,0000	0,0005	0,0040	0,0155	0,0401	0,0792	0,1281	0,1766	0,2124	0,2256
12	7	0,0000	0,0000	0,0006	0,0033	0,0115	0,0291	0,0591	0,1009	0,1489	0,1934
12	8	0,0000	0,0000	0,0001	0,0005	0,0024	0,0078	0,0199	0,0420	0,0762	0,1208
12	9	0,0000	0,0000	0,0000	0,0001	0,0004	0,0015	0,0048	0,0125	0,0277	0,0537
12	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0008	0,0025	0,0068	0,0161
12	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010	0,0029
12	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002
n	X	P									
13	0	0,5133	0,2542	0,1209	0,0550	0,0238	0,0097	0,0037	0,0013	0,0004	0,0001
13	1	0,3512	0,3672	0,2774	0,1787	0,1029	0,0540	0,0259	0,0113	0,0045	0,0016
13	2	0,1109	0,2448	0,2937	0,2680	0,2059	0,1388	0,0836	0,0453	0,0220	0,0095
13	3	0,0214	0,0997	0,1900	0,2457	0,2517	0,2181	0,1651	0,1107	0,0660	0,0349
13	4	0,0028	0,0277	0,0838	0,1535	0,2097	0,2337	0,2222	0,1845	0,1350	0,0873
13	5	0,0003	0,0055	0,0266	0,0691	0,1258	0,1803	0,2154	0,2214	0,1989	0,1571
13	6	0,0000	0,0008	0,0063	0,0230	0,0559	0,1030	0,1546	0,1968	0,2169	0,2095
13	7	0,0000	0,0001	0,0011	0,0058	0,0186	0,0442	0,0833	0,1312	0,1775	0,2095
13	8	0,0000	0,0000	0,0001	0,0011	0,0047	0,0142	0,0336	0,0656	0,1089	0,1571
13	9	0,0000	0,0000	0,0000	0,0001	0,0009	0,0034	0,0101	0,0243	0,0495	0,0873
13	10	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0022	0,0065	0,0162	0,0349
13	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0012	0,0036	0,0095
13	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0016
13	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
n	X	P									
14	0	0,4877	0,2288	0,1028	0,0440	0,0178	0,0068	0,0024	0,0008	0,0002	0,0001
14	1	0,3593	0,3559	0,2539	0,1539	0,0832	0,0407	0,0181	0,0073	0,0027	0,0009
14	2	0,1229	0,2570	0,2912	0,2501	0,1802	0,1134	0,0634	0,0317	0,0141	0,0056
14	3	0,0259	0,1142	0,2056	0,2501	0,2402	0,1943	0,1366	0,0845	0,0462	0,0222
14	4	0,0037	0,0349	0,0998	0,1720	0,2202	0,2290	0,2022	0,1549	0,1040	0,0611
14	5	0,0004	0,0078	0,0352	0,0860	0,1468	0,1963	0,2178	0,2066	0,1701	0,1222

14	6	0,0000	0,0013	0,0093	0,0322	0,0734	0,1262	0,1759	0,2066	0,2088	0,1833
14	7	0,0000	0,0002	0,0019	0,0092	0,0280	0,0618	0,1082	0,1574	0,1952	0,2095
14	8	0,0000	0,0000	0,0003	0,0020	0,0082	0,0232	0,0510	0,0918	0,1398	0,1833
14	9	0,0000	0,0000	0,0000	0,0003	0,0018	0,0066	0,0183	0,0408	0,0762	0,1222
14	10	0,0000	0,0000	0,0000	0,0000	0,0003	0,0014	0,0049	0,0136	0,0312	0,0611
14	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0010	0,0033	0,0093	0,0222
14	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0019	0,0056
14	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0009
14	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
n	**X**	**P**									
15	0	0,4633	0,2059	0,0874	0,0352	0,0134	0,0047	0,0016	0,0005	0,0001	0,0000
15	1	0,3658	0,3432	0,2312	0,1319	0,0668	0,0305	0,0126	0,0047	0,0016	0,0005
15	2	0,1348	0,2669	0,2856	0,2309	0,1559	0,0916	0,0476	0,0219	0,0090	0,0032
15	3	0,0307	0,1285	0,2184	0,2501	0,2252	0,1700	0,1110	0,0634	0,0318	0,0139
15	4	0,0049	0,0428	0,1156	0,1876	0,2252	0,2186	0,1792	0,1268	0,0780	0,0417
15	5	0,0006	0,0105	0,0449	0,1032	0,1651	0,2061	0,2123	0,1859	0,1404	0,0916
15	6	0,0000	0,0019	0,0132	0,0430	0,0917	0,1472	0,1906	0,2066	0,1914	0,1527
15	7	0,0000	0,0003	0,0030	0,0138	0,0393	0,0811	0,1319	0,1771	0,2013	0,1964
15	8	0,0000	0,0000	0,0005	0,0035	0,0131	0,0348	0,0710	0,1181	0,1647	0,1964
15	9	0,0000	0,0000	0,0001	0,0007	0,0034	0,0116	0,0298	0,0612	0,1048	0,1527
15	10	0,0000	0,0000	0,0000	0,0001	0,0007	0,0030	0,0096	0,0245	0,0515	0,0916
15	11	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0074	0,0191	0,0417
15	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0016	0,0052	0,0139
15	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010	0,0032
15	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005
15	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
n	**X**	**P**									
16	0	0,4401	0,1853	0,0743	0,0281	0,0100	0,0033	0,0010	0,0003	0,0001	0,0002
16	1	0,3706	0,3294	0,2097	0,1126	0,0535	0,0228	0,0087	0,0030	0,0009	0,0018
16	2	0,1463	0,2745	0,2775	0,2111	0,1336	0,0732	0,0353	0,0150	0,0056	0,0085
16	3	0,0359	0,1423	0,2285	0,2463	0,2079	0,1465	0,0888	0,0468	0,0215	0,0278
16	4	0,0061	0,0514	0,1311	0,2001	0,2252	0,2040	0,1553	0,1014	0,0572	0,0667
16	5	0,0008	0,0137	0,0555	0,1201	0,1802	0,2099	0,2008	0,1623	0,1123	0,1222
16	6	0,0001	0,0028	0,0180	0,0550	0,1101	0,1649	0,1982	0,1983	0,1684	0,1746
16	7	0,0000	0,0004	0,0045	0,0197	0,0524	0,1010	0,1524	0,1889	0,1969	0,1964
16	8	0,0000	0,0001	0,0009	0,0055	0,0197	0,0487	0,0923	0,1417	0,1812	0,1746
16	9	0,0000	0,0000	0,0001	0,0012	0,0058	0,0185	0,0442	0,0840	0,1318	0,1222
16	10	0,0000	0,0000	0,0000	0,0002	0,0014	0,0056	0,0167	0,0392	0,0755	0,0667
16	11	0,0000	0,0000	0,0000	0,0000	0,0002	0,0013	0,0049	0,0142	0,0337	0,0278
16	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0040	0,0115	0,0085
16	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0008	0,0029	0,0018
16	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0002
16	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000
16	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
n	**X**	**P**									
17	0	0,4181	0,1668	0,0631	0,0225	0,0075	0,0023	0,0007	0,0002	0,0000	0,0010
17	1	0,3741	0,3150	0,1893	0,0957	0,0426	0,0169	0,0060	0,0019	0,0005	0,0052
17	2	0,1575	0,2800	0,2673	0,1914	0,1136	0,0581	0,0260	0,0102	0,0035	0,0182
17	3	0,0415	0,1556	0,2359	0,2393	0,1893	0,1245	0,0701	0,0341	0,0144	0,0472
17	4	0,0076	0,0605	0,1457	0,2093	0,2209	0,1868	0,1320	0,0796	0,0411	0,0944
17	5	0,0010	0,0175	0,0668	0,1361	0,1914	0,2081	0,1849	0,1379	0,0875	0,1484
17	6	0,0001	0,0039	0,0236	0,0680	0,1276	0,1784	0,1991	0,1839	0,1432	0,1855
17	7	0,0000	0,0007	0,0065	0,0267	0,0668	0,1201	0,1685	0,1927	0,1841	0,1855
17	8	0,0000	0,0001	0,0014	0,0084	0,0279	0,0644	0,1134	0,1606	0,1883	0,1484
17	9	0,0000	0,0000	0,0003	0,0021	0,0093	0,0276	0,0611	0,1070	0,1540	0,0944
17	10	0,0000	0,0000	0,0000	0,0004	0,0025	0,0095	0,0263	0,0571	0,1008	0,0472
17	11	0,0000	0,0000	0,0000	0,0001	0,0005	0,0026	0,0090	0,0242	0,0525	0,0182
17	12	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0081	0,0215	0,0052
17	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0021	0,0068	0,0010
17	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0016	0,0001
17	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0000
17	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
17	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
n	**X**	**P**									
18	0	0,3972	0,1501	0,0536	0,0180	0,0056	0,0016	0,0004	0,0001	0,0000	0,0031
18	1	0,3763	0,3002	0,1704	0,0811	0,0338	0,0126	0,0042	0,0012	0,0003	0,0117
18	2	0,1683	0,2835	0,2556	0,1723	0,0958	0,0458	0,0190	0,0069	0,0022	0,0327
18	3	0,0473	0,1680	0,2406	0,2297	0,1704	0,1046	0,0547	0,0246	0,0095	0,0708
18	4	0,0093	0,0700	0,1592	0,2153	0,2130	0,1681	0,1104	0,0614	0,0291	0,1214
18	5	0,0014	0,0218	0,0787	0,1507	0,1988	0,2017	0,1664	0,1146	0,0666	0,1669
18	6	0,0002	0,0052	0,0301	0,0816	0,1436	0,1873	0,1941	0,1655	0,1181	0,1855
18	7	0,0000	0,0010	0,0091	0,0350	0,0820	0,1376	0,1792	0,1892	0,1657	0,1669
18	8	0,0000	0,0002	0,0022	0,0120	0,0376	0,0811	0,1327	0,1734	0,1864	0,1214
18	9	0,0000	0,0000	0,0004	0,0033	0,0139	0,0386	0,0794	0,1284	0,1694	0,0708

18	10	0,0000	0,0000	0,0001	0,0008	0,0042	0,0149	0,0385	0,0771	0,1248	0,0327
18	11	0,0000	0,0000	0,0000	0,0001	0,0010	0,0046	0,0151	0,0374	0,0742	0,0117
18	12	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0047	0,0145	0,0354	0,0031
18	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0045	0,0134	0,0006
18	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0039	0,0001
18	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0000
18	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000
18	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
18	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006
n	**X**	**P**									
19	0	0,3774	0,1351	0,0456	0,0144	0,0042	0,0011	0,0003	0,0001	0,0000	0,0074
19	1	0,3774	0,2852	0,1529	0,0685	0,0268	0,0093	0,0029	0,0008	0,0002	0,0222
19	2	0,1787	0,2852	0,2428	0,1540	0,0803	0,0358	0,0138	0,0046	0,0013	0,0518
19	3	0,0533	0,1796	0,2428	0,2182	0,1517	0,0869	0,0422	0,0175	0,0062	0,0961
19	4	0,0112	0,0798	0,1714	0,2182	0,2023	0,1491	0,0909	0,0467	0,0203	0,1442
19	5	0,0018	0,0266	0,0907	0,1636	0,2023	0,1916	0,1468	0,0933	0,0497	0,1762
19	6	0,0002	0,0069	0,0374	0,0955	0,1574	0,1916	0,1844	0,1451	0,0949	0,1762
19	7	0,0000	0,0014	0,0122	0,0443	0,0974	0,1525	0,1844	0,1797	0,1443	0,1442
19	8	0,0000	0,0002	0,0032	0,0166	0,0487	0,0981	0,1489	0,1797	0,1771	0,0961
19	9	0,0000	0,0000	0,0007	0,0051	0,0198	0,0514	0,0980	0,1464	0,1771	0,0518
19	10	0,0000	0,0000	0,0001	0,0013	0,0066	0,0220	0,0528	0,0976	0,1449	0,0222
19	11	0,0000	0,0000	0,0000	0,0003	0,0018	0,0077	0,0233	0,0532	0,0970	0,0074
19	12	0,0000	0,0000	0,0000	0,0000	0,0004	0,0022	0,0083	0,0237	0,0529	0,0018
19	13	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0024	0,0085	0,0233	0,0003
19	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0082	0,0000
19	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0022	0,0000
19	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0000
19	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0000
19	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003
19	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0018
n	**X**	**P**									
20	0	0,3585	0,1216	0,0388	0,0115	0,0032	0,0008	0,0002	0,0000	0,0000	0,0148
20	1	0,3774	0,2702	0,1368	0,0576	0,0211	0,0068	0,0020	0,0005	0,0001	0,0370
20	2	0,1887	0,2852	0,2293	0,1369	0,0669	0,0278	0,0100	0,0031	0,0008	0,0739
20	3	0,0596	0,1901	0,2428	0,2054	0,1339	0,0716	0,0323	0,0123	0,0040	0,1201
20	4	0,0133	0,0898	0,1821	0,2182	0,1897	0,1304	0,0738	0,0350	0,0139	0,1602
20	5	0,0022	0,0319	0,1028	0,1746	0,2023	0,1789	0,1272	0,0746	0,0365	0,1762
20	6	0,0003	0,0089	0,0454	0,1091	0,1686	0,1916	0,1712	0,1244	0,0746	0,1602
20	7	0,0000	0,0020	0,0160	0,0545	0,1124	0,1643	0,1844	0,1659	0,1221	0,1201
20	8	0,0000	0,0004	0,0046	0,0222	0,0609	0,1144	0,1614	0,1797	0,1623	0,0739
20	9	0,0000	0,0001	0,0011	0,0074	0,0271	0,0654	0,1158	0,1597	0,1771	0,0370
20	10	0,0000	0,0000	0,0002	0,0020	0,0099	0,0308	0,0686	0,1171	0,1593	0,0148
20	11	0,0000	0,0000	0,0000	0,0005	0,0030	0,0120	0,0336	0,0710	0,1185	0,0046
20	12	0,0000	0,0000	0,0000	0,0001	0,0008	0,0039	0,0136	0,0355	0,0727	0,0011
20	13	0,0000	0,0000	0,0000	0,0000	0,0002	0,0010	0,0045	0,0146	0,0366	0,0002
20	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0049	0,0150	0,0000
20	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0013	0,0049	0,0000
20	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0013	0,0000
20	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0000
20	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002
20	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0011
20	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0046

TABLA Nº 2
PROBABILIDADES DE POISSON



$Ejemplo: Para\ \lambda = 1\ y\ X = 0\ \Rightarrow P(X = 0) = 0,3679$

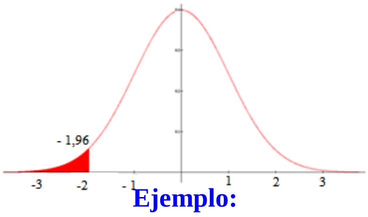
	λ									
X	**0,005**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0	0,9950	0,9900	0,9802	0,9704	0,9608	0,9512	0,9418	0,9324	0,9231	0,9139
1	0,0050	0,0099	0,0196	0,0291	0,0384	0,0476	0,0565	0,0653	0,0738	0,0823
2	0,0000	0,0000	0,0002	0,0004	0,0008	0,0012	0,0017	0,0023	0,0030	0,0037
3	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001
	λ									
X	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**	**1**
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,6381	0,3659	0,3679
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,5283	0,1647	0,1839
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,5896	0,0494	0,0613
4	0,0000	0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,5546	0,0111	0,0153
5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0004	0,0007	0,5743	0,0020	0,0031
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,5631	0,0003	0,0005
7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,5695	0,0000	0,0001
	λ									
X	**1,1**	**1,2**	**1,3**	**1,4**	**1,5**	**1,6**	**1,7**	**1,8**	**1,9**	**2**
0	0,3329	0,3012	0,2725	0,2466	0,2231	0,2019	0,1827	0,1653	0,1496	0,1353
1	0,3662	0,3614	0,3543	0,3452	0,3347	0,3230	0,3106	0,2975	0,2842	0,2707
2	0,2014	0,2169	0,2303	0,2417	0,2510	0,2584	0,2640	0,2678	0,2700	0,2707
3	0,0738	0,0867	0,0998	0,1128	0,1255	0,1378	0,1496	0,1607	0,1710	0,1804
4	0,0203	0,0260	0,0324	0,0395	0,0471	0,0551	0,0636	0,0723	0,0812	0,0902
5	0,0045	0,0062	0,0084	0,0111	0,0141	0,0176	0,0216	0,0260	0,0309	0,0361
6	0,0008	0,0012	0,0018	0,0026	0,0035	0,0047	0,0061	0,0078	0,0098	0,0120
7	0,0001	0,0002	0,0003	0,0005	0,0008	0,0011	0,0015	0,0020	0,0027	0,0034
8	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0003	0,0005	0,0006	0,0009
9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002
	λ									
X	**2,1**	**2,2**	**2,3**	**2,4**	**2,5**	**2,6**	**2,7**	**2,8**	**2,9**	**3**
0	0,1225	0,1108	0,1003	0,0907	0,0821	0,0743	0,0672	0,0608	0,0550	0,0498
1	0,2572	0,2438	0,2306	0,2177	0,2052	0,1931	0,1815	0,1703	0,1596	0,1494
2	0,2700	0,2681	0,2652	0,2613	0,2565	0,2510	0,2450	0,2384	0,2314	0,2240
3	0,1890	0,1966	0,2033	0,2090	0,2138	0,2176	0,2205	0,2225	0,2237	0,2240
4	0,0992	0,1082	0,1169	0,1254	0,1336	0,1414	0,1488	0,1557	0,1622	0,1680
5	0,0417	0,0476	0,0538	0,0602	0,0668	0,0735	0,0804	0,0872	0,0940	0,1008
6	0,0146	0,0174	0,0206	0,0241	0,0278	0,0319	0,0362	0,0407	0,0455	0,0504
7	0,0044	0,0055	0,0068	0,0083	0,0099	0,0118	0,0139	0,0163	0,0188	0,0216
8	0,0011	0,0015	0,0019	0,0025	0,0031	0,0038	0,0047	0,0057	0,0068	0,0081
9	0,0003	0,0004	0,0005	0,0007	0,0009	0,0011	0,0014	0,0018	0,0022	0,0027
10	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0004	0,0005	0,0006	0,0008
11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0002
12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	λ									
X	**3,1**	**3,2**	**3,3**	**3,4**	**3,5**	**3,6**	**3,7**	**3,8**	**3,9**	**4,0**
0	0,0450	0,0408	0,0369	0,0334	0,0302	0,0273	0,0247	0,0224	0,0202	0,0183
1	0,1397	0,1304	0,1217	0,1135	0,1057	0,0984	0,0915	0,0850	0,0789	0,0733
2	0,2165	0,2087	0,2008	0,1929	0,1850	0,1771	0,1692	0,1615	0,1539	0,1465
3	0,2237	0,2226	0,2209	0,2186	0,2158	0,2125	0,2087	0,2046	0,2001	0,1954
4	0,1733	0,1781	0,1823	0,1858	0,1888	0,1912	0,1931	0,1944	0,1951	0,1954
5	0,1075	0,1140	0,1203	0,1264	0,1322	0,1377	0,1429	0,1477	0,1522	0,1563
6	0,0555	0,0608	0,0662	0,0716	0,0771	0,0826	0,0881	0,0936	0,0989	0,1042
7	0,0246	0,0278	0,0312	0,0348	0,0385	0,0425	0,0466	0,0508	0,0551	0,0595
8	0,0095	0,0111	0,0129	0,0148	0,0169	0,0191	0,0215	0,0241	0,0269	0,0298
9	0,0033	0,0040	0,0047	0,0056	0,0066	0,0076	0,0089	0,0102	0,0116	0,0132
10	0,0010	0,0013	0,0016	0,0019	0,0023	0,0028	0,0033	0,0039	0,0045	0,0053
11	0,0003	0,0004	0,0005	0,0006	0,0007	0,0009	0,0011	0,0013	0,0016	0,0019
12	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005	0,0006

13	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002
14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	λ									
X	4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	4,9	5,0
0	0,0166	0,0150	0,0136	0,0123	0,0111	0,0101	0,0091	0,0082	0,0074	0,0067
1	0,0679	0,0630	0,0583	0,0540	0,0500	0,0462	0,0427	0,0395	0,0365	0,0337
2	0,1393	0,1323	0,1254	0,1188	0,1125	0,1063	0,1005	0,0948	0,0894	0,0842
3	0,1904	0,1852	0,1798	0,1743	0,1687	0,1631	0,1574	0,1517	0,1460	0,1404
4	0,1951	0,1944	0,1933	0,1917	0,1898	0,1875	0,1849	0,1820	0,1789	0,1755
5	0,1600	0,1633	0,1662	0,1687	0,1708	0,1725	0,1738	0,1747	0,1753	0,1755
6	0,1093	0,1143	0,1191	0,1237	0,1281	0,1323	0,1362	0,1398	0,1432	0,1462
7	0,0640	0,0686	0,0732	0,0778	0,0824	0,0869	0,0914	0,0959	0,1002	0,1044
8	0,0328	0,0360	0,0393	0,0428	0,0463	0,0500	0,0537	0,0575	0,0614	0,0653
9	0,0150	0,0168	0,0188	0,0209	0,0232	0,0255	0,0281	0,0307	0,0334	0,0363
10	0,0061	0,0071	0,0081	0,0092	0,0104	0,0118	0,0132	0,0147	0,0164	0,0181
11	0,0023	0,0027	0,0032	0,0037	0,0043	0,0049	0,0056	0,0064	0,0073	0,0082
12	0,0008	0,0009	0,0011	0,0013	0,0016	0,0019	0,0022	0,0026	0,0030	0,0034
13	0,0002	0,0003	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009	0,0011	0,0013
14	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005
15	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002
	λ									
X	5,1	5,2	5,3	5,4	5,5	5,6	5,7	5,8	5,9	6,0
0	0,0061	0,0055	0,0050	0,0045	0,0041	0,0037	0,0033	0,0030	0,0027	0,0025
1	0,0311	0,0287	0,0265	0,0244	0,0225	0,0207	0,0191	0,0176	0,0162	0,0149
2	0,0793	0,0746	0,0701	0,0659	0,0618	0,0580	0,0544	0,0509	0,0477	0,0446
3	0,1348	0,1293	0,1239	0,1185	0,1133	0,1082	0,1033	0,0985	0,0938	0,0892
4	0,1719	0,1681	0,1641	0,1600	0,1558	0,1515	0,1472	0,1428	0,1383	0,1339
5	0,1753	0,1748	0,1740	0,1728	0,1714	0,1697	0,1678	0,1656	0,1632	0,1606
6	0,1490	0,1515	0,1537	0,1555	0,1571	0,1584	0,1594	0,1601	0,1605	0,1606
7	0,1086	0,1125	0,1163	0,1200	0,1234	0,1267	0,1298	0,1326	0,1353	0,1377
8	0,0692	0,0731	0,0771	0,0810	0,0849	0,0887	0,0925	0,0962	0,0998	0,1033
9	0,0392	0,0423	0,0454	0,0486	0,0519	0,0552	0,0586	0,0620	0,0654	0,0688
10	0,0200	0,0220	0,0241	0,0262	0,0285	0,0309	0,0334	0,0359	0,0386	0,0413
11	0,0093	0,0104	0,0116	0,0129	0,0143	0,0157	0,0173	0,0190	0,0207	0,0225
12	0,0039	0,0045	0,0051	0,0058	0,0065	0,0073	0,0082	0,0092	0,0102	0,0113
13	0,0015	0,0018	0,0021	0,0024	0,0028	0,0032	0,0036	0,0041	0,0046	0,0052
14	0,0006	0,0007	0,0008	0,0009	0,0011	0,0013	0,0015	0,0017	0,0019	0,0022
15	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009
16	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003
17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001
	λ									
X	6,1	6,2	6,3	6,4	6,5	6,6	6,7	6,8	6,9	7,0
0	0,0022	0,0020	0,0018	0,0017	0,0015	0,0014	0,0012	0,0011	0,0010	0,0009
1	0,0137	0,0126	0,0116	0,0106	0,0098	0,0090	0,0082	0,0076	0,0070	0,0064
2	0,0417	0,0390	0,0364	0,0340	0,0318	0,0296	0,0276	0,0258	0,0240	0,0223
3	0,0848	0,0806	0,0765	0,0726	0,0688	0,0652	0,0617	0,0584	0,0552	0,0521
4	0,1294	0,1249	0,1205	0,1162	0,1118	0,1076	0,1034	0,0992	0,0952	0,0912
5	0,1579	0,1549	0,1519	0,1487	0,1454	0,1420	0,1385	0,1349	0,1314	0,1277
6	0,1605	0,1601	0,1595	0,1586	0,1575	0,1562	0,1546	0,1529	0,1511	0,1490
7	0,1399	0,1418	0,1435	0,1450	0,1462	0,1472	0,1480	0,1486	0,1489	0,1490
8	0,1066	0,1099	0,1130	0,1160	0,1188	0,1215	0,1240	0,1263	0,1284	0,1304
9	0,0723	0,0757	0,0791	0,0825	0,0858	0,0891	0,0923	0,0954	0,0985	0,1014
10	0,0441	0,0469	0,0498	0,0528	0,0558	0,0588	0,0618	0,0649	0,0679	0,0710
11	0,0244	0,0265	0,0285	0,0307	0,0330	0,0353	0,0377	0,0401	0,0426	0,0452
12	0,0124	0,0137	0,0150	0,0164	0,0179	0,0194	0,0210	0,0227	0,0245	0,0263
13	0,0058	0,0065	0,0073	0,0081	0,0089	0,0099	0,0108	0,0119	0,0130	0,0142
14	0,0025	0,0029	0,0033	0,0037	0,0041	0,0046	0,0052	0,0058	0,0064	0,0071
15	0,0010	0,0012	0,0014	0,0016	0,0018	0,0020	0,0023	0,0026	0,0029	0,0033
16	0,0004	0,0005	0,0005	0,0006	0,0007	0,0008	0,0010	0,0011	0,0013	0,0014
17	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006
18	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002
19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001
	λ									
X	7,1	7,2	7,3	7,4	7,5	7,6	7,7	7,8	7,9	8,0
0	0,0008	0,0007	0,0007	0,0006	0,0006	0,0005	0,0005	0,0004	0,0004	0,0003
1	0,0059	0,0054	0,0049	0,0045	0,0041	0,0038	0,0035	0,0032	0,0029	0,0027
2	0,0208	0,0194	0,0180	0,0167	0,0156	0,0145	0,0134	0,0125	0,0116	0,0107
3	0,0492	0,0464	0,0438	0,0413	0,0389	0,0366	0,0345	0,0324	0,0305	0,0286
4	0,0874	0,0836	0,0799	0,0764	0,0729	0,0696	0,0663	0,0632	0,0602	0,0573
5	0,1241	0,1204	0,1167	0,1130	0,1094	0,1057	0,1021	0,0986	0,0951	0,0916
6	0,1468	0,1445	0,1420	0,1394	0,1367	0,1339	0,1311	0,1282	0,1252	0,1221
7	0,1489	0,1486	0,1481	0,1474	0,1465	0,1454	0,1442	0,1428	0,1413	0,1396
8	0,1321	0,1337	0,1351	0,1363	0,1373	0,1381	0,1388	0,1392	0,1395	0,1396
9	0,1042	0,1070	0,1096	0,1121	0,1144	0,1167	0,1187	0,1207	0,1224	0,1241

10	0,0740	0,0770	0,0800	0,0829	0,0858	0,0887	0,0914	0,0941	0,0967	0,0993
11	0,0478	0,0504	0,0531	0,0558	0,0585	0,0613	0,0640	0,0667	0,0695	0,0722
12	0,0283	0,0303	0,0323	0,0344	0,0366	0,0388	0,0411	0,0434	0,0457	0,0481
13	0,0154	0,0168	0,0181	0,0196	0,0211	0,0227	0,0243	0,0260	0,0278	0,0296
14	0,0078	0,0086	0,0095	0,0104	0,0113	0,0123	0,0134	0,0145	0,0157	0,0169
15	0,0037	0,0041	0,0046	0,0051	0,0057	0,0062	0,0069	0,0075	0,0083	0,0090
16	0,0016	0,0019	0,0021	0,0024	0,0026	0,0030	0,0033	0,0037	0,0041	0,0045
17	0,0007	0,0008	0,0009	0,0010	0,0012	0,0013	0,0015	0,0017	0,0019	0,0021
18	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006	0,0006	0,0007	0,0008	0,0009
19	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0003	0,0004
20	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002
21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001
	λ									
X	8,1	8,2	8,3	8,4	8,5	8,6	8,7	8,8	8,9	9,0
0	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001
1	0,0025	0,0023	0,0021	0,0002	0,0017	0,0016	0,0014	0,0013	0,0012	0,0011
2	0,0100	0,0092	0,0086	0,0000	0,0074	0,0068	0,0063	0,0058	0,0054	0,0050
3	0,0269	0,0252	0,0237	0,0000	0,0208	0,0195	0,0183	0,0171	0,0160	0,0150
4	0,0544	0,0517	0,0491	0,0000	0,0443	0,0420	0,0398	0,0377	0,0357	0,0337
5	0,0882	0,0849	0,0816	0,0000	0,0752	0,0722	0,0692	0,0663	0,0635	0,0607
6	0,1191	0,1160	0,1128	0,0000	0,1066	0,1034	0,1003	0,0972	0,0941	0,0911
7	0,1378	0,1358	0,1338	0,0000	0,1294	0,1271	0,1247	0,1222	0,1197	0,1171
8	0,1395	0,1392	0,1388	0,0000	0,1375	0,1366	0,1356	0,1344	0,1332	0,1318
9	0,1256	0,1269	0,1280	0,0000	0,1299	0,1306	0,1311	0,1315	0,1317	0,1318
10	0,1017	0,1040	0,1063	0,0000	0,1104	0,1123	0,1140	0,1157	0,1172	0,1186
11	0,0749	0,0776	0,0802	0,0000	0,0853	0,0878	0,0902	0,0925	0,0948	0,0970
12	0,0505	0,0530	0,0555	0,0000	0,0604	0,0629	0,0654	0,0679	0,0703	0,0728
13	0,0315	0,0334	0,0354	0,0000	0,0395	0,0416	0,0438	0,0459	0,0481	0,0504
14	0,0182	0,0196	0,0210	0,0000	0,0240	0,0256	0,0272	0,0289	0,0306	0,0324
15	0,0098	0,0107	0,0116	0,0000	0,0136	0,0147	0,0158	0,0169	0,0182	0,0194
16	0,0050	0,0055	0,0060	0,0000	0,0072	0,0079	0,0086	0,0093	0,0101	0,0109
17	0,0024	0,0026	0,0029	0,0000	0,0036	0,0040	0,0044	0,0048	0,0053	0,0058
18	0,0011	0,0012	0,0014	0,0000	0,0017	0,0019	0,0021	0,0024	0,0026	0,0029
19	0,0005	0,0005	0,0006	0,0000	0,0008	0,0009	0,0010	0,0011	0,0012	0,0014
20	0,0002	0,0002	0,0002	0,0000	0,0003	0,0004	0,0004	0,0005	0,0005	0,0006
21	0,0001	0,0001	0,0001	0,0000	0,0001	0,0002	0,0002	0,0002	0,0002	0,0003
22	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	λ									
X	9,1	9,2	9,3	9,4	9,5	9,6	9,7	9,8	9,9	10
0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0000
1	0,0010	0,0009	0,0009	0,0008	0,0007	0,0007	0,0006	0,0005	0,0005	0,0005
2	0,0046	0,0043	0,0040	0,0037	0,0034	0,0031	0,0029	0,0027	0,0025	0,0023
3	0,0140	0,0131	0,0123	0,0115	0,0107	0,0100	0,0093	0,0087	0,0081	0,0076
4	0,0319	0,0302	0,0285	0,0269	0,0254	0,0240	0,0226	0,0213	0,0201	0,0189
5	0,0581	0,0555	0,0530	0,0506	0,0483	0,0460	0,0439	0,0418	0,0398	0,0378
6	0,0881	0,0851	0,0822	0,0793	0,0764	0,0736	0,0709	0,0682	0,0656	0,0631
7	0,1145	0,1118	0,1091	0,1064	0,1037	0,1010	0,0982	0,0955	0,0928	0,0901
8	0,1302	0,1286	0,1269	0,1251	0,1232	0,1212	0,1191	0,1170	0,1148	0,1126
9	0,1317	0,1315	0,1311	0,1306	0,1300	0,1293	0,1284	0,1274	0,1263	0,1251
10	0,1198	0,1210	0,1219	0,1228	0,1235	0,1241	0,1245	0,1249	0,1250	0,1251
11	0,0991	0,1012	0,1031	0,1049	0,1067	0,1083	0,1098	0,1112	0,1125	0,1137
12	0,0752	0,0776	0,0799	0,0822	0,0844	0,0866	0,0888	0,0908	0,0928	0,0948
13	0,0526	0,0549	0,0572	0,0594	0,0617	0,0640	0,0662	0,0685	0,0707	0,0729
14	0,0342	0,0361	0,0380	0,0399	0,0419	0,0439	0,0459	0,0479	0,0500	0,0521
15	0,0208	0,0221	0,0235	0,0250	0,0265	0,0281	0,0297	0,0313	0,0330	0,0347
16	0,0118	0,0127	0,0137	0,0147	0,0157	0,0168	0,0180	0,0192	0,0204	0,0217
17	0,0063	0,0069	0,0075	0,0081	0,0088	0,0095	0,0103	0,0111	0,0119	0,0128
18	0,0032	0,0035	0,0039	0,0042	0,0046	0,0051	0,0055	0,0060	0,0065	0,0071
19	0,0015	0,0017	0,0019	0,0021	0,0023	0,0026	0,0028	0,0031	0,0034	0,0037
20	0,0007	0,0008	0,0009	0,0010	0,0011	0,0012	0,0014	0,0015	0,0017	0,0019
21	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006	0,0006	0,0007	0,0008	0,0009
22	0,0001	0,0001	0,0002	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004
23	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002
24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001
25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
26	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
27	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
28	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

TABLA Nº 3
DISTRIBUCIÓN NORMAL

$$Z = \frac{X - \mu}{\sigma}$$

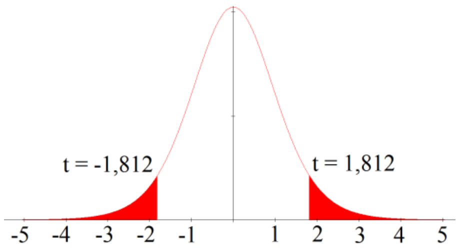


Ejemplo:

$P(Z \leq -1,96) = 0,0250$

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

TABLA Nº 4
DISTRIBUCIÓN t DE STUDENT



Ejemplos:

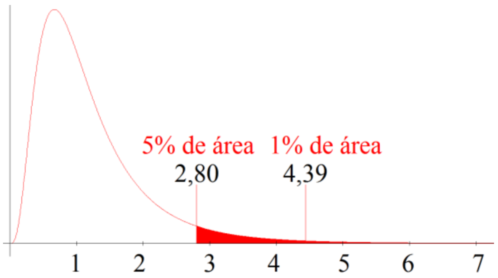
Para n-1 = 10 grados de libertad

$\boldsymbol{P(t \geq 1,812) = 0,05}$

$\boldsymbol{P(t \leq -1,812) = 0,05}$

α / n-1	0,25	0,2	0,15	0,1	0,05	0,025	0,01	0,005	0,0005
1	1,0000	1,3764	1,9626	3,0777	6,3138	12,7062	31,8205	63,6567	636,6192
2	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	6,9646	9,9248	31,5991
3	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	4,5407	5,8409	12,9240
4	0,7407	0,9410	1,1896	1,5332	2,1318	2,7764	3,7469	4,6041	8,6103
5	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	3,3649	4,0321	6,8688
6	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,1427	3,7074	5,9588
7	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	2,9980	3,4995	5,4079
8	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	2,8965	3,3554	5,0413
9	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	2,8214	3,2498	4,7809
10	0,6998	0,8791	1,0931	1,3722	1,8125	2,2281	2,7638	3,1693	4,5869
11	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	2,7181	3,1058	4,4370
12	0,6955	0,8726	1,0832	1,3562	1,7823	2,1788	2,6810	3,0545	4,3178
13	0,6938	0,8702	1,0795	1,3502	1,7709	2,1604	2,6503	3,0123	4,2208
14	0,6924	0,8681	1,0763	1,3450	1,7613	2,1448	2,6245	2,9768	4,1405
15	0,6912	0,8662	1,0735	1,3406	1,7531	2,1314	2,6025	2,9467	4,0728
16	0,6901	0,8647	1,0711	1,3368	1,7459	2,1199	2,5835	2,9208	4,0150
17	0,6892	0,8633	1,0690	1,3334	1,7396	2,1098	2,5669	2,8982	3,9651
18	0,6884	0,8620	1,0672	1,3304	1,7341	2,1009	2,5524	2,8784	3,9216
19	0,6876	0,8610	1,0655	1,3277	1,7291	2,0930	2,5395	2,8609	3,8834
20	0,6870	0,8600	1,0640	1,3253	1,7247	2,0860	2,5280	2,8453	3,8495
21	0,6864	0,8591	1,0627	1,3232	1,7207	2,0796	2,5176	2,8314	3,8193
22	0,6858	0,8583	1,0614	1,3212	1,7171	2,0739	2,5083	2,8188	3,7921
23	0,6853	0,8575	1,0603	1,3195	1,7139	2,0687	2,4999	2,8073	3,7676
24	0,6848	0,8569	1,0593	1,3178	1,7109	2,0639	2,4922	2,7969	3,7454
25	0,6844	0,8562	1,0584	1,3163	1,7081	2,0595	2,4851	2,7874	3,7251
26	0,6840	0,8557	1,0575	1,3150	1,7056	2,0555	2,4786	2,7787	3,7066
27	0,6837	0,8551	1,0567	1,3137	1,7033	2,0518	2,4727	2,7707	3,6896
28	0,6834	0,8546	1,0560	1,3125	1,7011	2,0484	2,4671	2,7633	3,6739
29	0,6830	0,8542	1,0553	1,3114	1,6991	2,0452	2,4620	2,7564	3,6594
30	0,6828	0,8538	1,0547	1,3104	1,6973	2,0423	2,4573	2,7500	3,6460
40	0,6807	0,8507	1,0500	1,3031	1,6839	2,0211	2,4233	2,7045	3,5510
50	0,6794	0,8489	1,0473	1,2987	1,6759	2,0086	2,4033	2,6778	3,4960
60	0,6786	0,8477	1,0455	1,2958	1,6706	2,0003	2,3901	2,6603	3,4602
70	0,6780	0,8468	1,0442	1,2938	1,6669	1,9944	2,3808	2,6479	3,4350
80	0,6776	0,8461	1,0432	1,2922	1,6641	1,9901	2,3739	2,6387	3,4163
90	0,6772	0,8456	1,0424	1,2910	1,6620	1,9867	2,3685	2,6316	3,4019
100	0,6770	0,8452	1,0418	1,2901	1,6602	1,9840	2,3642	2,6259	3,3905
110	0,6767	0,8449	1,0413	1,2893	1,6588	1,9818	2,3607	2,6213	3,3812
120	0,6765	0,8446	1,0409	1,2886	1,6577	1,9799	2,3578	2,6174	3,3735
1000000	0,6745	0,8416	1,0364	1,2816	1,6449	1,9600	2,3264	2,5758	3,2905

TABLA Nº 5
DISTRIBUCIÓN F DE FISHER



Ejemplos:

$Para\ n_1 = 9\ ;\ n_2 = 12\ grados\ de\ libertad$

$P(F > 2{,}80) = 0{,}05 = 5\%$

$\boldsymbol{P(F > 4{,}39) = 0{,}01 = 1\%}$

5% (normal) **y 1% (negritas)**

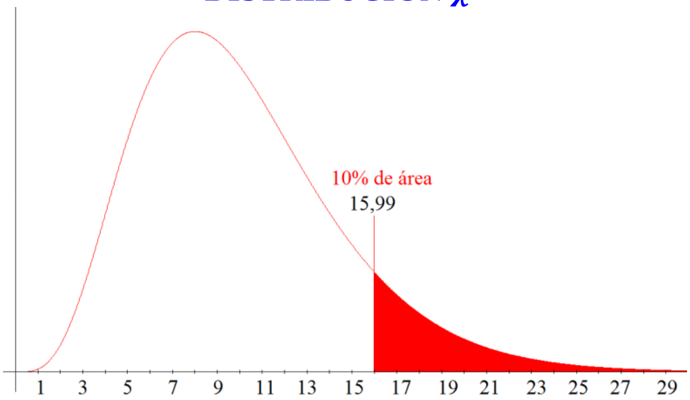
n_1 = grados de libertad del numerador

n_2 = grados de libertad del denominador

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	15	20	25	50	100
1	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88	240,54	241,88	245,95	248,01	249,26	251,77	253,04
	4052,2	**4999,5**	**5403,4**	**5624,6**	**5763,6**	**5859,0**	**5928,4**	**5981,1**	**6022,5**	**6055,8**	**6157,3**	**6208,7**	**6239,8**	**6302,5**	**6334,1**
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,43	19,45	19,46	19,48	19,49
	98,50	**99,00**	**99,17**	**99,25**	**99,30**	**99,33**	**99,36**	**99,37**	**99,39**	**99,40**	**99,43**	**99,45**	**99,46**	**99,48**	**99,49**
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66	8,63	8,58	8,55
	34,12	**30,82**	**29,46**	**28,71**	**28,24**	**27,91**	**27,67**	**27,49**	**27,35**	**27,23**	**26,87**	**26,69**	**26,58**	**26,35**	**26,24**
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80	5,77	5,70	5,66
	21,20	**18,00**	**16,69**	**15,98**	**15,52**	**15,21**	**14,98**	**14,80**	**14,66**	**14,55**	**14,20**	**14,02**	**13,91**	**13,69**	**13,58**
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56	4,52	4,44	4,41
	16,26	**13,27**	**12,06**	**11,39**	**10,97**	**10,67**	**10,46**	**10,29**	**10,16**	**10,05**	**9,72**	**9,55**	**9,45**	**9,24**	**9,13**
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87	3,83	3,75	3,71
	13,75	**10,92**	**9,78**	**9,15**	**8,75**	**8,47**	**8,26**	**8,10**	**7,98**	**7,87**	**7,56**	**7,40**	**7,30**	**7,09**	**6,99**
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44	3,40	3,32	3,27
	12,25	**9,55**	**8,45**	**7,85**	**7,46**	**7,19**	**6,99**	**6,84**	**6,72**	**6,62**	**6,31**	**6,16**	**6,06**	**5,86**	**5,75**
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15	3,11	3,02	2,97
	11,26	**8,65**	**7,59**	**7,01**	**6,63**	**6,37**	**6,18**	**6,03**	**5,91**	**5,81**	**5,52**	**5,36**	**5,26**	**5,07**	**4,96**
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94	2,89	2,80	2,76
	10,56	**8,02**	**6,99**	**6,42**	**6,06**	**5,80**	**5,61**	**5,47**	**5,35**	**5,26**	**4,96**	**4,81**	**4,71**	**4,52**	**4,41**
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77	2,73	2,64	2,59
	10,04	**7,56**	**6,55**	**5,99**	**5,64**	**5,39**	**5,20**	**5,06**	**4,94**	**4,85**	**4,56**	**4,41**	**4,31**	**4,12**	**4,01**
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65	2,60	2,51	2,46
	9,65	**7,21**	**6,22**	**5,67**	**5,32**	**5,07**	**4,89**	**4,74**	**4,63**	**4,54**	**4,25**	**4,10**	**4,01**	**3,81**	**3,71**
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54	2,50	2,40	2,35
	9,33	**6,93**	**5,95**	**5,41**	**5,06**	**4,82**	**4,64**	**4,50**	**4,39**	**4,30**	**4,01**	**3,86**	**3,76**	**3,57**	**3,47**
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46	2,41	2,31	2,26
	9,07	**6,70**	**5,74**	**5,21**	**4,86**	**4,62**	**4,44**	**4,30**	**4,19**	**4,10**	**3,82**	**3,66**	**3,57**	**3,38**	**3,27**
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39	2,34	2,24	2,19
	8,86	**6,51**	**5,56**	**5,04**	**4,69**	**4,46**	**4,28**	**4,14**	**4,03**	**3,94**	**3,66**	**3,51**	**3,41**	**3,22**	**3,11**
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33	2,28	2,18	2,12
	8,68	**6,36**	**5,42**	**4,89**	**4,56**	**4,32**	**4,14**	**4,00**	**3,89**	**3,80**	**3,52**	**3,37**	**3,28**	**3,08**	**2,98**
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28	2,23	2,12	2,07
	8,53	**6,23**	**5,29**	**4,77**	**4,44**	**4,20**	**4,03**	**3,89**	**3,78**	**3,69**	**3,41**	**3,26**	**3,16**	**2,97**	**2,86**
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23	2,18	2,08	2,02
	8,40	**6,11**	**5,18**	**4,67**	**4,34**	**4,10**	**3,93**	**3,79**	**3,68**	**3,59**	**3,31**	**3,16**	**3,07**	**2,87**	**2,76**
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19	2,14	2,04	1,98
	8,29	**6,01**	**5,09**	**4,58**	**4,25**	**4,01**	**3,84**	**3,71**	**3,60**	**3,51**	**3,23**	**3,08**	**2,98**	**2,78**	**2,68**
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16	2,11	2,00	1,94
	8,18	**5,93**	**5,01**	**4,50**	**4,17**	**3,94**	**3,77**	**3,63**	**3,52**	**3,43**	**3,15**	**3,00**	**2,91**	**2,71**	**2,60**
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12	2,07	1,97	1,91
	8,10	**5,85**	**4,94**	**4,43**	**4,10**	**3,87**	**3,70**	**3,56**	**3,46**	**3,37**	**3,09**	**2,94**	**2,84**	**2,64**	**2,54**
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,18	2,10	2,05	1,94	1,88
	8,02	**5,78**	**4,87**	**4,37**	**4,04**	**3,81**	**3,64**	**3,51**	**3,40**	**3,31**	**3,03**	**2,88**	**2,79**	**2,58**	**2,48**
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,15	2,07	2,02	1,91	1,85
	7,95	**5,72**	**4,82**	**4,31**	**3,99**	**3,76**	**3,59**	**3,45**	**3,35**	**3,26**	**2,98**	**2,83**	**2,73**	**2,53**	**2,42**
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,13	2,05	2,00	1,88	1,82
	7,88	**5,66**	**4,76**	**4,26**	**3,94**	**3,71**	**3,54**	**3,41**	**3,30**	**3,21**	**2,93**	**2,78**	**2,69**	**2,48**	**2,37**
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,11	2,03	1,97	1,86	1,80
	7,82	**5,61**	**4,72**	**4,22**	**3,90**	**3,67**	**3,50**	**3,36**	**3,26**	**3,17**	**2,89**	**2,74**	**2,64**	**2,44**	**2,33**
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,09	2,01	1,96	1,84	1,78
	7,77	**5,57**	**4,68**	**4,18**	**3,85**	**3,63**	**3,46**	**3,32**	**3,22**	**3,13**	**2,85**	**2,70**	**2,60**	**2,40**	**2,29**

26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,07	1,99	1,94	1,82	1,76
	7,72	**5,53**	**4,64**	**4,14**	**3,82**	**3,59**	**3,42**	**3,29**	**3,18**	**3,09**	**2,81**	**2,66**	**2,57**	**2,36**	**2,25**
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,06	1,97	1,92	1,81	1,74
	7,68	**5,49**	**4,60**	**4,11**	**3,78**	**3,56**	**3,39**	**3,26**	**3,15**	**3,06**	**2,78**	**2,63**	**2,54**	**2,33**	**2,22**
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,04	1,96	1,91	1,79	1,73
	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,75	2,60	2,51	2,30	2,19
29	4,18	6,94	4,76	6,39	4,39	6,16	4,21	6,04	4,10	5,96	4,62	5,80	4,52	5,70	4,41
	7,60	**5,42**	**4,54**	**4,04**	**3,73**	**3,50**	**3,33**	**3,20**	**3,09**	**3,00**	**2,73**	**2,57**	**2,48**	**2,27**	**2,16**
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01	1,93	1,88	1,76	1,70
	7,56	**5,39**	**4,51**	**4,02**	**3,70**	**3,47**	**3,30**	**3,17**	**3,07**	**2,98**	**2,70**	**2,55**	**2,45**	**2,25**	**2,13**
32	4,15	3,29	2,90	2,67	2,51	2,40	2,31	2,24	2,19	2,14	1,99	1,91	1,85	1,74	1,67
	7,50	**5,34**	**4,46**	**3,97**	**3,65**	**3,43**	**3,26**	**3,13**	**3,02**	**2,93**	**2,65**	**2,50**	**2,41**	**2,20**	**2,08**
36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	1,95	1,87	1,81	1,69	1,62
	7,40	**5,25**	**4,38**	**3,89**	**3,57**	**3,35**	**3,18**	**3,05**	**2,95**	**2,86**	**2,58**	**2,43**	**2,33**	**2,12**	**2,00**
42	4,07	3,22	2,83	2,59	2,44	2,32	2,24	2,17	2,11	2,06	1,91	1,83	1,77	1,65	1,57
	7,28	**5,15**	**4,29**	**3,80**	**3,49**	**3,27**	**3,10**	**2,97**	**2,86**	**2,78**	**2,50**	**2,34**	**2,25**	**2,03**	**1,91**
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	1,89	1,81	1,75	1,63	1,55
	7,23	**5,11**	**4,25**	**3,77**	**3,45**	**3,23**	**3,07**	**2,94**	**2,83**	**2,74**	**2,46**	**2,31**	**2,21**	**2,00**	**1,88**
54	4,02	3,17	2,78	2,54	2,39	2,27	2,18	2,12	2,06	2,01	1,86	1,77	1,71	1,58	1,51
	7,13	**5,02**	**4,17**	**3,69**	**3,38**	**3,16**	**2,99**	**2,86**	**2,76**	**2,67**	**2,39**	**2,24**	**2,14**	**1,92**	**1,79**
56	4,01	3,16	2,77	2,54	2,38	2,27	2,18	2,11	2,05	2,00	1,85	1,76	1,70	1,57	1,50
	7,11	**5,01**	**4,15**	**3,67**	**3,36**	**3,14**	**2,98**	**2,85**	**2,74**	**2,66**	**2,38**	**2,22**	**2,12**	**1,90**	**1,78**
57	4,01	3,16	2,77	2,53	2,38	2,26	2,18	2,11	2,05	2,00	1,85	1,76	1,70	1,57	1,49
	7,10	**5,00**	**4,15**	**3,67**	**3,36**	**3,14**	**2,97**	**2,84**	**2,74**	**2,65**	**2,37**	**2,22**	**2,12**	**1,90**	**1,77**
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	1,81	1,72	1,66	1,53	1,45
	7,01	**4,92**	**4,07**	**3,60**	**3,29**	**3,07**	**2,91**	**2,78**	**2,67**	**2,59**	**2,31**	**2,15**	**2,05**	**1,83**	**1,70**
72	3,97	3,12	2,73	2,50	2,34	2,23	2,14	2,07	2,01	1,96	1,81	1,72	1,66	1,53	1,44
	7,00	**4,91**	**4,07**	**3,59**	**3,28**	**3,06**	**2,90**	**2,77**	**2,66**	**2,58**	**2,30**	**2,14**	**2,04**	**1,82**	**1,69**
76	3,97	3,12	2,72	2,49	2,33	2,22	2,13	2,06	2,01	1,96	1,80	1,71	1,65	1,52	1,43
	6,98	**4,90**	**4,05**	**3,58**	**3,27**	**3,05**	**2,88**	**2,75**	**2,65**	**2,56**	**2,28**	**2,13**	**2,03**	**1,80**	**1,67**
84	3,95	3,11	2,71	2,48	2,32	2,21	2,12	2,05	1,99	1,95	1,79	1,70	1,64	1,50	1,42
	6,95	**4,87**	**4,02**	**3,55**	**3,24**	**3,02**	**2,86**	**2,73**	**2,63**	**2,54**	**2,26**	**2,10**	**2,00**	**1,78**	**1,64**
87	3,95	3,10	2,71	2,48	2,32	2,20	2,12	2,05	1,99	1,94	1,78	1,69	1,63	1,50	1,41
	6,94	**4,86**	**4,02**	**3,54**	**3,24**	**3,02**	**2,85**	**2,72**	**2,62**	**2,53**	**2,25**	**2,10**	**1,99**	**1,77**	**1,63**
95	3,94	3,09	2,70	2,47	2,31	2,20	2,11	2,04	1,98	1,93	1,77	1,68	1,62	1,48	1,40
	6,91	**4,84**	**3,99**	**3,52**	**3,22**	**3,00**	**2,83**	**2,70**	**2,60**	**2,51**	**2,23**	**2,08**	**1,98**	**1,75**	**1,61**
96	3,94	3,09	2,70	2,47	2,31	2,19	2,11	2,04	1,98	1,93	1,77	1,68	1,62	1,48	1,40
	6,91	**4,83**	**3,99**	**3,52**	**3,21**	**3,00**	**2,83**	**2,70**	**2,60**	**2,51**	**2,23**	**2,07**	**1,97**	**1,74**	**1,61**
116	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,75	1,66	1,60	1,46	1,37
	6,86	**4,79**	**3,96**	**3,49**	**3,18**	**2,96**	**2,80**	**2,67**	**2,56**	**2,48**	**2,20**	**2,04**	**1,94**	**1,71**	**1,57**
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,75	1,66	1,60	1,46	1,37
	6,85	**4,79**	**3,95**	**3,48**	**3,17**	**2,96**	**2,79**	**2,66**	**2,56**	**2,47**	**2,19**	**2,03**	**1,93**	**1,70**	**1,56**
145	3,91	3,06	2,67	2,43	2,28	2,16	2,07	2,00	1,94	1,90	1,74	1,64	1,58	1,44	1,35
	6,81	**4,75**	**3,92**	**3,45**	**3,15**	**2,93**	**2,76**	**2,64**	**2,53**	**2,45**	**2,16**	**2,01**	**1,90**	**1,67**	**1,53**
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,72	1,62	1,56	1,41	1,32
	6,76	**4,71**	**3,88**	**3,41**	**3,11**	**2,89**	**2,73**	**2,60**	**2,50**	**2,41**	**2,13**	**1,97**	**1,87**	**1,63**	**1,48**

TABLA Nº 6
DISTRIBUCIÓN χ^2



Ejemplo:
Para 10 grados de libertad
$P(\chi^2 > 15,99) = 0,10 = 10\%$

	0,995	0,990	0,975	0,950	0,900	0,750	0,500	0,250	0,100	0,050	0,025	0,010	0,005
1	0,000	0,000	0,001	0,004	0,016	0,102	0,455	1,323	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	0,575	1,386	2,773	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	1,213	2,366	4,108	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	1,923	3,357	5,385	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	2,675	4,351	6,626	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	3,455	5,348	7,841	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	4,255	6,346	9,037	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	5,071	7,344	10,219	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	5,899	8,343	11,389	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	6,737	9,342	12,549	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	7,584	10,341	13,701	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	8,438	11,340	14,845	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	9,299	12,340	15,984	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	10,165	13,339	17,117	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	11,037	14,339	18,245	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	11,912	15,338	19,369	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	12,792	16,338	20,489	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,390	10,865	13,675	17,338	21,605	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	14,562	18,338	22,718	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	15,452	19,337	23,828	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	16,344	20,337	24,935	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	17,240	21,337	26,039	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	18,137	22,337	27,141	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	19,037	23,337	28,241	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	19,939	24,337	29,339	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	20,843	25,336	30,435	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	21,749	26,336	31,528	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	22,657	27,336	32,620	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	23,567	28,336	33,711	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	24,478	29,336	34,800	40,256	43,773	46,979	50,892	53,672
40	20,707	22,164	24,433	26,509	29,051	33,660	39,335	45,616	51,805	55,758	59,342	63,691	66,766
50	27,991	29,707	32,357	34,764	37,689	42,942	49,335	56,334	63,167	67,505	71,420	76,154	79,490
60	35,534	37,485	40,482	43,188	46,459	52,294	59,335	66,981	74,397	79,082	83,298	88,379	91,952
70	43,275	45,442	48,758	51,739	55,329	61,698	69,334	77,577	85,527	90,531	95,023	100,425	104,215
80	51,172	53,540	57,153	60,391	64,278	71,145	79,334	88,130	96,578	101,879	106,629	112,329	116,321
90	59,196	61,754	65,647	69,126	73,291	80,625	89,334	98,650	107,565	113,145	118,136	124,116	128,299
100	67,328	70,065	74,222	77,929	82,358	90,133	99,334	109,141	118,498	124,342	129,561	135,807	140,169
110	75,550	78,458	82,867	86,792	91,471	99,666	109,334	119,608	129,385	135,480	140,917	147,414	151,948
120	83,852	86,923	91,573	95,705	100,624	109,220	119,334	130,055	140,233	146,567	152,211	158,950	163,648

FORMULARIO CON EJEMPLOS ILUSTRATIVOS RESUELTOS SOBRE CONOCIMIENTOS ESTADÍSTICOS BÁSICOS

TAMAÑO DE LA MUESTRA: Para que una muestra sea fiable, es necesario que su tamaño sea obtenido mediante procesos matemáticos que eliminen la incidencia del error. Para calcular el tamaño de la muestra suele utilizarse la siguiente fórmula:

$$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2}$$

Donde:

$\sigma =$ Desviación estándar de la población que rara vez conoce su valor. Generalmente cuando no se tiene su valor, suele utilizarse un valor constante de 0,5.

Z = Valor obtenido mediante niveles de confianza. Es un valor constante que, si no se tiene su valor, se lo toma en relación al 95% de confianza equivale a 1,96 (como más usual) o en relación al 99% de confianza equivale 2,58, valor que queda a criterio del investigador de acuerdo al nivel de profundidad de la investigación.

e = Límite aceptable de error muestral que, generalmente cuando no se tiene su valor, suele utilizarse un valor que varía entre el 1% (0,01) y 9% (0,09), valor que queda a criterio del encuestador de acuerdo al nivel de profundidad de la investigación, sin embargo, se aconseja emplear el 5% o el 1%, por ser valores que guardan relación con el 95% y 99% de confianza, respectivamente.

Ejemplo ilustrativo: Calcular el tamaño de la muestra de una población de 1000 elementos.

Solución: *En Excel:*

	A	B	C	D	E	F	G
1	N	**1000**					
2	σ	0,5					
3	Z	1,96					
4	e	0,05					
5	$n = \frac{N\sigma^2 Z^2}{(N-1)e^2 + \sigma^2 Z^2}$	**277,74083**	=(B1*B2^2*B3^2)/((B1-1)*B4^2+B2^2*B3^2)				
6							

DISTRIBUCIÓN DE FRECUENCIAS

- *Frecuencia Absoluta* (f).- Es el número de veces que se repite el valor de cada variable. La suma de frecuencias absolutas es siempre al total de datos observados.

- *Frecuencia Relativa* (fr).- Indica la proporción con que se repite un valor. Es el cociente entre la frecuencia absoluta y el número total de datos. La suma de las frecuencias relativas es siempre 1

$$fr = \frac{f}{n}$$

- *Frecuencia Acumulada* (fa).- Indica el número de valores que son menores o iguales que el valor dado. Al sumar las frecuencias absolutas desde el menor puntaje hacia arriba tenemos la frecuencia acumulada, es decir, es la suma de la frecuencia absoluta primera con la segunda, este valor con la tercera, y así sucesivamente.

- ***Frecuencia Porcentual*** **(*$f\%$*).-** Llamada también frecuencia relativa porcentual. Se obtiene multiplicando la frecuencia relativa por 100. La suma de las frecuencias porcentuales es siempre 100%. Se calcula así:

$$f\% = fr \cdot 100$$

- ***Frecuencia Relativa Acumulada*** **(*fra*).-** Es la suma de la frecuencia relativa primera con la segunda, este valor con la tercera, y así sucesivamente.

- ***Frecuencia Relativa Acumulada Porcentual*** **(*$fra\%$*).-** Indica el número de valores que son menores o iguales que el valor dado. Se obtiene multiplicando la frecuencia relativa acumulada por 100. Se calcula así:

$$fra\% = fra \cdot 100$$

Ejemplo ilustrativo: Calcular las diferentes frecuencias de las siguientes calificaciones evaluadas sobre 10 obtenidas de 40 estudiantes en la asignatura de Estadística sin agrupar en clases:

10	8	9	8	7	8	9	10
6	7	10	9	8	8	10	8
6	5	6	8	10	5	9	9
8	10	9	7	6	7	7	6
8	10	7	8	5	9	8	5

Solución:

En Excel:

	A	B	C	D	E	F	G	H	I	J	K	L
1	10	8	9	8	7	8	9	10				
2	6	7	10	9	8	8	10	8				
3	6	5	6	8	10	5	9	9				
4	8	10	9	7	6	7	7	6				
5	8	10	7	8	5	9	8	5				
6												
7	Calificación	f	fr		fa		f%		fra		fra%	
8	5	4	0,1	=B8/$B$14	4	=B8	10	=C8*100	0,1	=C8	10	=I8*100
9	6	5	0,125	=B9/$B$14	9	=B9+E8	12,5	=C9*100	0,225	=C9+J8	22,5	=I9*100
10	7	6	0,15	=B10/$B$14	15	=B10+E9	15	=C10*100	0,375	=C10+J9	37,5	=I10*100
11	8	11	0,275	=B11/$B$14	26	=B11+E10	27,5	=C11*100	0,65	=C11+J10	65	=I11*100
12	9	7	0,175	=B12/$B$14	33	=B12+E11	17,5	=C12*100	0,825	=C12+J11	82,5	=I12*100
13	10	7	0,175	=B13/$B$14	40	=B13+E12	17,5	=C13*100	1	=C13+J12	100	=I13*100
14	Total	40	1	=SUMA(C8:C13)			100	=SUMA(G8:G13)				
15		=SUMA(B8:B13)										

- ***Rango (R).-*** También se llama recorrido o amplitud total. Es la diferencia entre el valor mayor y el menor de los datos.

$$R = x_{máx} - x_{mín}$$

- ***Número de Intervalos de Clase*** **(*n_i*).-** No debe ser menor de 5 y mayor de 12, ya que un número mayor o menor de clases podría oscurecer el comportamiento de los datos. Para calcular el número de intervalos se aplica la regla de Sturges, propuesta por Herberth Sturges en 1926:

$$n_i = 1 + 3{,}32 \cdot log(n)$$

Siendo n el tamaño de la muestra.

- ***Ancho del Intervalo (i).-*** Se obtiene dividiendo el Rango para el número de intervalos

$$i = \frac{R}{ni}$$

Cuando el valor de i no es exacto, se debe redondear al valor superior más cercano. Esto altera el valor de rango por lo que es necesario efectuar un ajuste así:

$$Nuevo\ R = ni \cdot i$$

- ***Intervalos de Clase agregando $i-1$*** al límite inferior de cada clase, comenzando por el $x_{mín}$ del rango.

- ***Marca de Clase (x_m).-*** Es el valor medio de cada clase, se obtiene sumando los límites superior (L_s) e inferior (L_i) del intervalo y dividiendo ésta suma entre 2

$$x_m = \frac{L_s + L_i}{2}$$

Ejemplo ilustrativo: A 40 estudiantes se les pidió que estimen el número de horas que habrían dedicado a estudiar la semana pasada (tanto en clase como fuera de ella), obteniéndose los siguientes resultados:

36	30	47	60	32	35	40	50
54	35	45	52	48	58	60	38
32	35	56	48	30	55	49	39
58	50	65	35	56	47	37	56
58	50	47	58	55	39	58	45

Solución:

1) Calculando el Rango se obtiene:

$R = x_{máx} - x_{mín} = 65 - 30 = 35$

2) Calculando el número de intervalos se obtiene:

$n_i = 1 + 3{,}32 \cdot log(n) = 1 + 3{,}32 \cdot log40 = 6{,}32 = 6$

3) Calculando el ancho se obtiene:

$i = \frac{R}{ni} = \frac{35}{6} = 5{,}83$

Redondeando se obtiene: i = 6, por lo que es necesario realizar un ajuste al rango.

En Excel:

	A	B	C	D	E	F	G	H
1	36	30	47	60	32	35	40	50
2	54	35	45	52	48	58	60	38
3	32	35	56	48	30	55	49	39
4	58	50	65	35	56	47	37	56
5	58	50	47	58	55	39	58	45
6								
7	n	40	=CONTAR(A1:H5)					
8	$x_{máx}$	65	=MAX(A1:H5)					
9	x_{min}	30	=MIN(A1:H5)					
10	R	35	=B8-B9	=MAX(A1:H5)-MIN(A1:H5)				
11	n_i	6	=ENTERO(1+3,32*LOG10(B7))					
12	i	6	=B10/B11					

4) Calculando el nuevo rango se obtiene:

$Nuevo\ R = ni \cdot i = 6 \cdot 6 = 36$

El exceso de 1 que se tiene en este caso se distribuye entre $x_{máx}$ y $X_{mín}$. En este ejemplo, se podría agregar 1 al valor mayor y no quitar nada al valor menor, o no agregar nada al mayor y quitar 1 al menor. Al elegir la primera opción se obtiene:

$x_{máx} = 65 + 1 = 66$
$x_{min} = 30 - 0 = 30$

5) Formando los intervalos de clase agregando $i - 1$ (6-1=5) al límite inferior de cada clase, comenzando por el $x_{mín}$ del rango se obtiene:
30+5 = 35; 36+5 = 41; 42+5 = 47; 48+5 = 53; 54+5 = 59; 60+5 = 65

En Excel:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	36	30	47	60	32	35	40	50								
2	54	35	45	52	48	58	60	38								
3	32	35	56	48	30	55	49	39								
4	58	50	65	35	56	47	37	56								
5	58	50	47	58	55	39	58	45								
6																
7	Clases		f	xm			fr		fa		f%		fra		fra%	
8	30	35	8	32,5	=(A8+B8)/2		0,2	=C8/$C$14	8	=C8	20	=G8*100	0,2	=G8	20	=M8*100
9	36	41	6	38,5	=(A9+B9)/2		0,15	=C9/$C$14	14	=C9+I8	15	=G9*100	0,35	=G9+M8	35	=M9*100
10	42	47	5	44,5	=(A10+B10)/2		0,125	=C10/$C$14	19	=C10+I9	12,5	=G10*100	0,475	=G10+M9	47,5	=M10*100
11	48	53	7	50,5	=(A11+B11)/2		0,175	=C11/$C$14	26	=C11+I10	17,5	=G11*100	0,65	=G11+M10	65	=M11*100
12	54	59	11	56,5	=(A12+B12)/2		0,275	=C12/$C$14	37	=C12+I11	27,5	=G12*100	0,925	=G12+M11	92,5	=M12*100
13	60	65	3	62,5	=(A13+B13)/2		0,075	=C13/$C$14	40	=C13+I12	7,5	=G13*100	1	=G13+M12	100	=M13*100
14	Total		40	=SUMA(C8:C13)			1	=SUMA(H8:H13)			100					

MEDIDAS DE TENDENCIA CENTRAL

MEDIA ARITMÉTICA

MEDIA ARITMÉTICA SIMPLE

a) Para Datos sin Agrupar

La media de una población es el parámetro μ (que se lee "miu"). Si hay N observaciones en el conjunto de datos de la población, la media se calcula así:

$$\mu = \frac{x_1 + x_2 + x_3 + \cdots + x_N}{N} = \frac{\sum x_i}{N}$$

La media de una muestra es un estadístico $\bar{x}$ (que se lee "x barra"). Con n observaciones en el conjunto de datos de la muestra $(x_1, x_2, \ldots)$, la media se determina así:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

b) Para Datos Agrupados en Tablas de Frecuencias.- Cuando una serie se la agrupa en *serie simple con frecuencias* para obtener la media aritmética, se multiplica la variable por la frecuencia respectiva (f), luego se obtiene la suma de todos estos productos y luego a este valor se lo divide para el número de elementos (n). Todo esto puede representarse mediante una fórmula matemática, así:

$$\bar{x} = \frac{f_1 \cdot x_1 + f_2 \cdot x_2 + f_3 \cdot x_3 + \cdots + f_n \cdot x_n}{f_1 + f_2 + f_3 + \cdots f_n} = \frac{\sum f_i \cdot x_i}{\sum f} = \frac{\sum fx}{n}$$

Donde $n = \sum f$ es la frecuencia total (o sea, el número total de casos)

c) Para Datos Agrupados en Intervalos.- Cuando una serie se la agrupa en *intervalos* para obtener la media aritmética, se multiplica la marca de clase de intervalo (xm) por la frecuencia respectiva (f), luego se obtiene la suma de todos estos productos y luego a este valor se lo divide para el número de elementos. Todo esto se representa mediante la siguiente fórmula matemática:

$$\bar{x} = \frac{f_1 \cdot xm_1 + f_2 \cdot xm_2 + f_3 \cdot xm_3 + \cdots f_n \cdot xm_n}{f_1 + f_2 + f_3 + \cdots f_n} = \frac{\sum f_i \cdot xm_i}{\sum f} = \frac{\sum f \cdot xm}{n}$$

Ejemplo ilustrativo: Calcular la media aritmética de las siguientes calificaciones de Estadística tomadas de una muestra de 20, sin agrupar, agrupando en tablas de frecuencias y agrupando en intervalos.

4, 8, 10, 10, 5, 10, 9, 8, 6, 8, 10, 8, 5, 7, 4, 4, 8, 8, 6 y 6

Solución:

1) Sin agrupar

$$\bar{x} = \frac{\sum x_i}{n} = \frac{4+8+10+10+5+10+9+8+6+8+10+8+5+7+4+4+8+8+6+6}{20}$$

$$\bar{x} = \frac{144}{20} = 7{,}2$$

En Excel:

	A	B	C	D
1	4	8	10	10
2	5	10	9	8
3	6	8	10	8
4	5	7	4	4
5	8	8	6	6
6				
7	$\bar{x}$	7,2	=PROMEDIO(A1:D5)	

2) Agrupando en tablas de frecuencias

x	f
4	3
5	2
6	3
7	1
8	6
9	1
10	4
Total	20

$$\bar{x} = \frac{\sum fx}{n} = \frac{3\cdot 4 + 2\cdot 5 + 3\cdot 6 + 1\cdot 7 + 6\cdot 8 + 1\cdot 9 + 4\cdot 10}{3+2+3+1+6+1+4} = \frac{144}{20} = 7{,}2$$

En Excel:

	A	B	C	D	E	F
1	4	8	10	10		
2	5	10	9	8		
3	6	8	10	8		
4	5	7	4	4		
5	8	8	6	6		
6						
7	x	f				
8	4	3	=CONTAR.SI($A$1:$D$5;A8)			
9	5	2	=CONTAR.SI($A$1:$D$5;A9)			
10	6	3	=CONTAR.SI($A$1:$D$5;A10)			
11	7	1	=CONTAR.SI($A$1:$D$5;A11)			
12	8	6	=CONTAR.SI($A$1:$D$5;A12)			
13	9	1	=CONTAR.SI($A$1:$D$5;A13)			
14	10	4	=CONTAR.SI($A$1:$D$5;A14)			
15	Total	20	=SUMA(B8:B14)			
16	$\bar{x}$	**7,2**	**=SUMAPRODUCTO(A8:A14;B8:B14)/B15**			

3) Agrupando en intervalos

Intervalos	f	xm
4- 5	5	4,5
6 -7	4	6,5
8- 9	7	8,5
10-11	4	10,5

$$\bar{x} = \frac{\sum f \cdot xm}{n} = \frac{5 \cdot 4{,}5 + 4 \cdot 6{,}5 + 7 \cdot 8{,}5 + 4 \cdot 10{,}5}{5 + 4 + 7 + 4} = \frac{150}{20} = 7{,}5$$

Nota: Cuando se agrupa en intervalos los cálculos son sólo aproximaciones

En Excel:

	A	B	C	D	E	F	G
1	4	8	10	10			
2	5	10	9	8			
3	6	8	10	8			
4	5	7	4	4			
5	8	8	6	6			
6							
7	$X_{máx}$	10	=MAX(A1:D5)				
8	$X_{mín}$	4	=MIN(A1:D5)				
9	Rango	6	=B7-B8				
10	n	20	=CONTAR(A1:D5)				
11	n_i	5	=ENTERO(1+3,32*LOG(B10))				
12	i	1,2	=B9/B11				
13							
14	Intervalos		xm	f	{=FRECUENCIA(A1:D5;B15:B18)}		
15	4	5	4,5	5			
16	6	7	6,5	4			
17	8	9	8,5	7			
18	10	11	10,5	4			
19				20	=SUMA(D15:D18)		
20							
21	$\bar{x}$	**7,5**	**=SUMAPRODUCTO(C15:C18;D15:D18)/D19**				

Ejemplo ilustrativo de un problema:

A un estudiante le han realizado cinco evaluaciones en Matemática y su media es 7,8. Si en otras dos evaluaciones obtiene 7 y 10, ¿cuál es el nuevo valor medio?.

Solución:

Simbolizando, "A un estudiante le han realizado cinco evaluaciones en Estadística y su media es 8", se obtiene:

$$\frac{x_1+x_2+x_3+x_4+x_5}{5}=7{,}8 \Rightarrow x_1+x_2+x_3+x_4+x_5=7{,}8\cdot 5=39$$

Simbolizando, "Si en otras dos pruebas obtiene 7 y 10, el nuevo valor medio será", se obtiene:

$$\frac{x_1+x_2+x_3+x_4+x_5+7+10}{7}=\bar{x}$$

Remplazando $x_1+x_2+x_3+x_4+x_5=39$ en la expresión anterior se obtiene:

$$\frac{39+7+10}{7}=\bar{x} \Longrightarrow \frac{56}{7}=\bar{x}=8$$

MEDIA ARITMÉTICA PONDERADA

Cuando los números $x_1, x_2, x_3, \ldots x_k$ se les asocian ciertos factores peso (o pesos) $p_1, p_2, p_3, \ldots p_k$, dependientes de la relevancia asignada a cada número, en tal caso se requiere calcular la media aritmética ponderada, la cual se calcula así:

$$\bar{x}=\frac{p_1\cdot x_1+p_2\cdot x_2+p_3\cdot x_3+\cdots p_k\cdot x_k}{p_1+p_2+p_3+\cdots p_k}=\frac{\sum p\cdot x}{\sum p}$$

Ejemplo ilustrativo: Se tiene una información acerca de las utilidades por pan y cantidades vendidas de panes de tres tiendas. Calcular la media aritmética promedio de la utilidad por pan.

Tienda	Utilidad/pan	Cantidad vendida
1	1	2000
2	0,8	1800
3	0,9	2100

Solución:

$$\bar{x}=\frac{\sum p\cdot x}{\sum p}=\frac{2000\cdot 1+1800\cdot 0{,}8+2100\cdot 0{,}9}{2000+1800+2100}=\frac{5330}{5900}=0{,}90339$$

En Excel:

	A	B	C	D	E
1	Tienda	Utilidad/pan	Cantidad vendida		
2	1	1	2000		
3	2	0,8	1800		
4	3	0,9	2100		
5			5900	=SUMA(C2:C4)	
6	$\bar{x}$	**0,9033898**	**=SUMAPRODUCTO(B2:B4;C2:C4)/C5**		

Ejemplo ilustrativo de un problema:

Una estudiante de secundaria de ecuador de la Unidad Educativa "Ibarra" obtiene en el primer quimestre, 6 en la primera parcial, 9 en la segunda parcial, 6 en la tercera parcial. Si el promedio del quimestre es de 7,2, ¿cuál fue la calificación del examen?. Recuerde que el sistema educativo ecuatoriano secundario las tres parciales aportan al promedio con el 80% y la nota del examen con el 20%

Solución

Evaluación	Calificación(x)	Ponderación(p)	$p \cdot x$
1ra parcial	6	80/3	160
2da parcial	9	80/3	240
3ra parcial	6	80/3	160
examen	x	20	$20x$
Total		100	$560 + 20x$

$$\bar{x} = \frac{\sum p \cdot x}{\sum p} = 7{,}2 = \frac{560 + 20x}{100} \Longrightarrow 100 \cdot 7{,}2 = 560 + 20x \Longrightarrow 720 - 560 = 20x \Longrightarrow 160 = 20x$$

$$x = \frac{160}{20} \Longrightarrow x = 8$$

MEDIA GEOMÉTRICA

a) Para Datos No Agrupados

Se emplea la ecuación:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdots x_n}$$

O aplicando logaritmos la ecuación:

$$logG = \frac{\log x_1 + \log x_2 + \log x_3 + \cdots \log x_n}{n}$$

Ejemplo ilustrativo: La media geométrica es útil en el cálculo de tasas de crecimiento; por ejemplo, si el crecimiento de las ventas en un pequeño negocio son 3%, 4%,8%,9% y 10%, hallar la media de crecimiento.

Solución:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdots x_n} = \sqrt[5]{3 \cdot 4 \cdot 8 \cdot 9 \cdot 10} = 6{,}128$$

Respuesta: 6,128%

O utilizando logaritmos:

$$logG = \frac{\log x_1 + \log x_2 + \log x_3 + \cdots \log x_n}{n} = \frac{\log 3 + \log 4 + \log 8 + \log 9 + \log 10}{5}$$

$$logG = \frac{0{,}4771 + 0{,}6021 + 0{,}9031 + 0{,}9542 + 1}{5} = \frac{3{,}9365}{5} = 0{,}7873$$

$$G = antilog\ 0{,}7873 = 6{,}128$$

En Excel:

	A	B	C	D
1	3			
2	4			
3	8			
4	9			
5	10			
6	G	6,12777413	=MEDIA.GEOM(A1:A5)	

b) Para Datos Agrupados en Tablas de Frecuencias

Se emplea la siguiente ecuación:

$$\log G = \frac{\sum \log x_i \cdot f_i}{n}$$

Donde:

$f_i = frecuencia\ absoluta\ de\ cada\ dato\ x_i$

Ejemplo ilustrativo: Calcular la media geométrica para las siguientes calificaciones de Estadística:

x_i	f_i
4	5
6	8
8	9
9	10
10	8

Solución:

Se llena la siguiente tabla, realizando los cálculos respectivos:

x_i	f_i	$\log x_i$	$\log x_i \cdot f_i$
4	5	0,602	3,010
6	8	0,778	6,225
8	9	0,903	8,128
9	10	0,954	9,542
10	8	1,000	8,000
Total	40		34,906

Se aplica la siguiente ecuación para obtener la respuesta.

$$\log G = \frac{\sum \log x_i \cdot f_i}{n} = \frac{34{,}906}{40} = 0{,}873$$

G = anti log 0,873 = 7,458

En Excel:

	A	B	C	D	E	F	G	H
1	x_i	f_i						
2	4	5						
3	6	8						
4	8	9						
5	9	10						
6	10	8						
7								
8	G	7,46	=10^(SUMAPRODUCTO(LOG10(A2:A6);B2:B6)/SUMA(B2:B6))					

c) Para Datos Agrupados en Intervalos

Se emplea la ecuación:

$$\log G = \frac{\sum \log xm \cdot f_i}{n}$$

Donde:

xm $= marca\ de\ clase$

MEDIA ARMÓNICA

a) Para Datos No Agrupados

Sean los números $x_1, x_2, \ldots x_n$.La media armónica H se obtiene con la siguiente ecuación:

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots \frac{1}{x_n}}$$

O con la siguiente ecuación:

$$H = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots \frac{1}{x_n}\right)}$$

Ejemplo ilustrativo: La velocidad de producción de azúcar de tres máquinas procesadoras son 0,5, 0,3 y 0,4 minutos por kilogramo. Hallar el tiempo promedio de producción después de una jornada de 4800 minutos del proceso.

Solución:

Como en la razón minutos/kilogramos (min/kg) cada máquina trabaja 4800 min, la razón contante es el tiempo de trabajo (4800 min), es decir la contante es la unidad del numerador, por lo tanto se debe emplear el promedio armónico.

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots \frac{1}{x_n}} = \frac{3}{\frac{1}{0,5} + \frac{1}{0,3} + \frac{1}{0,4}} = 0,383$$

O empleando la otra ecuación**:**

$$H = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots \frac{1}{x_n}\right)} = \frac{1}{\frac{1}{3}\left(\frac{1}{0,5} + \frac{1}{0,3} + \frac{1}{0,4}\right)} = 0,383$$

El tiempo promedio de producción es 0,383 minutos por kilogramo de azúcar.

En Excel :

	A	B	C	D
1	0,5			
2	0,3			
3	0,4			
4	H	0,38298	=MEDIA.ARMO(A1:A3)	

b) Para Datos Agrupados en Tablas de Frecuencias

Se emplea cualquiera de las siguientes ecuaciones:

$$H = \frac{n}{\sum_{i=1}^{n} \frac{f_i}{x_i}} = \frac{n}{\frac{f_1}{x_1} + \frac{f_2}{x_2} + \cdots \frac{f_n}{x_n}} \qquad H = \frac{1}{\frac{1}{n}\left(\frac{f_1}{x_1} + \frac{f_2}{x_2} + \cdots \frac{f_n}{x_n}\right)}$$

Ejemplo ilustrativo: En la siguiente tabla se presentan los datos sobre el tiempo en horas que se demoran en realizar la misma obra determinados obreros. Calcular el tiempo promedio que se demora en realizar la obra un obrero tipo (un obrero promedio).

Tiempo	Obreros
4	4
5	5
6	7
7	2
9	2

Solución:

$$H=\frac{n}{\sum_{i=1}^{n}\frac{f_i}{x_i}}=\frac{n}{\frac{f_1}{x_1}+\frac{f_2}{x_2}+\cdots\frac{f_n}{x_n}}=\frac{20}{\frac{4}{4}+\frac{5}{5}+\frac{7}{6}+\frac{2}{7}+\frac{2}{9}}=\frac{20}{\frac{463}{126}}=\frac{2520}{463}=5{,}44$$

En Excel:

	A	B	C	D	E	F	G
1	Tiempo	Obreros					
2	4	4					
3	5	5					
4	6	7					
5	7	2					
6	9	2					
7							
8	H	5,44276	=SUMA(B2:B6)/SUMAPRODUCTO((1/A2:A6);B2:B6)				

c) Para Datos Agrupados en Intervalos

Se emplea la siguiente ecuación:

$$H=\frac{n}{\sum_{i=1}^{n}\frac{f_i}{xm_i}}=\frac{n}{\frac{f_1}{xm_1}+\frac{f_2}{xm_2}+\cdots\frac{f_n}{xm_n}}$$

Ejemplo ilustrativo: En la siguiente tabla se presentan los datos sobre el tiempo en minutos que se demoran para resolver una prueba de Estadística determinados estudiantes. Calcular el tiempo promedio que se demora en resolver la prueba un estudiante tipo.

Tiempo	Estudiantes
[40-50)	4
[50-60)	8
[60-70)	10
[70-80)	7
[80-90]	11

Solución:
Realizando los cálculos respectivos se obtiene:

x_i	f_i	xm_i	f_i/xm_i
[40-50)	4	45	0,089
[50-60)	8	55	0,145
[60-70)	10	65	0,154
[70-80)	7	75	0,093
[80-90]	11	85	0,129
Total	40		0,611

Aplicado la ecuación se obtiene:

$$H = \frac{n}{\sum_{i=1}^{n} \frac{f_i}{xm_i}} = \frac{n}{\frac{f_1}{xm_1} + \frac{f_2}{xm_2} + \cdots \frac{f_n}{xm_n}} = \frac{40}{0{,}611} = 65{,}47$$

En Excel:

	A	B	C	D	E	F	G
1	Intervalos		*f*	*xm*			
2	40	50	4	45	=PROMEDIO(A2:B2)		
3	50	60	8	55	=PROMEDIO(A3:B3)		
4	60	70	10	65	=PROMEDIO(A4:B4)		
5	70	80	7	75	=PROMEDIO(A5:B5)		
6	80	90	11	85	=PROMEDIO(A6:B6)		
7			40	=SUMA(C2:C6)			
8							
9	H	=	65,473	=C7/SUMAPRODUCTO(1/(D2:D6);C2:C6)			

LA MEDIANA

La mediana, llamada algunas veces media posicional, es el valor del término medio que divide una distribución de datos ordenados en dos partes iguales, es decir, el 50% de los datos se ubican sobre la mediana o hacia los puntajes altos y el 50% restante hacia los puntajes bajos.

a) Para Datos No Agrupados

1) Si el número n de datos es impar, la mediana es el dato que se encuentra a la mitad de la lista. Para calcular su posición se aplica la siguiente ecuación:

$$Md = x_{\frac{n}{2}+\frac{1}{2}} = x_{\frac{n+1}{2}}$$

Ejemplo ilustrativo:
Calcular la mediana de las siguientes calificaciones del curso de Estadística evaluadas sobre diez: 10, 8, 6, 4, 9, 7, 10, 9 y 6

Solución:

Se ordena los datos de menor a mayor:

4	6	6	7	8	9	9	10	10
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9

Se aplica la ecuación:

$Md = x_{\frac{n+1}{2}}$

$Md = x_{\frac{9+1}{2}} = Md = x_5$

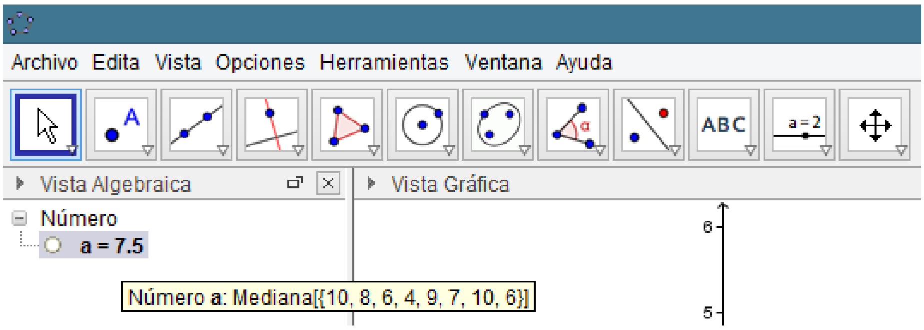
La mediana es el valor de x_5 (quinto dato), es decir, Md=8

En Excel :

	A	B	C	D	E	F	G	H
1	10	8	6	4	9	7	10	6
2								
3	Md	7,5	=MEDIANA(A1:H1)					

En GeoGebra :

Escribir los datos: Mediana[10,8,6,4,9,7,10,6]. Enter



2) Si el número n de datos es par, la mediana es la media aritmética de los dos datos que se encuentran a la mitad de la lista. Para calcular su posición se aplica la siguiente ecuación:

$$Md = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Ejemplo ilustrativo: Calcular la mediana de las siguientes calificaciones del curso de Matemática evaluadas sobre diez: 10, 8, 9, 6, 4, 8, 9, 7, 10 y 9

Solución:

Se ordena los datos de menor a mayor:

4	6	7	8	8	9	9	9	10	10
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}

$$Md = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_5 + x_6}{2} = \frac{8+9}{2} = 8{,}5$$

b) Para Datos Agrupados en Tablas de Frecuencia

Para calcular la posición de la mediana se aplica la siguiente ecuación:

$$Md = \frac{n+1}{2}$$

Ejemplo ilustrativo:

Dados los siguientes 20 números: 1, 3, 3, 5, 5, 5, 5, 2, 2, 2, 6, 6, 4, 4, 4 ,4, 5, 5, 5, 5

Agrupar los datos en tabla de frecuencia y calcular la mediana.

Solución:

Agrupando en frecuencias

x	f
1	1
2	3
3	2
4	4
5	8
6	2
Total	20

Calculando la posición de la mediana se obtiene:

$$Md = \frac{n+1}{2} = \frac{20+1}{2} = 10{,}5$$

Como la posición de la mediana es 10,5, su valor es el promedio de los datos décimo y undécimo. Para observar con claridad cuáles son los datos décimo y undécimo se aconseja calcular la frecuencia acumulada.

x	f	fa
1	1	1
2	3	4
3	2	6
4	4	10
5	8	18
6	2	20
Total	20	

Se observa que el décimo dato es 4 y el undécimo es 5, por lo tanto:

$$Md = \frac{4+5}{2} = 4{,}5$$

c) Para Datos Agrupados en Intervalos

$$Md = Li_{md} + \left(\frac{\frac{n}{2} - Fa}{f_{md}}\right) \cdot c$$

En donde:

$Li_{md} =$ Límite inferior del intervalo de clase de la mediana

$n =$ Número total de datos

$Fa =$Frecuencia acumulada del intervalo de clase que antecede al intervalo de clase de la mediana.

$fmd =$ Frecuencia absoluta del intervalo de clase de la mediana

$c =$ Ancho del intervalo

Ejemplo ilustrativo: Calcular la mediana empleando la fórmula y mediante un histograma de frecuencias acumuladas.
Se calcula la frecuencia acumulada como se muestra en la siguiente tabla:

Intervalos	f	fa
[45,55)	6	6
[55,65)	10	16
[65,75)	19	35
[75,85)	11	46
[85,95)	4	50

Solución:
Se calcula la posición de la mediana de la siguiente manera:

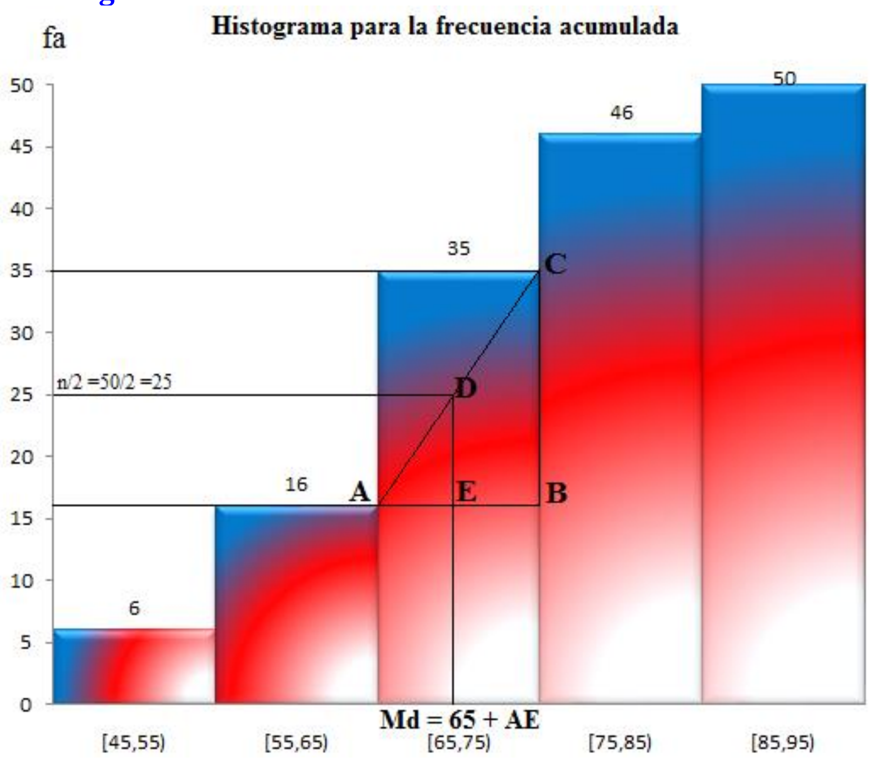
$$\frac{n}{2} = \frac{50}{2} = 25$$

Por lo tanto el intervalo o clase de la mediana es [65,75).

Al aplicar la ecuación respectiva se obtiene:

$$Md = Li_{md} + \left(\frac{\frac{n}{2} - Fa}{f_{md}}\right) \cdot c \Rightarrow Md = 65 + \left(\frac{\frac{50}{2} - 16}{19}\right) \cdot 10 = 65 + \left(\frac{9}{19}\right) \cdot 10 = 65 + \frac{90}{19} = 69{,}737$$

Resolviendo de manera gráfica



Observando el gráfico se determina que Md = 65+AE

Los triángulos ABC y AED son semejantes, por lo que se cumple:

$$\frac{AB}{CB} = \frac{AE}{DE}$$

$$\frac{75-65}{35-16} = \frac{AE}{25-16} \Longrightarrow \frac{10}{19} = \frac{AE}{9}$$

Despejando AE se obtiene:

$$\frac{10}{19} \cdot 9 = AE \Longrightarrow AE = \frac{90}{19} = 4{,}737$$

Entonces, Md = 65+AE = 65+4,737= →Md = 69,737

MEDIDAS DE POSICIÓN

Son similares a la mediana en que también subdividen una distribución de mediciones de acuerdo con la proporción de frecuencias observadas. Mientas que la mediana divide a una distribución en mitades, los cuartiles (Q) la dividen en cuartos, los deciles (D) la dividen en décimos y los puntos percentiles (P) la dividen en centésimos.
Colectivamente, cuartiles, deciles y percentiles se denominan cuantiles. Puesto que sirven para ubicar datos particulares dentro de ciertas porciones de una distribución de datos, toman el nombre de medidas de posición.

CUARTILES.- Son cada uno de los 3 valores Q_1, Q_2, Q_3 que dividen a la distribución de los datos en 4 partes iguales.

Primer cuartil: $Q_1 = P_{25}$, segundo cuartil: $Q_2 = D_5 = P_{50}$ = Mediana, tercer cuartil: $Q_3 = P_{75}$

a) Para Datos No Agrupados

La posición o ubicación de los cuartiles se encuentra aplicando la siguiente ecuación:

$$Q_k = X_{\left[\frac{n\cdot k}{4}+\frac{1}{2}\right]} = X_{\left[\frac{n\cdot k+2}{4}\right]}$$

Donde:
n = número total de datos
k = número del cuartil

Ejemplo ilustrativo: Encuentre los cuartiles dada la siguiente distribución: 6, 9, 9, 12, 12, 12, 15 y 17

Solución:
Para calcular los cuartiles se ordena los datos de menor a mayor

6	9	9	12	12	12	15	17
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8

Aplicando la ecuación para el cuartil uno se obtiene:

$$Q_k = X_{\left[\frac{n\cdot k+2}{4}\right]} = X_{\left[\frac{n+2}{4}\right]} = X_{\left[\frac{8+2}{4}\right]} = X_{\left[\frac{10}{4}\right]} = X_{2,5}$$

Como la posición del cuartil 1 es 2,5, su valor es el promedio de los datos segundo y tercero

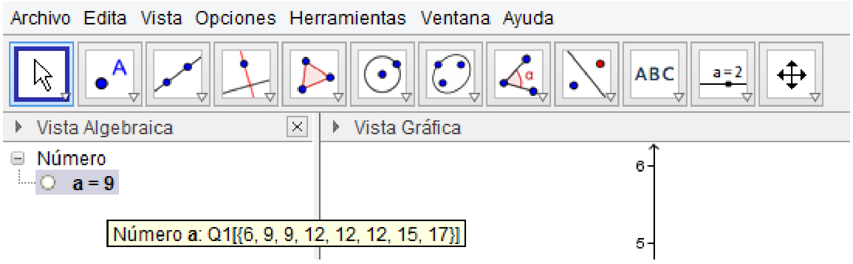
$$Q_1 = X_{2,5} = \frac{x_2 + x_3}{2} = \frac{9+9}{2} = 9$$

O también la posición 2,5 dice que el cuartil 1 está ubicado al 50% del trayecto comprendido entre el segundo dato, que es 9 y el tercer dato que es 9, es decir, Q_1= 9+0,5(9-9) = 9

Interpretación: Este resultado indica que el 25% de los datos es inferior a 9
En Excel:

	A	B	C	D
1	6			
2	9			
3	9			
4	12			
5	12			
6	12			
7	15			
8	17			
9				
10	Q_1	9	=CUARTIL.INC(A1:A8;1)	

En GeoGebra:



Aplicando la ecuación para el cuartil dos se obtiene:

$$Q_k = X_{\left[\frac{n\cdot k+2}{4}\right]} \Longrightarrow Q_2 = X_{\left[\frac{n\cdot 2+2}{4}\right]} = X_{\left[\frac{2n+2}{4}\right]} = X_{\left[\frac{2\cdot 8+2}{4}\right]} = X_{\left[\frac{16+2}{4}\right]} = X_{4,5} = \frac{x_4 + x_5}{2} = \frac{12+12}{2} = 12$$

O también la posición 4,5 dice que el cuartil 2 está ubicado al 50% del trayecto comprendido entre el cuarto dato, que es 12 y el quinto dato que también es 12, es decir,

$Q_2 = 12 + 0{,}5(12 - 12) = 12$

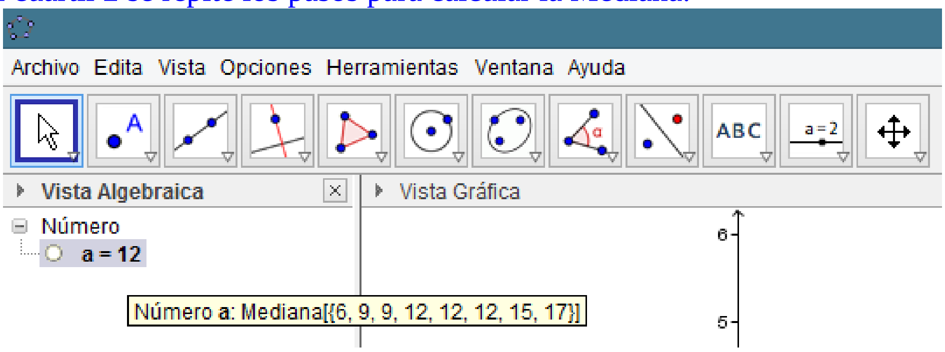
Interpretación: Este resultado indica que el 50% de los datos es inferior a 12

En Excel:

	A	B	C	D
1	6			
2	9			
3	9			
4	12			
5	12			
6	12			
7	15			
8	17			
9				
10	Q_2	12	=CUARTIL.INC(A1:A8;2)	

En GeoGebra:

Para calcular el cuartil 2 se repite los pasos para calcular la Mediana:



Aplicando la ecuación para el cuartil tres se obtiene:

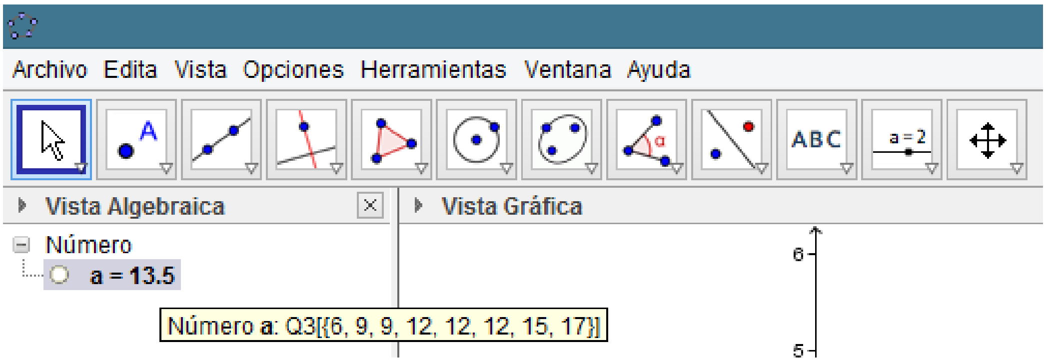
$$Q_k = X_{\left[\frac{n\cdot k+2}{4}\right]} \Longrightarrow Q_3 = X_{\left[\frac{3n+2}{4}\right]} = X_{\left[\frac{3\cdot 8+2}{4}\right]} = X_{\left[\frac{24+2}{4}\right]} = x_{\frac{26}{4}} = X_{6,5} = \frac{x_6 + x_7}{2} = \frac{12+15}{2} = 13{,}5$$

O también la posición 6,5 dice que el cuartil 2 está ubicado al 50% del trayecto comprendido entre el doceavo dato, que es 12 y el quinceavo dato que es 15, es decir, Q_3 = 12+0,5(15-12)

Q_3= 12+0,5(3)=12+1,5=13,5

Interpretación: Este resultado indica que el 75% de los datos es inferior a 13,5

En GeoGebra:



En Excel:

Repetir los pasos para el cuartil 1, y en la opción de cuartil escribir 3.

	A	B	C	D
1	6			
2	9			
3	9			
4	12			
5	12			
6	12			
7	15			
8	17			
9				
10	Q_3	12,75	=CUARTIL.INC(A1:A8;3)	

Notas importantes:

-Los cálculos en Excel para un número impar de datos coinciden con los cálculos realizados con las ecuaciones.

-Para un número par de datos, aunque en ciertas ocasiones coinciden, suele existir diferencias en los cálculos del Q_1 y Q_3 realizados con Excel. Este error de cálculo es: e = 0,25d, en donde d es la distancia de separación de los datos

-Para el Q_1 se resta el error al valor obtenido con Excel

-Para el Q_3 se suma el error al valor obtenido con Excel

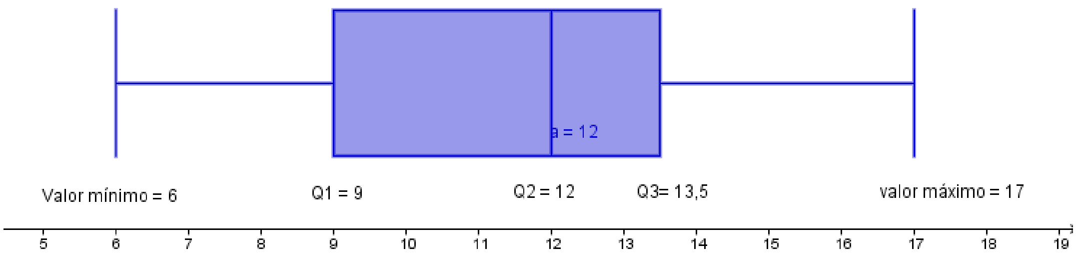
En nuestro ejemplo $e = 0{,}25(x_7 - x_6) = 0{,}25(15 - 12) = 0{,}25(3) = 0{,}75$. Al sumar el error al valor Q_3 inicialmente calculado con Excel se obtiene el valor correcto como se muestra en la siguiente figura:

	A	B	C	D	E
1	6				
2	9				
3	9				
4	12				
5	12				
6	12				
7	15				
8	17				
9					
10	Q_3	13,5	=CUARTIL.INC(A1:A8;3)+0,25*(A7-A6)		

Diagrama de caja y bigotes

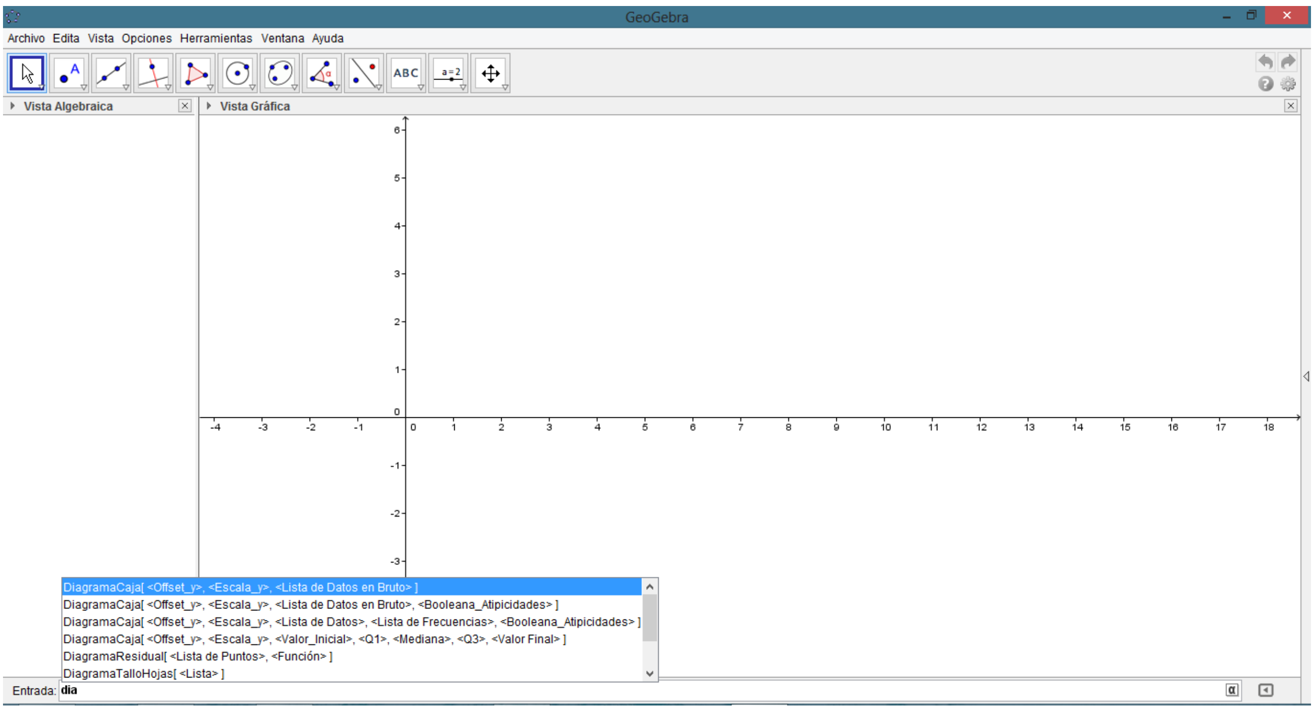
Un diagrama de caja y bigotes es una representación gráfica que ayuda a visualizar una distribución de datos: caja desde Q_1 a Q_3 (50% de los datos), y bigotes el recorrido (distancia desde valor mínimo hasta el valor máximo).

De acuerdo al ejemplo ilustrativo del cálculo de cuartiles para datos sin agrupar de la distribución de datos 6, 9, 9, 12, 12, 12, 15 y 17 se obtienen:
Valor mínimo = 6
$Q_1 = 9; Q_2 = 12;\ Q_3 = 13{,}5$
Valor máximo = 17
Por lo tanto el diagrama de caja y bigotes es:

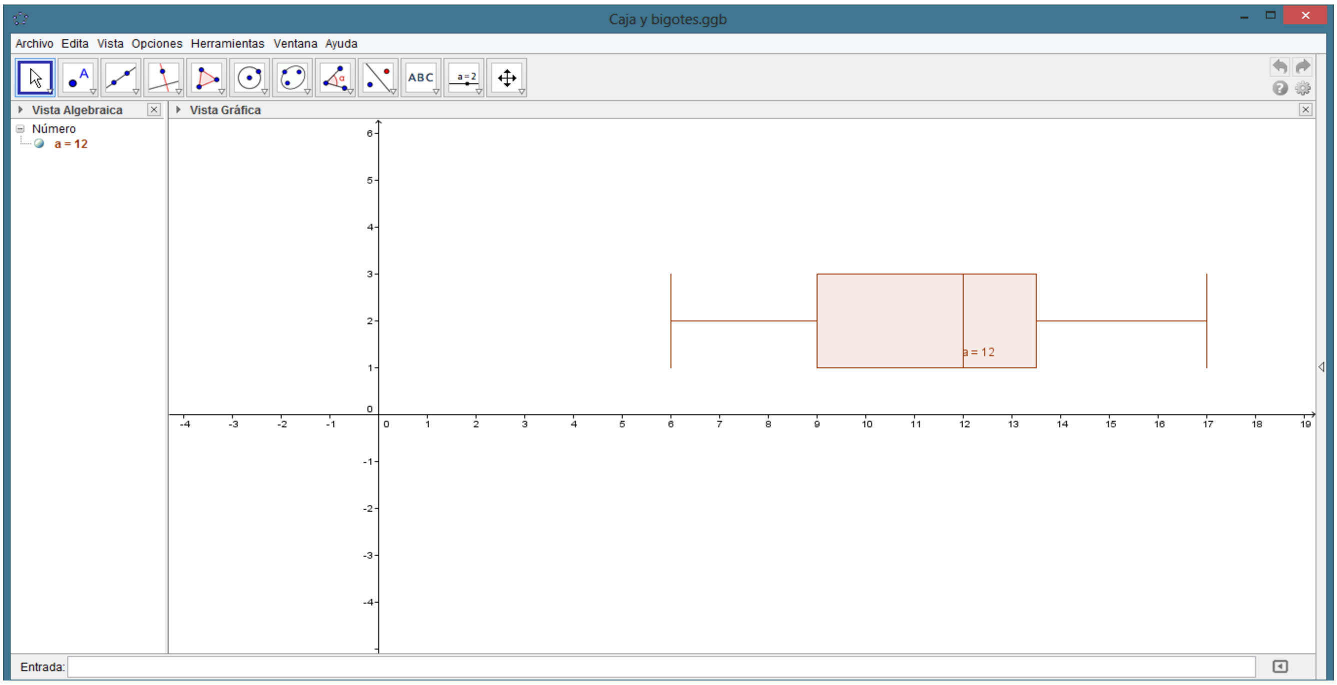


El diagrama de caja y bigotes en GeoGebra se elabora de la siguiente manera:

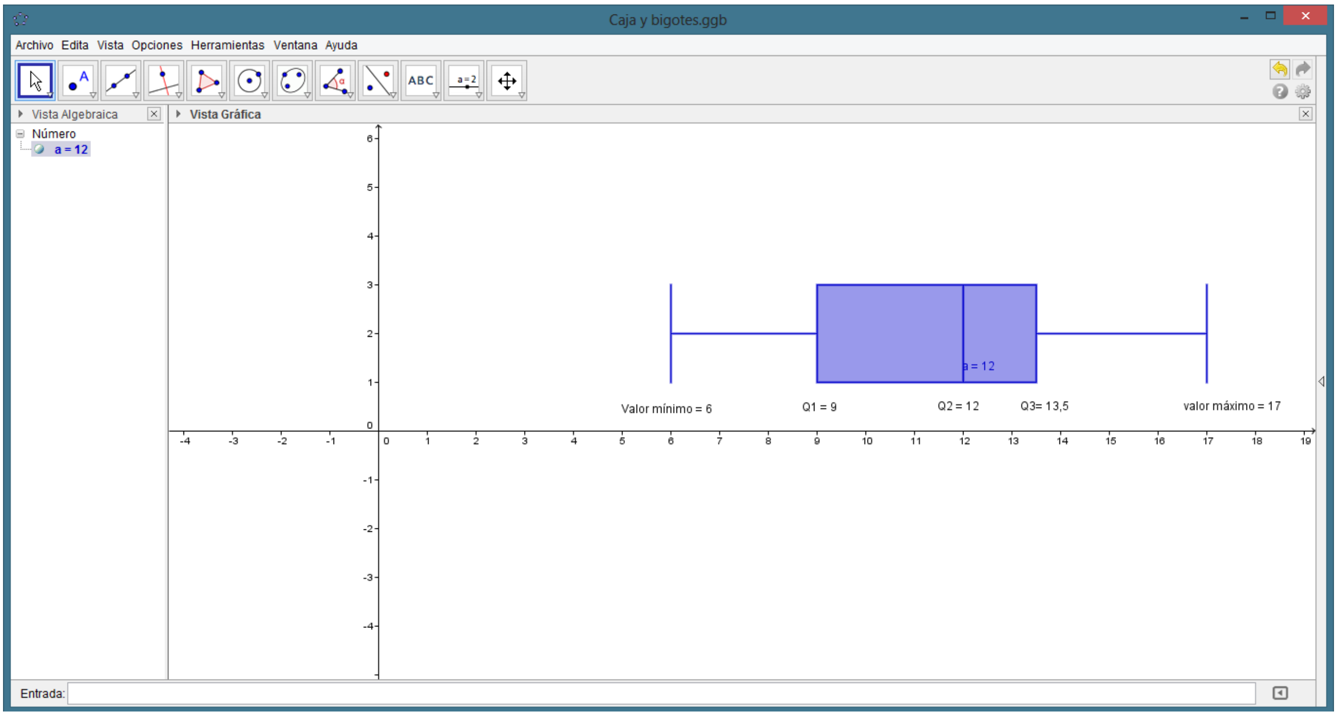
a) Ingrese al programa.En la casilla Entrada escriba las primeras letras de DiagramaCaja



b) Seleccione DiagramaCaja[<Offset_y>, <Escala_y>, <Lista de Datos en Bruto>] y dicha opción escriba DiagramaCaja[2,1,{6,9,9,12,12,12,15,17}]. Enter



c) Editando el diagrama de caja y bigotes se obtiene:



b) Para Datos Agrupados en Tablas de Frecuencias
Se aplica la misma ecuación empleada para el cálculo en los datos no agrupados

Ejemplo ilustrativo: Dada la siguiente tabla, calcular el cuartil 2:

x	f
6	1
9	2
12	3
15	1
17	1

Solución:
1) Cálculo del cuartil 2
Aplicando la primera ecuación para el cuartil dos se obtiene:

$$Q_k = X_{\left[\frac{n\cdot k+2}{4}\right]} \Longrightarrow Q_2 = X_{\left[\frac{n\cdot 2+2}{4}\right]} = X_{\left[\frac{2(n+1)}{4}\right]} = X_{\left[\frac{n+1}{2}\right]} = X_{\left[\frac{8+1}{2}\right]} = X_{\left[\frac{9}{2}\right]} = X_{4,5}$$

Como la posición del cuartil 2 es 4,5, su valor es el promedio de los datos cuarto y quinto
Para observar con claridad cuáles son los datos cuarto y quinto se aconseja calcular la frecuencia acumulada

x	f	fa
6	1	1
9	2	3
12	3	6
15	1	7
17	1	8

Se observa que el cuarto dato es 12 y el quinto dato es 12, por lo tanto

$$Q_2 = X_{4,5} = \frac{x_4 + x_5}{2} = \frac{12+12}{2} = 12$$

c) Para Datos Agrupados en Intervalos
Se emplea la siguiente ecuación:

$$\mathrm{Q_k} = \mathrm{Li_Q} + \left(\frac{\frac{\mathrm{nk}}{4} - \mathrm{Fa}}{\mathrm{f_Q}}\right)\cdot c$$

Donde:
$Li_Q =$ Límite inferior del intervalo de clase del cuartil
$n =$ Número total de datos
$Fa =$ Frecuencia acumulada del intervalo de clase que antecede al intervalo de clase del cuartil
$f_Q =$ Frecuencia absoluta del intervalo de clase del cuartil
$c =$ Ancho del intervalo de clase del cuartil

Ejemplo ilustrativo: Dado los siguientes datos sobre pesos de un grupo de 50 personas:

Intervalos	f
45- 55	6
55- 65	10
65- 75	19
75- 85	11
85- 95	4

1) Calcular los cuartiles empleando la ecuación
2) Calcular los cuartiles empleando un histograma para $fra(\%)$ (Frecuencia relativa acumulada mediada en porcentajes)

Solución:
1) Cálculo de los cuartiles empleando la ecuación

1.1) Cálculo del primer cuartil

Primero se calcula nk/4 y después se averigua el intervalo en el que está el cuartil, este intervalo recibe el nombre de intervalo o clase del primer cuartil. Para averiguar el intervalo en el que están los cuartiles se aconseja calcular la frecuencia acumulada

$$\frac{n \cdot k}{4} = \frac{50 \cdot 1}{4} = 12{,}5$$

Intervalos	f	fa
45- 55	6	6
55- 65	10	16
65- 75	19	35
75- 85	11	46
85- 95	4	50
n	50	

Por lo tanto en este ejemplo:
El intervalo del segundo cuartil es 55-65.
El número total de datos es n =10
Se observa que 6 valores están por debajo del valor 55, es decir Fa = 6.
La frecuencia absoluta f_Q del intervalo del cuartil es 10
El ancho del intervalo del cuartil es c = 65-55 =10.
Al aplicar la ecuación se obtiene:

$$Q_k = Li_Q + \left(\frac{\frac{nk}{4} - Fa}{f_Q}\right) \cdot c = 55 + \left(\frac{\frac{50 \cdot 1}{4} - 6}{10}\right) \cdot 10 = 55 + \left(\frac{\frac{50}{4} - 6}{10}\right) \cdot 10$$

$$Q_1 = 55 + \left(\frac{13}{20}\right) \cdot 10 = 55 + 6{,}5 = 61{,}5$$

1.2) Cálculo del segundo cuartil
Primero se calcula nk/4 y después se averigua el intervalo en el que está el cuartil, este intervalo recibe el nombre de intervalo o clase del cuartil.

$$\frac{n \cdot 2}{4} = \frac{50 \cdot 2}{4} = 25$$

Por lo tanto para el segundo cuartil se tiene:
Intervalo: 65-75
n=10
Fa=16
f_Q =19
c =75-65 =10

Al aplicar la ecuación se obtiene:

$$Q_k = Li_Q + \left(\frac{\frac{nk}{4} - Fa}{f_Q}\right) \cdot c$$

$$Q_2 = 65 + \left(\frac{\frac{50 \cdot 2}{4} - 16}{19}\right) \cdot 10 = 65 + \left(\frac{\frac{100}{4} - 16}{19}\right) \cdot 10 = 65 + \left(\frac{9}{19}\right) \cdot 10 = 65 + 4{,}737 = 69{,}937$$

1.3) Cálculo del tercer cuartil

Primero se calcula nk/4 y después se averigua el intervalo en el que está el cuartil, este intervalo recibe el nombre de intervalo o clase del cuartil.

$$\frac{n \cdot 3}{4} = \frac{50 \cdot 3}{4} = 37{,}5$$

Por lo tanto para el segundo cuartil se tiene:
Intervalo: 75-85

$n = 10$
$Fa = 35$
$f_Q = 11$
$c = 85 - 75 = 10$

Al aplicar la ecuación se obtiene:

$$Q_k = Li_Q + \left(\frac{\frac{nk}{4} - Fa}{f_Q}\right) \cdot c$$

$$Q_3 = 75 + \left(\frac{\frac{50 \cdot 3}{4} - 35}{11}\right) \cdot 10 = 75 + \left(\frac{\frac{150}{4} - 35}{11}\right) \cdot 10 = 75 + \left(\frac{5}{22}\right) \cdot 10 = 75 + 2{,}273 = 77{,}273$$

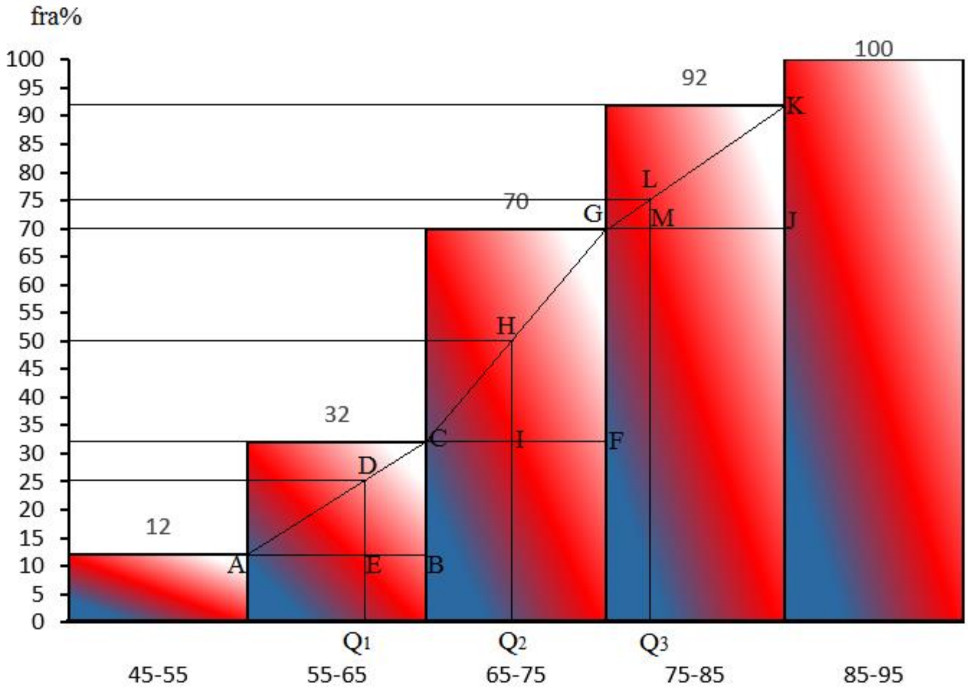
2) Cálculo de los cuartiles empleando un histograma para $fra(\%)$

2.1) Calculando la $fra(\%)$ se obtiene:

Intervalos	f	fa	fr	fra (%)
45- 55	6	6	0,12	12
55- 65	10	16	0,20	32
65- 75	19	35	0,38	70
75- 85	11	46	0,22	92
85- 95	4	50	0,08	100
N	50			

2.2) Elaborando el histograma en Excel y en Paint se obtiene la siguiente figura:

Histograma para la $fra(\%)$



2.3) Cálculo del primer cuartil

Observando en gráfico tenemos que el $Q_1 = 55 + AE$

Los triángulos ABC y AED son semejantes, por lo que se cumple:

$$\frac{AB}{CB} = \frac{AE}{DE}$$

$$\frac{65-55}{32-12} = \frac{AE}{25-12} \Longrightarrow \frac{10}{20} = \frac{AE}{13}$$

Despejando AE se obtiene:

$$\frac{10}{20} \cdot 13 = AE \Longrightarrow AE = 6{,}5$$

Entonces, $Q_1 = 55 + 6{,}5 = 61{,}5$

2.3) Cálculo del segundo cuartil

Observando en gráfico tenemos que el $Q_2 = 65 + CI$

Los triángulos CFG y CIH son semejantes, por lo que se cumple:

$$\frac{CF}{FG} = \frac{CI}{HI} \Longrightarrow \frac{75-65}{70-32} = \frac{CI}{50-32} \Longrightarrow \frac{10}{38} = \frac{CI}{18}$$

Despejando CI se obtiene:

$$\frac{10}{38} \cdot 18 = AE \Longrightarrow AE = 4{,}737$$

Entonces, $Q_2 = 65 + 4{,}737 = 69{,}737$

2.3) Cálculo del tercer cuartil
Observando en gráfico tenemos que el $Q_3 = 75 + GM$

Los triángulos GJK y GML son semejantes, por lo que se cumple:

$$\frac{GJ}{JK} = \frac{GM}{ML} \Longrightarrow \frac{85-75}{92-70} = \frac{CI}{75-70} \Longrightarrow \frac{10}{22} = \frac{CI}{5}$$

Despejando CI se obtiene:

$$\frac{10}{22} \cdot 5 = CI \Longrightarrow CI = 2{,}273$$

Entonces, $Q_3 = 75 + 2{,}273 = 77{,}273$

B) DECILES

a) Para Datos No Agrupados
La posición o ubicación de los deciles se encuentra aplicando la siguiente ecuación:

$$D_k = X_{\left[\frac{n\cdot k}{10}+\frac{1}{2}\right]} = X_{\left[\frac{n\cdot k+5}{10}\right]}$$

Donde:
n = número total de datos.
k = número del decil.

Ejemplo ilustrativo: Calcular el quinto decil de la siguiente distribución: 6, 9, 9, 12, 12, 12, 15 y 17

Solución: Para calcular los deciles se ordena los datos de menor a mayor.

6	9	9	12	12	12	15	17
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8

Aplicando la ecuación para el quinto decil se obtiene:

$$D_k = X_{\left[\frac{n\cdot k+5}{10}\right]}$$

$$D_5 = X_{\left[\frac{n\cdot 5+5}{10}\right]} = X_{\left[\frac{5n+5}{10}\right]} = X_{\left[\frac{5\cdot 8+10}{10}\right]} = X_{\left[\frac{40+5}{10}\right]} = X_{4,5} = \frac{x_4+x_5}{2} = \frac{12+12}{2} = 12$$

O también la posición 4,5 dice que el decil 5 está ubicado al 50% del trayecto comprendido entre el cuarto dato, que es 12 y el quinto dato que también es 12, es decir,
$D_5 = 12+0{,}5(12-12) = 12$

En Excel:
Como D_5 es igual a P_{50} :

	A	B	C	D	E
1	6				
2	9				
3	9				
4	12				
5	12				
6	12				
7	15				
8	17				
9					
10	$D_5 = P_{50}$	12	=PERCENTIL.INC(A1:A8;0,5)		

b) Para Datos Agrupados en Tablas de Frecuencia
Se emplea la misma ecuación utilizada en el cálculo de los deciles para datos sin agrupar.

c) Para Datos Agrupados en Intervalos
Se emplea la siguiente ecuación:

$$D_k = Li_D + \left(\frac{\frac{nk}{10} - Fa}{f_D}\right) \cdot c$$

Donde:
Li_D = Límite inferior del intervalo de clase del decil.
n = número total de datos.
F_a = Frecuencia acumulada del intervalo de clase que antecede al intervalo de clase del decil.
f_D = Frecuencia absoluta del intervalo de clase del decil.
c = Ancho del intervalo de clase del decil.

C) PERCENTILES O CENTILES
a) Para Datos No Agrupados
La posición o ubicación de los percentiles se encuentra aplicando la siguiente ecuación:

$$P_k = X_{\left[\frac{n\cdot k}{100}+\frac{1}{2}\right]} = X_{\left[\frac{n\cdot k+50}{100}\right]}$$

Donde:
n = número total de datos
k = número del percentil

Ejemplo ilustrativo: Calcular los percentiles de orden 20 y 33 del peso de diez personas que pesan (en kg)

80, 78, 65, 73, 65, 67, 72, 68, 70 y 72

Solución:
Se ordena los datos de menor a mayor se tiene:

65	65	67	68	70	72	72	73	78	80
x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}

1) Cálculo del percentil de orden 20 se obtiene:

$$P_k = X_{\left[\frac{n\cdot k+50}{100}\right]} = X_{\left[\frac{n\cdot 20+50}{100}\right]} = X_{\left[\frac{10\cdot 20+50}{100}\right]} = X_{\left[\frac{250}{100}\right]} = X_{2,5} = \frac{x_2 + x_3}{2} = \frac{65+67}{2} = 66$$

En Excel se obtiene un valor aproximado insertando la función PERCENTIL.INC(A1:A10:0,2) *como se muestra en la siguiente figura:*

	A	B	C	D	E
1	65				
2	65				
3	67				
4	68				
5	70				
6	72				
7	72				
8	73				
9	78				
10	80				
11					
12	P_{20}	66,6	=PERCENTIL.INC(A1:A10;0,2)		

2) Cálculo del percentil de orden 33 se obtiene:

$$P_k = X_{\left[\frac{n\cdot k+50}{100}\right]} = X_{\left[\frac{n\cdot 33+50}{100}\right]} = X_{\left[\frac{10\cdot 33+50}{100}\right]} = X_{\left[\frac{380}{100}\right]} = X_{3,8} = \frac{x_3 + x_4}{2} = \frac{67+68}{2} = 67,5 = 68$$

En Excel:

	A	B	C	D	E
1	65				
2	65				
3	67				
4	68				
5	70				
6	72				
7	72				
8	73				
9	78				
10	80				
11					
12	P_{33}	67,97	=PERCENTIL.INC(A1:A10;0,33)		

b) Para Datos Agrupados en Tablas de Frecuencia
Se emplea la misma ecuación utilizada en el cálculo de los percentiles para datos sin agrupar.

c) Para Datos Agrupados en Intervalos
Se emplea la ecuación:

$$\mathrm{P_k} = \mathrm{Li_P} + \left(\frac{\frac{\mathrm{nk}}{100} - \mathrm{Fa}}{\mathrm{f_P}}\right)\cdot \mathrm{c}$$

Donde:
Li_p =Límite inferior del intervalo de clase del percentil.
n = número total de datos.
F_a = Frecuencia acumulada del intervalo de clase que antecede al intervalo de clase del percentil.
f_p = Frecuencia absoluta del intervalo de clase del percentil.
c = Ancho del intervalo de clase del percentil.

MODA

a) Para Datos No Agrupados
Se observa el dato que tiene mayor frecuencia

Ejemplo ilustrativo N° 1: Determinar la moda del conjunto de datos 2, 4, 6, 8, 8 y 10

Solución: Mo = 8, porque es el dato que ocurre con mayor frecuencia. A este conjunto de datos se le llama unimodal

En Excel:

	A	B	C	D
1	2			
2	4			
3	6			
4	8			
5	8			
6	10			
7	Mo	8	=MODA.UNO(A1:A6)	

Ejemplo ilustrativo N° 2: Determinar la moda del conjunto de datos: 2, 4, 6, 8 y 10
Solución: Este conjunto de datos no tiene moda, porque todos los datos tienen la misma frecuencia.

Ejemplo ilustrativo N° 3: Determinar la moda del conjunto de datos: 8, 4, 6, 6, 8, 2 y 10
Solución: Este conjunto de datos tiene dos modas, 8 y 6, y se llama bimodal.

En Excel:
Se inserta la función MODA.VARIOS, la cual debe especificarse como fórmula de matriz, para lo cual se selecciona las celdas donde aparecerá la respuesta (B9:B10). Luego se inserta la función MODA .VARIOS, se selecciona las celdas respectivas (A1:A7). Finalmente, se presiona Ctrl+Blog Mayús+Enter.

B9 {=MODA.VARIOS(A1:A7)}

	A	B	C
1	8		
2	4		
3	6		
4	6		
5	8		
6	2		
7	10		
8	Mo	8	
9		6	

b) Para Datos Agrupados en Tablas de Frecuencia
Se observa el dato tiene mayor frecuencia
Ejemplo ilustrativo: Calcule la moda o modas (si las hay) de los siguientes datos:

x	f
2	1
4	2
6	3
8	1
10	1

Solución:
Se observa que el dato con mayor frecuencia es 6, por lo tanto Mo = 6

c) Para Datos Agrupados en Intervalos

Se halla en el intervalo o clase que tenga la frecuencia más alta, llamada intervalo o clase modal. Se emplea la siguiente ecuación:

$$Mo = L_{Mo} + \left(\frac{D_a}{D_a + D_b}\right) \cdot c$$

$L_{M0=}$ Límite inferior de la clase modal.
D_a = Diferencia entre la frecuencia absoluta de la clase modal y la clase que la antecede.
D_b = Diferencia entre la frecuencia absoluta de la clase modal y la clase que le sigue.
c = ancho de la clase modal.

Ejemplo ilustrativo: Calcule la moda o modas (si las hay) de los siguientes datos:

Intervalo o Clase	f
10-19	3
20-29	7
30-39	15
40-49	12
50-59	8

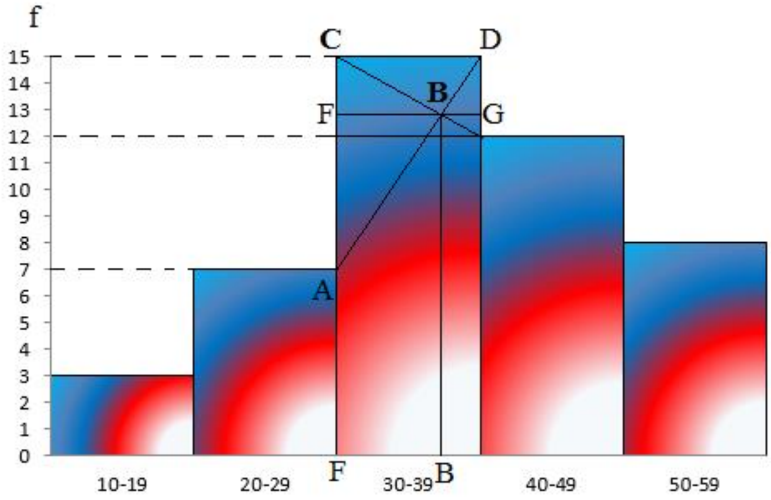
Solución: Se observa que la clase modal es 30-39, ya que es el intervalo con la mayor frecuencia.
Aplicando la ecuación

$$Mo = L_{Mo} + \left(\frac{D_a}{D_a + D_b}\right) \cdot c$$

Se tiene:

$$Mo = 30 + \left(\frac{15-7}{(15-7)+(15-12)}\right) \cdot 10 = 30 + \left(\frac{8}{8+3}\right) \cdot 10 = 30 + \frac{80}{11} = 37{,}27$$

Gráficamente empleando un histograma se calcula la moda de la siguiente manera:



La clase modal es 30-39, ya que es el intervalo con la mayor frecuencia
Observando el histograma se tiene que Mo = 30 + FB
Los triángulos ABC y EBD son semejantes, por lo que se cumple:

$$\frac{FB}{AC} = \frac{BG}{DE}$$

Donde:

AC = Diferencia entre la frecuencia absoluta de la clase modal y la clase que la antecede.
BG es igual al ancho del intervalo 30-39 menos FB.
DE = Diferencia entre la frecuencia absoluta de la clase modal y la clase que le sigue.

Remplazando valores y despejando FB se tiene:

$$\frac{FB}{15-7}=\frac{10-FB}{15-12}\Longrightarrow\frac{FB}{8}=\frac{10-FB}{3}\Longrightarrow 3FB=8(10-FB)\Longrightarrow 3FB=80-8FB$$

$$3FB+8FB=80\Longrightarrow 11FB=80\Longrightarrow FB=\frac{80}{11}=7{,}27$$

Por lo tanto Mo = 30 + FB = 30+7,27 = 37,27

MEDIDAS DE DISPERSIÓN

DESVIACIÓN MEDIA O DESVIACIÓN PROMEDIO

La desviación media o desviación promedio es la media aritmética de los valores absolutos de las desviaciones respecto a la media aritmética.

a) Para Datos No Agrupados

Se emplea la fórmula:

$$\mathrm{DM}=\frac{\sum|\mathrm{x}-\bar{\mathrm{x}}|}{\mathrm{n}}$$

Ejemplo ilustrativo: Calcular la desviación media de la distribución: 3, 8, 8, 8, 9, 9, 9, 18

Solución: Se calcula la media aritmética.

$$\bar{x}=\frac{\sum x_i}{n}=\frac{3+8+8+8+9+9+9+18}{8}=9$$

Se calcula la desviación media.

$$\mathrm{DM}=\frac{\sum|\mathrm{x}-\bar{\mathrm{x}}|}{\mathrm{n}}$$

$$\mathrm{DM}=\frac{|3-9|+|8-9|+|8-9|+|8-9|+|9-9|+|9-9|+|9-9|+|18-9|}{8}$$

$$\mathrm{DM}=\frac{6+1+1+1+0+0+0+9}{8}=\frac{18}{8}=2{,}25$$

Empleando Excel:

	A	B	C
1	3		
2	8		
3	8		
4	8		
5	9		
6	9		
7	9		
8	18		
9	DM	2,25	=DESVPROM(A1:A8)

b) Para Datos Agrupados en Tablas de Frecuencia

Se emplea la fórmula:

$$DM=\frac{\sum f|x-\bar{x}|}{n}$$

Ejemplo ilustrativo: Calcular la desviación media en base a la siguiente tabla sobre las calificaciones de un estudiante en 12 asignaturas evaluadas sobre 10**.**

Calificación	Cantidad de asignaturas
6	4
7	2
8	3
9	2
10	1
Total	12

Solución:

Se calcula la media aritmética.

$$\bar{x} = \frac{\sum fx}{n} = \frac{4\cdot 6 + 2\cdot 7 + 3\cdot 8 + 2\cdot 9 + 1\cdot 10}{12} = \frac{24+14+24+18+10}{12} = \frac{90}{12} = 7{,}5$$

Se llena la siguiente tabla:

x	f	$\lvert x-\bar{x}\rvert$	$f\lvert x-\bar{x}\rvert$
6	4	1,5	6
7	2	0,5	1
8	3	0,5	1,5
9	2	1,5	3
10	1	2,5	2,5
Total	12		14

Se emplea la ecuación de la desviación media.

$$DM = \frac{\sum f\lvert x-\bar{x}\rvert}{n} = \frac{14}{12} = 1{,}167$$

c) Para Datos Agrupados en Intervalos

Se emplea la fórmula:

$$DM = \frac{\sum f\lvert xm-\bar{x}\rvert}{n}$$

Donde xm es la marca de clase**.**

Ejemplo ilustrativo: Calcular la desviación media de un curso de 40 estudiantes en la asignatura de Estadística en base a la siguiente tabla:

Calificación	Cantidad de estudiantes
2-4	6
4-6	8
6-8	16
8-10	10
Total	40

Solución: Para calcular la media aritmética se llena la siguiente tabla:

Intervalo	f	xm	$f\cdot xm$
2-4	6	3	18
4-6	8	5	40
6-8	16	7	112
8-10	10	9	90
Total	40		260

Calculando la media aritmética se obtiene:

$$\bar{x} = \frac{\sum f \cdot xm}{n} = \frac{260}{40} = 6,5$$

Para calcular la desviación media se llena la siguiente tabla:

Intervalo	f	xm	$\lvert xm - \bar{x} \rvert$	$f\lvert xm - \bar{x} \rvert$
2-4	6	3	3,5	21
4-6	8	5	2,5	12
6-8	16	7	0,5	8
8-10	10	9	2,5	25
Total	40			66

$$DM = \frac{66}{40} = 1,65$$

VARIANZA Y DESVIACIÓN ESTÁNDAR

El teorema de Chebyshev, de autoría del matemático ruso Pafnuty Lvovich Chebyshev, establece que para todo conjunto de datos, por lo menos $1 - 1/k^2$ de las observaciones están dentro de k desviaciones estándar de la media, en donde k es cualquier número mayor que 1. Este teorema se expresa de la siguiente manera:

$$1 - \frac{1}{k^2}$$

Así por ejemplo, si se forma una distribución de datos con k =3 desviaciones estándar por debajo de la media hasta 3 desviaciones estándar por encima de la media, entonces por lo menos

$$1 - \frac{1}{3^2} = 1 - \frac{1}{9} = \frac{9-1}{9} = \frac{8}{9} = 0,8889 = 88,89\%$$

Interpretación: El 88,89% de todas las observaciones estarán dentro ± 3 desviaciones de la media.

a) Para Datos No Agrupados

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$$

Donde:

$x_i = observaciones\ individuales\ de\ la\ población$

$\mu =$ media aritmética poblacional

$N = número\ de\ observaciones\ de\ la\ población$

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

Donde:

$x_i = observaciones\ individuales\ de\ la\ muestra$

$\bar{x} = media\ aritmética\ de\ la\ muestra$

$n = número\ de\ obsevaciones\ de\ la\ muestra$

La desviación estándar de una muestra se calculó con:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

Ejemplo ilustrativo: Considere que los siguientes datos corresponden al sueldo de una población: \$350, \$400, \$500, \$700 y \$1000

1) Calcular la desviación estándar.

2) ¿Cuál es el intervalo que está dentro de k = 2 desviaciones estándar de la media?. ¿Qué porcentaje de las observaciones se encuentran dentro de ese intervalo?

Solución:

1) Para la calcular la desviación estándar se sigue los siguientes pasos:

a) Se calcula la media aritmética.

$$\mu = \frac{\sum x_i}{N} = \frac{350 + 400 + 500 + 700 + 1000}{5} = \frac{2950}{5} = \$\ 590$$

b) Se aplica la respectiva fórmula para calcular la varianza

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$$

$$\sigma^2 = \frac{(350-590)^2 + (400-590)^2 + (500-590)^2 + (700-590)^2 + (1000-590)^2}{5}$$

$$\sigma^2 = \frac{57600 + 36100 + 8100 + 12100 + 168100}{5} = \frac{282000}{5} = \$^2\ 56400$$

c) Se calcula la desviación estándar.

$$\sigma = \sqrt{\sigma^2} = \sqrt{\$^2 56400} = \$237{,}4868$$

En Excel:

	A	B	C	D
1	350			
2	400			
3	500			
4	700			
5	1000			
6	σ	237,49	=DESVESTP(A1:A5)	

2) Cálculo del intervalo de k = 2 desviaciones estándar de la media.

Se transportan 2 desviaciones estándar (2 x \$ 237,4868) = \$ 474,97 por encima y por debajo de la media $\mu = \$590$

Por lo tanto se tiene un intervalo desde \$ 590 - \$474,97 = \$ 115,03 hasta \$ 590 + \$474,97 = \$ 1064,97

Aplicando el Teorema de Chebyshev

$$1 - \frac{1}{k^2} = 1 - \frac{1}{2^2} = 1 - \frac{1}{4} = \frac{4-1}{4} = \frac{3}{4} = 0{,}75 = 75\%$$

Interpretación: Se puede afirmar de que por lo menos el 75% los sueldos están entre \$ 115,03 y \$ 1064,97

b) Para Datos Agrupados en Tablas de Frecuencia

La varianza para una población se calcula con:

$$\sigma^2 = \frac{\sum f(x_i - \mu)^2}{N}$$

Donde:

f = frecuencia absoluta.

La desviación estándar poblacional se calcula con:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum f(x_i - \mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2 = \frac{\sum f(x_i - \bar{x})^2}{n-1}$$

La desviación estándar de una muestra se calcula con:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum f(x_i - \bar{x})^2}{n-1}}$$

Ejemplo ilustrativo: Calcular la desviación estándar de los siguientes datos correspondientes a una muestra.

Calificaciones	f
4	3
5	6
6	4
7	13
8	7
10	6
Total	39

Solución: Se llena la siguiente tabla:

Calificaciones	f	fx
4	3	12
5	6	30
6	4	24
7	13	91
8	7	56
10	6	60
Total	39	273

Se calcula la media aritmética.

$$\bar{x} = \frac{\sum f \cdot x_i}{n} = \frac{273}{39} = 7$$

Se llena la siguiente tabla:

Calificaciones	f	fx_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f(x_i - \bar{x})^2$
4	3	12	-3	9	27
5	6	30	-2	4	24
6	4	24	-1	1	4
7	13	91	0	0	0
8	7	56	1	1	7
10	6	60	3	9	54
Total	39	273			116

Se calcula la desviación estándar

$$s=\sqrt{\frac{116}{39-1}}=\sqrt{\frac{116}{38}}=\sqrt{3{,}0526}=1{,}747$$

c) Para Datos Agrupados en Intervalos

La varianza para una población se calcula con:

$$\sigma^2=\frac{\sum f(xm_i-\mu)^2}{N}$$

Donde:

$f=frecuencia\ absoluta\ ;\ xm=marca\ de\ clase$

La desviación estándar poblacional se calcula con:

$$\sigma=\sqrt{\sigma^2}=\sqrt{\frac{\sum f(xm_i-\mu)^2}{N}}$$

La varianza de la muestra se calcula con:

$$s^2=\frac{\sum f(xm_i-\bar{x})^2}{n-1}$$

La desviación estándar de una muestra se calcula con:

$$s=\sqrt{s^2}=\sqrt{\frac{\sum f(xm_i-\bar{x})^2}{n-1}}$$

Ejemplo ilustrativo: Calcular la desviación estándar de los siguientes datos correspondientes a una muestra.

Intervalo	f
60-65	5
65-70	20
70-75	40
80-85	27
85-90	8
Total	100

Solución: Se llena la siguiente tabla:

Intervalo	f	xm	$f\cdot xm$
60-65	5	62,5	312,5
65-70	20	67,5	1350
70-75	40	72,5	2900
80-85	27	82,5	2227,5
85-90	8	87,5	700
Total	100		7490

Se calcula la media aritmética.

$$\bar{x}=\frac{\sum f\cdot xm}{n}=\frac{7490}{100}=74{,}9$$

Se llena la siguiente tabla:

Intervalo	f	xm	$f\cdot xm$	$xm_i-\bar{x}$	$(xm_i-\bar{x})^2$	$f(xm_i-\bar{x})^2$
60-65	5	62,5	312,5	-12,4	153,76	768,8
65-70	20	67,5	1350	-7,4	54,76	1095,2
70-75	40	72,5	2900	-2,4	5,76	230,4
80-85	27	82,5	2227,5	7,6	57,76	1559,52
85-90	8	87,5	700	12,6	158,76	1270,08
Total	100		7490			4924

Se calcula la desviación estándar.

$$s=\sqrt{\frac{4924}{100-1}}=\sqrt{\frac{4924}{99}}=\sqrt{49{,}737}=7{,}052$$

OTRAS MEDIDAS DE DISPERSIÓN

a) RANGO, AMPLITUD TOTAL O RECORRIDO

Dada una serie de valores $x_1, x_2, \ldots . x_n$, su recorrido es la diferencia aritmética entre el máximo y el mínimo de estos valores.

$$Re = xmáx - xmín$$

Ejemplo ilustrativo: Calcula el rango de las siguientes distribuciones:
1) 4, 6, 8, 10, 12, 14, 16
2) 5, 10, 13, 13, 14, 15, 17

Solución:
$Re_1 = 16 - 4 = 12$; $Re_2 = 17 - 5 = 12$

En Excel:

	A	B	C	D	E	F	G
1	4				5		
2	6				10		
3	8				13		
4	10				13		
5	12				14		
6	14				15		
7	16				17		
8							
9	Recorrido	12	=MAX(A1:A7)-MIN(A1:A7)		Recorrido	12	=MAX(E1:E7)-MIN(E1:E7)

Ambas series tienen rango 12, pero están desigualmente distribuidas, pues mientras la primera se distribuye uniformemente a lo largo de todo el recorrido, la segunda tiene una mayor concentración en el centro.

La amplitud es una medida de dispersión cuya ventaja es la facilidad con que se calcula. Tiene en cambio las siguientes desventajas:

- En su cálculo sólo intervienen dos elementos del conjunto.

- Al aumentar el número de observaciones, puede esperarse que aumente la variabilidad. Puesto que la amplitud no tiene en cuenta el tamaño del conjunto, no es una medida adecuada para comparar la variabilidad de dos grupos de observaciones, a menos que éstos sean del mismo tamaño.

Nota: Cuando los datos están agrupados en intervalos se calcula la amplitud sacando la diferencia entre la marca de clase mayor y la marca de clase menor.

b) AMPLITUD INTERCUARTÍLICA

La amplitud intercuartílica es la distancia entre el tercer cuartil Q_3 y el primer cuartil Q_1.
Amplitud intercuartílica = tercer cuartil - primer cuartil = $Q_3 - Q_1$

c) RANGO SEMI-INTERCUARTIL O DESVIACIÓN CUARTÍLICA

La desviación cuartílica es la mitad de la distancia entre el tercer cuartil y el primero

$$DQ = \frac{Q_3 - Q_1}{2}$$

Ejemplo ilustrativo: Si el tercer cuartil = 24 y el primer cuartil = 10. ¿Cuál es la desviación cuartílica?

Solución: La amplitud intercuartílica es 24 - 10 = 14. Por lo tanto, la desviación cuartílica es:

$$DQ = \frac{14}{2}$$

d) RANGO PERCENTIL O AMPLITUD CUARTÍLICA

Cada conjunto de datos tiene 99 percentiles, que dividen el conjunto en 100 partes iguales.
La amplitud cuartílica es la distancia entre dos percentiles establecidos.

$$Rango\ percentil\ = P_{90} - P_{10}$$

DISPERSIÓN RELATIVA O COEFICIENTE DE VARIACIÓN

El Coeficiente de variación (*CV*) es una medida de la dispersión relativa de un conjunto de datos, que se obtiene dividiendo la desviación estándar del conjunto entre su media aritmética y se expresa generalmente en términos porcentuales. Para una población se emplea la siguiente fórmula:

$$CV = \frac{\sigma}{\mu} \cdot 100\%$$

Donde:

$CV = Coeficiente\ de\ variación;\ \sigma = desviación\ estándar\ de\ la\ población$

$\mu = media\ aritmética\ de\ la\ población$

Para una muestra se emplea la siguiente fórmula:

$$CV = \frac{s}{\bar{x}} \cdot 100\%$$

Donde:

$CV = Coeficiente\ de\ variación; s = desviación\ estándar\ de\ la\ muestra$

$\bar{x} = media\ aritmética\ de\ la\ muestra$

Ejemplo ilustrativo: Mathías, un estudiante universitario, tiene las siguientes calificaciones en las 10 asignaturas que recibe en su carrera: 8, 7, 10, 9, 8, 7, 8, 10, 9 y 10. Emily, una compañero de Mathías, tiene las siguientes calificaciones: 8, 9, 8, 7, 8, 9, 10, 7, 8 y 10. ¿Cuál estudiante tiene menor variabilidad en sus calificaciones?

Solución: Como se está tomando en cuenta todas las asignaturas, se debe calcular el coeficiente de variación poblacional.

En Excel:

	A	B	C	D	E	F
1	Mathias				Emily	
2	8				7	
3	7				7	
4	10				8	
5	9				8	
6	8				8	
7	7				8	
8	8				9	
9	10				9	
10	9				10	
11	10				10	
12	μ	8,60	=PROMEDIO(A2:A11)	8,40		=PROMEDIO(E2:E11)
13	σ	1,11	=DESVESTP(A2:A11)	1,02		=DESVESTP(E2:E11)
14	CV	12,95	=(B13/B12)*100	12,14		=(E13/E12)*100

Agrupando los datos en tablas de frecuencias:
Para Mathías se obtiene:
En Excel:

	A	B	C	D	E	F
1	x_i	f	$f(x_i - \mu)^2$			
2	7	2	5,12	=B2*(A2-$B$7)^2		
3	8	3	1,08	=B3*(A3-$B$7)^2		
4	9	2	0,32	=B4*(A4-$B$7)^2		
5	10	3	5,88	=B5*(A5-$B$7)^2		
6	Total	10	12,4	=SUMA(C2:C5)		
7	μ	8,6	=SUMAPRODUCTO(A2:A5;B2:B5)/SUMA(B2:B5)			
8						
9	σ	1,114	=RAIZ(C6/SUMA(B2:B5))			
10						
11	CV	12,948	=(B9/B7)*100			

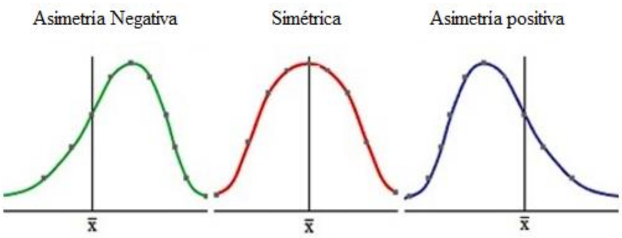
Para Emily se obtiene:
En Excel:

	A	B	C	D	E	F
1	x_i	f	$f(x_i - \mu)^2$			
2	7	2	3,92	=B2*(A2-$B$7)^2		
3	8	4	0,64	=B3*(A3-$B$7)^2		
4	9	2	0,72	=B4*(A4-$B$7)^2		
5	10	2	5,12	=B5*(A5-$B$7)^2		
6	Total	10	10,4	=SUMA(C2:C5)		
7	μ	8,4	=SUMAPRODUCTO(A2:A5;B2:B5)/SUMA(B2:B5)			
8						
9	σ	1,020	=RAIZ(C6/SUMA(B2:B5))			
10						
11	CV	12,141	=(B9/B7)*100			

Interpretación: Por lo tanto, Emily tiene menor variabilidad en sus calificaciones

ASIMETRÍA

Es una medida de forma de una distribución que permite identificar y describir la manera como los datos tiende a reunirse de acuerdo con la frecuencia con que se hallen dentro de la distribución. Permite identificar las características de la distribución de datos sin necesidad de generar el gráfico.



a) Coeficiente de Karl Pearson

$$As = \frac{3(\bar{x} - Md)}{s}$$

Donde: $\bar{x}$ = media aritmética ; Md = Mediana ; s = desviación típica o estándar.
Nota:
El Coeficiente de Pearson varía entre -3 y 3
Si As < 0 → la distribución será asimétrica negativa.
Si As = 0 → la distribución será simétrica.
Si As > 0 → la distribución será asimétrica positiva.

b) Medida de Yule Bowley o Medida Cuartílica

$$As = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$$

Donde: Q_1 = Cuartil uno; Q_2 = Cuartil dos = Mediana; Q_3 = Cuartil tres.

Nota:

La Medida de Bowley varía entre -1 y 1

Si As < 0 → la distribución será asimétrica negativa.

Si As = 0 → la distribución será simétrica.

Si As > 0 → la distribución será asimétrica positiva.

c) Medida de Fisher

Para datos sin agrupar se emplea la siguiente fórmula:

$$As = \frac{\sum(x_i - \bar{x})^3}{n\sigma^3}$$

Para datos agrupados en tablas de frecuencias se emplea la siguiente fórmula:

$$As = \frac{\sum f(x_i - \bar{x})^3}{n\sigma^3}$$

Para datos agrupados en intervalos se emplea la siguiente fórmula:

$$As = \frac{\sum f(xm - \bar{x})^3}{n\sigma^3}$$

Donde:

x_i = cada uno de los valores; n = número de datos; $\bar{x}$ = media aritmética; f = frecuencia absoluta

σ^3 = cubo de la desviación estándar poblacional; xm = marca de clase

Nota:

Si As < 0 →Indica que existe presencia de la minoría de datos en la parte izquierda de la mediaSi As = 0 → la distribución será simétrica

Si As > 0 → Indica que existe presencia de la minoría de datos en la parte derecha de la media, aunque en algunos casos no necesariamente indicará que la distribución sea asimétrica positiva

Ejemplo ilustrativo: Calcular el Coeficiente de Pearson, Medida Cuartílica y la Medida de Fisher dada la siguiente distribución: 6, 9, 9, 12, 12, 12, 15 y 17

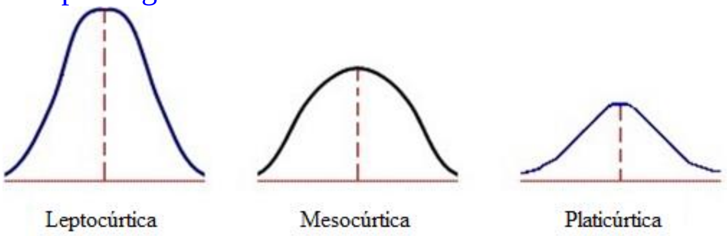
Solución:

	A	B	C	D	E	F
1	Datos	$(x_i - \bar{x})^3$				
2	6	-166,375				
3	9	-15,625				
4	9	-15,625				
5	12	0,125				
6	12	0,125				
7	12	0,125				
8	15	42,875				
9	17	166,375				
10	Total	12	=SUMA(B2:B9)			
11	n	8	=CONTAR(A2:A9)			
12	Media aritmética	11,5	=PROMEDIO(A2:A9)			
13	Desviación estándar	3,5050983	=DESVEST.M(A2:A9)			
14	Desviación poblacional	3,2787193	=DESVEST.P(A2:A9)			
15	Cuartil 1	9	=CUARTIL.INC(A2:A9;1)			
16	Cuartil 2	12	=CUARTIL.INC(A2:A9;2)			
17	Cuartil 3	13,5	=CUARTIL.INC(A2:A9;3)+0,25*(A8-A7)			
18	Coeficiente de Pearson					
19-20	$As = \frac{3(\bar{x} - Md)}{s}$	-0,4279481	=3*(B12-B16)/B13			
21	Medida de Bowley					
22-23	$As = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$	-0,3333333	=(B15+B17-2*B16)/(B17-B15)			
24	Medida de Fisher					
25-26	$As = \frac{\sum(x_i - \bar{x})^3}{n\sigma^3}$	0,0425577	=B10/(B11*B14^3)			
27	**Coeficiente de Asimetría en Excel**	0,0530788	=COEFICIENTE.ASIMETRIA(A2:A9)			

Nota: El COEFICIENTE.ASIMETRIA(A2:A9) es un valor que tiene consideraciones semejantes a la Medida de Fisher

CURTOSIS O APUNTAMIENTO

La curtosis mide el grado de agudeza o achatamiento de una distribución con relación a la distribución normal, es decir, mide cuán puntiaguda es una distribución.



a) Medida de Fisher

Para datos sin agrupar se emplea la siguiente fórmula:

$$\alpha = \frac{\sum (x_i - \bar{x})^4}{n\sigma^4}$$

Para datos agrupados en tablas de frecuencias se emplea la siguiente fórmula:

$$\alpha = \frac{\sum f(x_i - \bar{x})^4}{n\sigma^4}$$

Para datos agrupados en intervalos se emplea la siguiente fórmula:

$$\alpha = \frac{\sum f(xm - \bar{x})^4}{n\sigma^4}$$

Donde: x_i = cada uno de los valores; n = número de datos; $\bar{x}$ = media aritmética; σ^4 = Cuádruplo de la desviación estándar poblacional; f = frecuencia absoluta; xm = marca de clase

Nota: Si $\alpha < 3$ → la distribución es platicúrtica; Si $\alpha = 3$ → la distribución es normal o mesocúrtica
Si $\alpha > 3$ → la distribución es leptocúrtica

b) Medida basada en Cuartiles y Percentiles

$$\kappa = \frac{\text{Desviación cuartílica}}{\text{Amplitud cuartílica}} = \frac{\frac{Q_3 - Q_1}{2}}{P_{90} - P_{10}} = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

κ (letra griega minúscula kappa) = Coeficiente percentil de curtosis

Nota: Si $\kappa < 0{,}263$ → la distribución es platicúrtica; Si $\kappa = 0{,}263$ → la distribución es normal o mesocúrtica; Si $\kappa > 0{,}263$ → la distribución es leptocúrtica

Ejemplo ilustrativo: Determinar qué tipo de curtosis tiene la siguiente distribución: 6, 9, 9, 12, 12, 12, 15 y 17. Emplear la medida de Fisher y el coeficiente percentil de curtosis.

Solución: *En Excel:*

	A	B	C	D	E	F
1	Datos	$(x_i - \bar{x})^4$				
2	6	915,063				
3	9	39,063				
4	9	39,063				
5	12	0,063				
6	12	0,063				
7	12	0,063				
8	15	150,063				
9	17	915,063				
10	Total	2058,500	=SUMA(B2:B9)			
11	n	8	=CONTAR(A2:A9)			
12	Media aritmética	11,5	=PROMEDIO(A2:A9)			
13	Desviación poblacional	3,279	=DESVEST.P(A2:A9)			
14	Cuartil 1	9	=CUARTIL.INC(A2:A9;1)			
15	Cuartil 3	13,5	=CUARTIL.INC(A2:A9;3)+0,25*(A8-A7)			
16	$\alpha = \frac{\sum (x_i - \bar{x})^4}{n\sigma^4}$	2,226609	=B10/(B11*B13^4)			
17						
18	Perceltil 10	7,4	=PERCENTIL.INC(A2:A9;0,1)-0,25*(A3-A2)			
19	Perceltil 90	16,350	=PERCENTIL.INC(A2:A9;0,9)+0,25*(A8-A7)			
20	$\kappa = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$	0,2500	=(B15-B14)/(2*(B19-B18))			
21						
22						
23	**Curtosis en Excel**	-0,224121	=CURTOSIS(A2:A9)			
24	Valor semejante a la α	2,7758789	=B23+3			

CORRELACIÓN Y REGRESIÓN

Cuando se estudian en forma conjunta dos características (variables estadísticas) de una población o muestra, se dice que estamos analizando una variable estadística bidimensional. La correlación es el grado de relación que existe entre ambas características, y la regresión es la forma de expresar matemáticamente dicha relación.

COEFICIENTE DE CORRELACIÓN DE KARL PEARSON

Llamando también coeficiente de correlación producto-momento.

a) Para datos no agrupados se calcula aplicando la siguiente ecuación:

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

r = Coeficiente producto-momento de correlación lineal; $x = X - \bar{X}$; $y = Y - \bar{Y}$

Ejemplo ilustrativo: Con los datos sobre las temperaturas en dos días diferentes en una ciudad, determinar el tipo de correlación que existe entre ellas mediante el coeficiente de PEARSON.

X	18	17	15	16	14	12	9	15	16	14	16	18	ΣX =180
Y	13	15	14	13	9	10	8	13	12	13	10	8	ΣY= 138

Solución:

Se calcula la media aritmética

$$\bar{x} = \frac{\sum x_i}{n}$$

Para X:

$$\bar{X}_X = \frac{180}{12} = 15$$

Para Y:

$$\bar{Y}_Y = \frac{138}{12} = 11{,}5$$

Se llena la siguiente tabla:

X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	x^2	xy	y^2
18	13	3	1,5	9	4,5	2,25
17	15	2	3,5	4	7	12,25
15	14	0	2,5	0	0	6,25
16	13	1	1,5	1	1,5	2,25
14	9	-1	-2,5	1	2,5	6,25
12	10	-3	-1,5	9	4,5	2,25
9	8	-6	-3,5	36	21	12,25
15	13	0	1,5	0	0	2,25
16	12	1	0,5	1	0,5	0,25
14	13	-1	1,5	1	-1,5	2,25
16	10	1	-1,5	1	-1,5	2,25
18	8	3	-3,5	9	-10,5	12,25
180	**138**			**72**	**28**	**63**

Se aplica la fórmula:

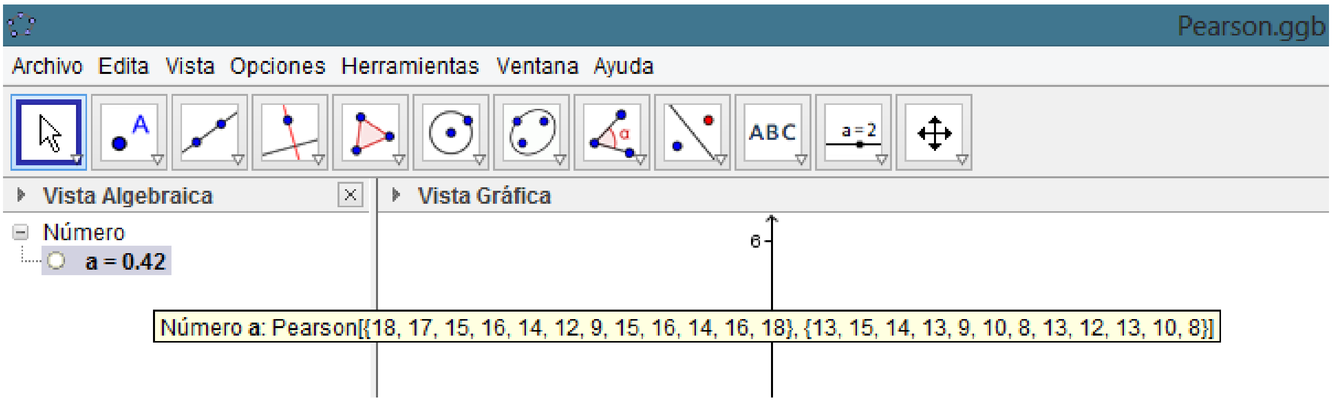
$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} = \frac{28}{\sqrt{(72)(63)}} = 0{,}416$$

Existe una correlación moderada

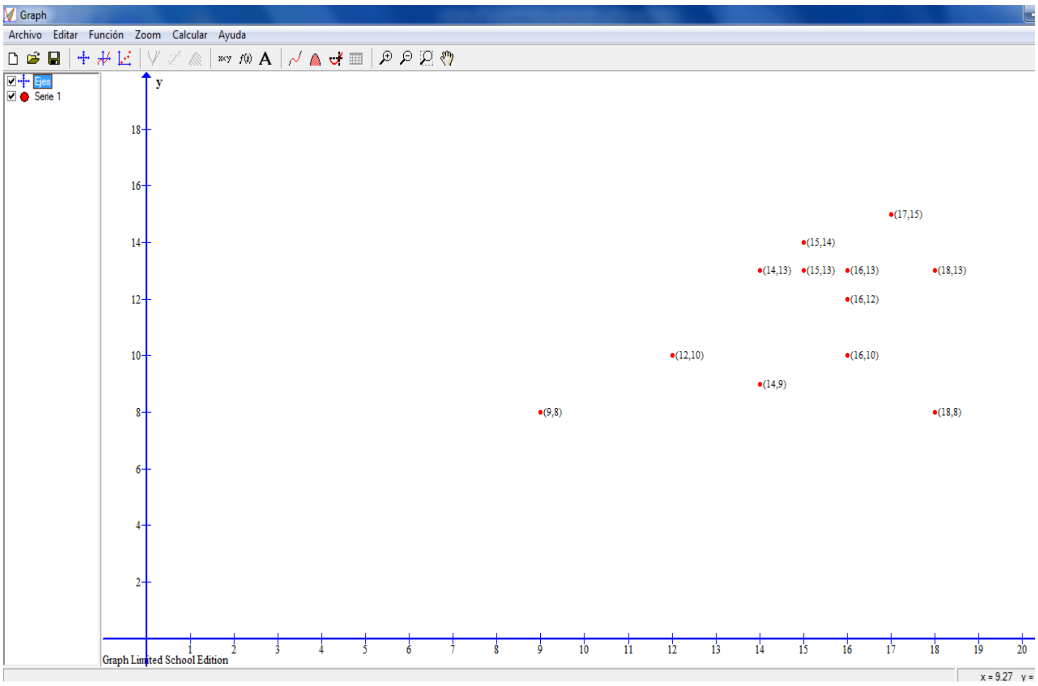
En Excel:

	A	B	C	D	E	F
1	**X**	**Y**				
2	18	13				
3	17	15				
4	15	14				
5	16	13				
6	14	9				
7	12	10				
8	9	8				
9	15	13				
10	16	12				
11	14	13				
12	16	10				
13	18	8				
14						
15	r	0,416	=COEF.DE.CORREL(A2:A13;B2:B13)			

En GeoGebra:



El Diagrama de dispersión en Graph:



b) Para datos agrupados, el coeficiente de Correlación de Pearson se calcula aplicando la siguiente fórmula:

$$r = \frac{n \cdot \sum f \cdot dx \cdot dy - (\sum fx \cdot dx)(\sum fy \cdot dy)}{\sqrt{[n \cdot \sum fx \cdot dx^2 - (\sum fx \cdot dx)^2][n \cdot \sum fy \cdot dy^2 - (\sum fy \cdot dy)^2]}}$$

Donde:
n = número de datos; f = frecuencia de celda; fx = frecuencia de la variable X; fy = frecuencia de la variable Y; dx = valores codificados o cambiados para los intervalos de la variable X, procurando que al intervalo central le corresponda dx = 0, para que se hagan más fáciles los cálculos; dy = valores codificados o cambiados para los intervalos de la variable X, procurando que al intervalo central le corresponda dy = 0, para que se hagan más fáciles los cálculos.

Ejemplo ilustrativo: Con los siguientes datos sobre los Coeficientes Intelectuales (X) y de las calificaciones en una prueba de conocimiento (Y) de 50 estudiantes:

N° de estudiante	X	Y	N° de estudiante	X	Y
1	76	28	26	88	40
2	77	24	27	88	31
3	78	18	28	88	35
4	79	41	29	88	26
5	79	43	30	89	30
6	80	45	31	89	24
7	80	34	32	90	18
8	80	18	33	90	11
9	82	40	34	90	15
10	82	35	35	91	38
11	83	30	36	92	34
12	83	21	37	92	31
13	83	22	38	93	33
14	83	23	39	93	35
15	84	25	40	93	24
16	84	11	41	94	40
17	84	15	42	96	35
18	85	31	43	97	36
19	85	35	44	98	40
20	86	26	45	99	33
21	86	30	46	100	51
22	86	24	47	101	54
23	86	16	48	101	55
24	87	20	49	102	41
25	88	36	50	102	45

1) Elaborar una tabla de dos variables
2) Calcular el coeficiente de correlación

Solución: En la *tabla de frecuencias de dos variables*, cada recuadro de esta tabla se llama una *celda* y corresponde a un par de intervalos, y el número indicado en cada celda se llama *frecuencia de celda*. Todos los totales indicados en la última fila y en la última columna se llaman *totales marginales o frecuencias marginales*, y corresponden, respectivamente, a las frecuencias de intervalo de las distribuciones de frecuencia separadas de la variable X y Y.

Para elaborar la tabla se recomienda:
- Agrupar las variables X y Y en un igual número de intervalos.
- Los intervalos de la variable X se ubican en la parte superior de manera horizontal (fila) y en orden ascendente.

- Los intervalos de la variable Y se ubican en la parte izquierda de manera vertical (columna) y en orden descendente.

Para elaborar los intervalos se procede a realizar los cálculos respectivos:

En la variable X:

Calculando el Rango se obtiene:

$R = x_{máx} - x_{mín} = 102 - 76 = 26$

Calculando el número de intervalos se obtiene:

$n_i = 1 + 3{,}32 \cdot log(n) = 1 + 3{,}32 \cdot log50 = 6{,}6 = 7$

Calculando el ancho se obtiene:

$$i = \frac{R}{ni} = \frac{26}{6{,}6} = 3{,}93 = 4$$

En la variable Y:

Calculando el Rango se obtiene:

$R = y_{máx} - y_{mín} = 55 - 11 = 44$

Calculando el número de intervalos se obtiene:

$n_i = 1 + 3{,}32 \cdot log(n) = 1 + 3{,}32 \cdot log50 = 6{,}64 = 7$

Calculando el ancho se obtiene:

$$i = \frac{R}{ni} = \frac{44}{6{,}64} = 6{,}62 = 7$$

Nota: Para la variable X se tomará un ancho de intervalo igual a 4 y para la variable Y un ancho de intervalo igual a 7. Debe quedar igual número de intervalos para cada variable, que en este ejemplo es igual a 7.

Contando las frecuencias de celda para cada par de intervalos de las variables X y Y se obtiene la siguiente tabla de frecuencias de dos variables:

		Coeficientes Intelectuales (X)							
		76-79	80-83	84-87	88-91	92-95	96-99	100-103	fy
Calificaciones (Y)	53-59							2	2
	46-52							1	1
	39-45	2	2		1	1	1	2	9
	32-38		2	1	3	3	3		12
	25-31	1	1	4	3	1			10
	18-24	2	4	2	2	1			11
	11-17			3	2				5
	fx	5	9	10	11	6	4	5	50

Interpretación:

- El número 2 es la frecuencia de la celda correspondiente al par de intervalos 76-79 en Coeficiente Intelectual y 39-45 en Calificación obtenida en la prueba de conocimiento.
- El número 5 en la fila de fx es el total marginal o frecuencia marginal del intervalo 76-79 en Coeficiente Intelectual.
- El número 2 en la columna de fy es el total marginal o frecuencia marginal del intervalo 53-59 en Calificación obtenida en la prueba de conocimiento.
- El número 50 es total de frecuencias marginales y representa al número total de estudiantes.

2) Realizando los cálculos respectivos se obtiene la siguiente tabla:

		Coeficientes Intelectuales (X)										
		76-79	80-83	84-87	88-91	92-95	96-99	100-103				
	dx / dy	-3	-2	-1	0	1	2	3	fy	$fy \cdot dy$	$fy \cdot dy^2$	$f \cdot dx \cdot dy$
Calificaciones (Y) 53-59	3							2 / 18	2	6	18	18
46-52	2							1 / 6	1	2	4	6
39-45	1	2 / -6	2 / -4		1 / 0	1 / 1	1 / 2	2 / 6	9	9	9	-1
32-38	0		2 / 0	1 / 0	3 / 0	3 / 0	3 / 0		12	0	0	0
25-31	-1	1 / 3	1 / 2	4 / 4	3 / 0	1 / -1			10	-10	10	8
18-24	-2	2 / 12	4 / 16	2 / 4	2 / 0	1 / -2			11	-22	44	30
11-17	-3			3 / 9	2 / 0				5	-15	45	9
fx		5	9	10	11	6	4	5	50	-30	130	70
$fx \cdot dx$		-15	-18	-10	0	6	8	15	-14			
$fx \cdot dx^2$		45	36	10	0	6	16	45	158			
$f \cdot dx \cdot dy$		9	14	17	0	-2	2	30	70			

Nota:
Los números de las esquinas de cada celda en la anterior tabla representan el producto $f \cdot dx \cdot dy$, así por ejemplo, para obtener el número el número -6 de los intervalos 76-79 en X y 39-45 en Y se obtiene multiplicando 2·(-3)·1 = -6. Para obtener el número 18 de los intervalos 100-103 en X y 53-59 en Y se obtiene multiplicando 2·3·3 = 18

-Los números de la última columna (18, 6, -1, 0, 8, 30 y 9) se obtienen sumando los números de las esquinas en cada fila, así por ejemplo, para obtener el número -1 se suma (-6) + (-4) + 0 + 1 +2 + 6 = -1
-Los números de la última fila (9, 14, 17, 0, -2, 2 y 30) se obtienen sumando los números de las esquinas en cada columna, así por ejemplo, para obtener el número 9 se suma (-6) + 3 + 12 = 9.
-Para obtener el número -30 de la antepenúlmina columna se obtiene sumando los resultados de $fy \cdot dy$, es decir, representa la $\sum fy \cdot dy$
-Para obtener el número -14 de la antepenúlmina fila se obtiene sumando los resultados de $fx \cdot dx$, es decir, representa la $\sum fx \cdot dx$
-Para obtener el número 130 de la penúltima columna se obtiene sumando los resultados de $fy \cdot dy^2$, es decir, representa $\sum fy \cdot dy^2$
-Para obtener el número 158 de la penúltima fila se obtiene sumando los resultados de $fx \cdot dx^2$, es decir, representa $\sum fx \cdot dx^2$
-Para obtener último número 70 de la última columna se obtiene sumando los resultados de la última columna 18 +6 +(-1) +0 + 8 + 30 +9 = 70, es decir, representa $\sum fx \cdot dx \cdot dy$
-Para obtener último número 70 de la última fila se obtiene sumando los resultados de la última fila 9 + 14 + 17 + 0 +(-2) +2 + 30 = 70 , es decir, representa $\sum fx \cdot dx \cdot dy$. Por lo tanto tiene que ser igual al último número de la última columna como comprobación que los cálculos de la tabla han sido correctos.

Observando los datos en la tabla anterior se reemplaza los valores en la ecuación del Coeficiente de Correlación de Pearson para datos agrupados, obteniéndose:

$$r=\frac{n\cdot\sum f\cdot dx\cdot dy-(\sum fx\cdot dx)(\sum fy\cdot dy)}{\sqrt{[n\cdot\sum fx\cdot dx^2-(\sum fx\cdot dx)^2][n\cdot\sum fy\cdot dy^2-(\sum fy\cdot dy)^2]}}$$

$$r=\frac{50\cdot 70-(-14)(-30)}{\sqrt{[50\cdot 158-(-14)^2][50\cdot 130-(-30)^2]}}=\frac{3500-420}{\sqrt{[7900-196][6500-900]}}=\frac{3080}{\sqrt{[7704][5600]}}$$

$$r=\frac{3080}{\sqrt{43142400}}=\frac{3080}{6568{,}287448}=0{,}469$$

Existe una correlación positiva moderada

COEFICIENTE DE CORRELACIÓN POR RANGOS DE SPEARMAN

Este coeficiente se emplea cuando una o ambas escalas de medidas de las variables son ordinales, es decir, cuando una o ambas escalas de medida son posiciones. Ejemplo: Orden de llegada en una carrera y peso de los atletas. Se calcula aplicando la siguiente ecuación:

$$r_s=1-\frac{6\sum d^2}{n(n^2-1)}$$

r_s = Coeficiente de correlación por rangos de Spearman;d = Diferencia entre los rangos (X menos Y)
n = número de datos

Ejemplo ilustrativo N° 1: La siguiente tabla muestra el rango u orden obtenido en la primera evaluación (X) y el rango o puesto obtenido en la segunda evaluación (Y) de 8 estudiantes universitarios en la asignatura de Estadística. Calcular el coeficiente de correlación por rangos de Spearman.

Estudiante	X	Y
Dyanita	1	3
Elizabeth	2	4
Mario	3	1
Orlando	4	5
Mathías	5	6
Josué	6	2
Emily	7	8
Monserrath	8	7

Para calcular el coeficiente de correlación por rangos de Spearman de se llena la siguiente tabla:

Estudiante	X	Y	$d=X-Y$	$d^2=(X-Y)^2$
Dyanita	1	3	-2	4
Elizabeth	2	4	-2	4
Mario	3	1	2	4
Orlando	4	5	-1	1
Mathías	5	6	-1	1
Josué	6	2	4	16
Emily	7	8	-1	1
Monserrath	8	7	1	1
				$\Sigma d^2=32$

Se aplica la fórmula:

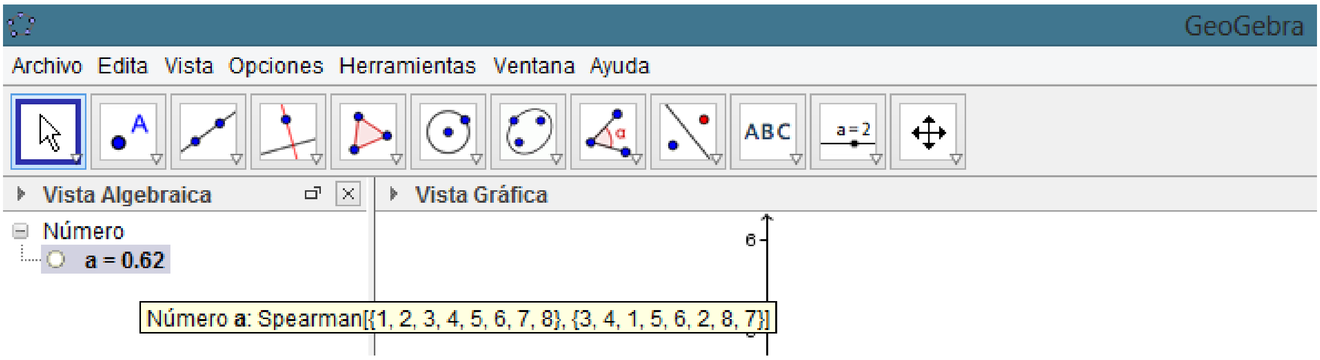
$$r_s=1-\frac{6\sum d^2}{n(n^2-1)}=1-\frac{6\cdot 32}{8(8^2-1)}=1-\frac{192}{504}=\frac{504-192}{504}=\frac{312}{504}=0{,}619$$

Por lo tanto existe una correlación positiva moderada entre la primera y segunda evaluación de los 8 estudiantes.

En Excel:

	A	B	C	D	E	F	G
1	Estudiante	X	Y				
2	Dyanita	1	3				
3	Elizabeth	2	4				
4	Mario	3	1				
5	Orlando	4	5				
6	Mathías	5	6				
7	Josué	6	2				
8	Emily	7	8				
9	Monserrat	8	7				
10				0,619	=COEF.DE.CORREL(B2:B9;C2:C9)		

En GeoGebra:



Ejemplo ilustrativo N° 2: La siguiente tabla muestra las calificaciones de 8 estudiantes universitarios en las asignaturas de Matemática y Estadística. Calcular el coeficiente de correlación por rangos de Spearman.

N°	Estudiante	Matemática	Estadística
1	Dyana	10	8
2	Elizabeth	9	6
3	Mario	8	10
4	Orlando	7	9
5	Mathías	7	8
6	Josué	6	7
7	Emily	6	6
8	Monserrath	4	9

Solución:

Para calcular el coeficiente de correlación por rangos de Spearman se procede a clasificar u ordenar los datos en rangos (X para Matemática y Y para Estadística) tomando en cuenta las siguientes observaciones:

En la asignatura de Matemática se observa:

- Dyana tiene la más alta calificación, ocupando el primer puesto, por lo que su rango es 1
- Elizabeth ocupa el segundo puesto, por lo que su rango es 2
- Mario se encuentra ubicado en el tercer lugar, por lo que su rango es 3
- Orlando y Mathías ocupan el cuarto y quinto puesto, por lo que su rango es la media aritmética de 4 y 5 que da por resultado 4,5
- Josué y Emily ocupan el sexto y séptimo lugar, por lo que su rango es la media aritmética de 6 y 7 que da por resultado 6,5
- Monserrath se encuentra ubicada en el octavo lugar, por lo que su rango es 8

En la asignatura de Estadística se observa:

- Mario tiene la más alta calificación, ocupando el primer puesto, por lo que su rango es 1
- Orlando y Monserrath ocupan el segundo y tercer puesto, por lo que su rango es la media aritmética de 2 y 3 que da por resultado 2,5

- Dyana y Mathías ocupan el cuarto y quinto puesto, por lo que su rango es la media aritmética de 4 y 5 que da por resultado 4,5
- Josué se encuentra ubicado en el sexto lugar, por lo que su rango es 6
- Elizabeth y Emily ocupan el séptimo y octavo lugar, por lo que su rango es la media aritmética de 7 y 8 que da por resultado 7,5

Los rangos X y Y se presentan en la siguiente tabla:

N°	Estudiante	Matemática	Estadística	X	Y
1	Dyana	10	8	1	4,5
2	Elizabeth	9	6	2	7,5
3	Mario	8	10	3	1
4	Orlando	7	9	4,5	2,5
5	Mathías	7	8	4,5	4,5
6	Josué	6	7	6,5	6
7	Emily	6	6	6,5	7,5
8	Monserrath	4	9	8	2,5

Calculando d, d^2 y Σd^2 se obtiene los siguientes resultados:

N°	Estudiante	Matemática	Estadística	X	Y	$d = X - Y$	$d^2 = (X - Y)^2$
1	Dyana	10	8	1	4,5	-3,5	12,25
2	Elizabeth	9	6	2	7,5	-5,5	30,25
3	Mario	8	10	3	1	2	4
4	Orlando	7	9	4,5	2,5	2	4
5	Mathías	7	8	4,5	4,5	0	0
6	Josué	6	7	6,5	6	0,5	0,25
7	Emily	6	6	6,5	7,5	-1	1
8	Monserrath	4	9	8	2,5	5,5	30,25
							$\Sigma d^2 = 82$

Aplicando la fórmula se obtiene:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \cdot 82}{8(8^2-1)} = 1 - \frac{492}{504} = \frac{504-492}{504} = \frac{12}{504} = 0{,}024$$

COEFICIENTE DE DETERMINACIÓN

Revela qué porcentaje del cambio en Y se explica por un cambio en X. Se calcula elevando al cuadrado el coeficiente de correlación.

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

$x = X - \bar{X}$; $y = Y - \bar{Y}$; r = Coeficiente de correlación de Pearson; r^2 = Coeficiente de determinación

La ecuación del coeficiente producto-momento (Coeficiente de Pearson) $r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$ puede escribirse en la forma equivalente:

$$Coeficiente\ de\ Pearson = r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}}$$

De donde coeficiente de determinación = r^2 = (Coeficiente de Pearson $)^2$

Ejemplo ilustrativo: Con los datos de la siguiente tabla sobre las temperaturas, calcular el coeficiente de determinación empleando la ecuación obtenida de la forma equivalente del coeficiente de Pearson.

X	18	17	15	16	14	12	9	15	16	14	16	18
Y	13	15	14	13	9	10	8	13	12	13	10	8

Solución:
Se calcula el coeficiente de Pearson llenando la siguiente tabla:

X	Y	XY	X^2	Y^2
18	13	234	324	169
17	15	255	289	225
15	14	210	225	196
16	13	208	256	169
14	9	126	196	81
12	10	120	144	100
9	8	72	81	64
15	13	195	225	169
16	12	192	256	144
14	13	182	196	169
16	10	160	256	100
18	8	144	324	64
$\boldsymbol{\Sigma X = 180}$	$\boldsymbol{\Sigma Y = 138}$	$\boldsymbol{\Sigma XY = 2098}$	$\boldsymbol{\Sigma X^2 = 2772}$	$\boldsymbol{\Sigma Y^2 = 1650}$

Se aplica la ecuación para calcular el coeficiente de Pearson.

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} = \frac{12\cdot 2098 - 180\cdot 138}{\sqrt{[12\cdot 2772 - (180)^2][12\cdot 1650 - (138)^2]}}$$

$$r = \frac{25176 - 24840}{\sqrt{[33264 - 32400][19800 - 19044]}} = \frac{336}{\sqrt{[864][756]}} = \frac{336}{\sqrt{653184}} = \frac{336}{808{,}198} = 0{,}4157$$

Elevando al cuadrado coeficiente de Pearson queda calculado el coeficiente de determinación.

Coeficiente de determinación = $r^2 = (0{,}4157)^2 = 0{,}1728$

Esto establece que 17,28% del cambio en Y se explica mediante un cambio en X.

Nota: El r^2 tiene significado sólo para las relaciones lineales. Dos variables pueden tener $r^2 = 0$ y sin embargo estar relacionadas en sentido curvilíneo. El valor de r^2 no se interpreta como si la variable Y fuera causado por un cambio de la variable X, ya que la correlación no significa causa.

ANÁLISIS DE REGRESIÓN

La regresión examina la relación entre dos variables, pero restringiendo una de ellas con el objeto de estudiar las variaciones de una variable cuando la otra permanece constante. En otras palabras, la regresión es un método que se emplea para predecir el valor de una variable en función de valores dados a la otra variable.

a) LA RECTA DE LOS MÍNIMOS CUADRADOS

Se llama línea de mejor ajuste y se define como la línea que hace mínima la suma de los cuadrados de las desviaciones respecto a ella de todos los puntos que corresponden a la información recogida.

La recta de los mínimos cuadrados que aproxima el conjunto de puntos $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \ldots\ldots (X_N, Y_N)$ tomando en cuenta a Y como variable dependiente tiene por ecuación

$$Y = a_0 + a_1 X$$

A esta ecuación suele llamarse recta de regresión de Y sobre X, y se usa para estimar los valores de Y para valores dados de X

Si a la recta de regresión $Y = a_0 + a_1X$ se le suma en ambos lados $\sum Y = \sum(a_0 + a_1X)$ se obtiene $\sum Y = a_0N + a_1 \sum X$

Si a la recta de regresión $Y = a_0 + a_1X$ se multiplica por X a ambos lados y luego se suma $\sum X\,Y = \sum X(a_0 + a_1X)$ se obtiene $\sum X\,Y = a_0 \sum X + a_1 \sum X^2$

Las constantes a_0 y a_1 quedan fijadas al resolver simultáneamente las ecuaciones anteriormente encontradas, es decir, al resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \Sigma Y = a_0N + a_1\Sigma X \\ \Sigma XY = a_0\Sigma X + a_1\Sigma X^2 \end{cases}$$

Que se llaman las ecuaciones normales para la recta de mínimos cuadrados.

Las constantes a_0 y a_1 de las anteriores ecuaciones también se pueden calcular empleando las siguientes fórmulas:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} \qquad a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2}$$

Otra ecuación para los mínimos cuadrados para $x = X - \bar{X}$, $y = Y - \bar{Y}$ de la recta de regresión de Y sobre X es:

$$y = \left(\frac{\sum xy}{\sum x^2}\right)x$$

La recta de los mínimos cuadrados que aproxima el conjunto de puntos $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3), \dots \dots (X_N, Y_N)$ tomando en cuenta a X como variable dependiente tiene por ecuación:

$$X = b_0 + b_1Y$$

A esta ecuación suele llamarse recta de regresión de X sobre Y, y se usa para estimar los valores de X para valores dados de Y. Las constantes b_0 y b_1 quedan fijadas al resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \Sigma X = b_0N + b_1\Sigma Y \\ \Sigma XY = b_0\Sigma Y + b_1\Sigma Y^2 \end{cases}$$

Las constantes b_0 y b_1 del sistema de ecuaciones anterior se pueden calcular empleando las siguientes fórmulas:

$$b_0 = \frac{\sum X \cdot \sum Y^2 - \sum Y \cdot \sum XY}{N \sum Y^2 - (\sum Y)^2} \qquad b_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum Y^2 - (\sum Y)^2}$$

Otra ecuación para los mínimos cuadrados para $x = X - \bar{X}$, $y = Y - \bar{Y}$ es:

$$x = \left(\frac{\sum xy}{\sum y^2}\right)y$$

El punto de intersección entre las rectas $Y = a_0 + a_1X$ con $X = b_0 + b_1Y$ se simboliza $(\bar{X}, \bar{Y})$ y se llama centroide o centro de gravedad

Ejemplo ilustrativo: Con los datos de la siguiente tabla sobre la altura en centímetros (X) y los pesos en kilogramos (Y) de una muestra de 8 estudiantes varones tomada al azar del segundo semestre de una universidad.

X	152	157	162	167	173	178	182	188
Y	56	61	67	72	70	72	83	92

1) Ajustar la recta de mínimos cuadrados para Y como variable dependiente resolviendo el sistema:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases}$$

2) Ajustar la recta de mínimos cuadrados para Y como variable dependiente empleando las fórmulas:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} \qquad a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2}$$

3) Ajustar la recta de mínimos cuadrados para Y como variable dependiente empleando la fórmula:

$$y = \left(\frac{\sum xy}{\sum x^2} \right) x$$

4) Ajustar la recta de mínimos cuadrados para X como variable dependiente resolviendo el sistema:

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases}$$

5) Calcular el punto centroide.

6) Elaborar el diagrama de dispersión. Y en el mismo diagrama graficar las dos rectas de mínimos cuadrados obtenidas en los pasos anteriores.

7) Estimar el valor de Y cuando X = 200 en el diagrama de dispersión de Y como variable dependiente.

8) Estimar el valor de X cuando Y= 100 en el diagrama de dispersión X como variable dependiente.

Solución: Se llena la siguiente tabla:

X	Y	XY	X^2	Y^2
152	56	8512	23104	3136
157	61	9577	24649	3721
162	67	10854	26244	4489
167	72	12024	27889	5184
173	70	12110	29929	4900
178	72	12816	31684	5184
182	83	15106	33124	6889
188	92	17296	35344	8464
$\Sigma X = 1359$	$\Sigma Y = 573$	$\Sigma XY = 98295$	$\Sigma X^2 = 231967$	$\Sigma Y^2 = 41967$

1) Remplazando valores en el sistema se tiene:

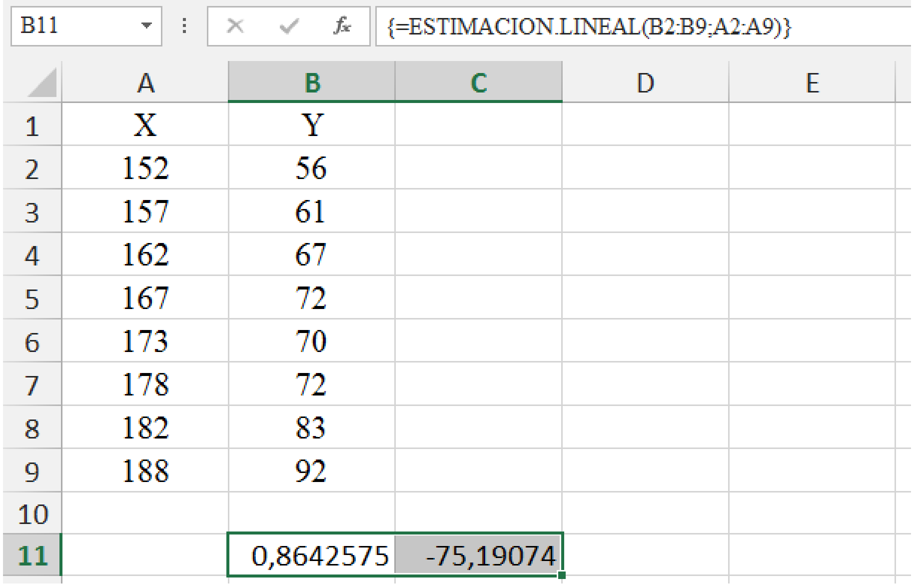
$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 \end{cases} \Rightarrow \begin{matrix} 573 = a_0 \cdot 8 + a_1 \cdot 1359 \\ 98295 = a_0 \cdot 1359 + a_1 \cdot 231967 \end{matrix} \Rightarrow \begin{cases} 8a_0 + 1359a_1 = 573 \\ 1359a_0 + 231967a_1 = 98295 \end{cases}$$

Resolviendo el sistema por determinantes (regla de Cramer) se obtiene:

$$a_0 = \frac{\Delta a_0}{\Delta} = \frac{\begin{vmatrix} 573 & 1359 \\ 98295 & 231967 \end{vmatrix}}{\begin{vmatrix} 8 & 1359 \\ 1359 & 231967 \end{vmatrix}} = \frac{573 \cdot 231967 - 98295 \cdot 1359}{8 \cdot 231967 - 1359 \cdot 1359} = \frac{-665814}{8855} = -75{,}191$$

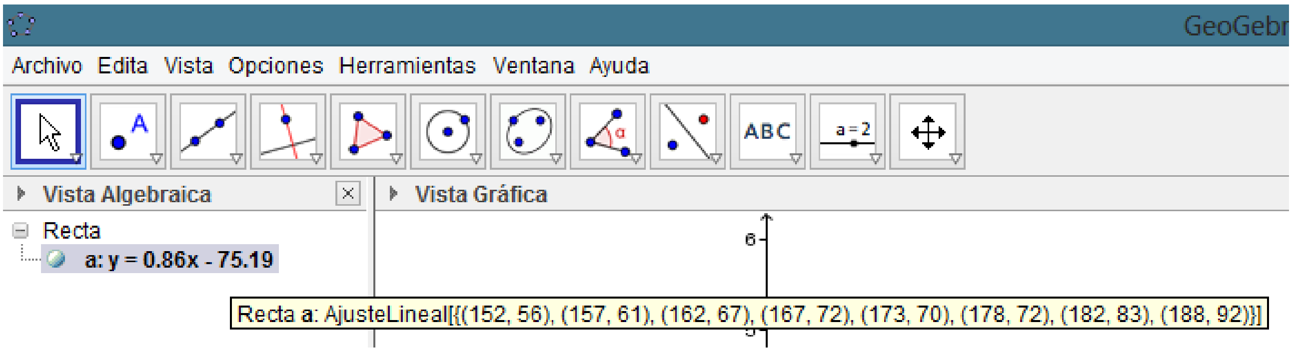
$$a_1 = \frac{\Delta a_1}{\Delta} = \frac{\begin{vmatrix} 8 & 573 \\ 1359 & 98295 \end{vmatrix}}{8855} = \frac{8 \cdot 98295 - 1359 \cdot 573}{8855} = \frac{7653}{8855} = 0{,}864$$

Para calcular los valores de a_1 y a_0 en Excel se calcula de la siguiente manera:


B11 | {=ESTIMACION.LINEAL(B2:B9;A2:A9)}

	A	B	C	D	E
1	X	Y			
2	152	56			
3	157	61			
4	162	67			
5	167	72			
6	173	70			
7	178	72			
8	182	83			
9	188	92			
10					
11		0,8642575	-75,19074		

Los cálculos en GeoGebra se muestran en la siguiente figura:



Remplazando valores en la ecuación respectiva se obtiene:
$Y = a_0 + a_1X \Rightarrow Y = -75{,}191 + 0{,}864X$

Interpretación:
- El valor $a_1 = 0{,}864$ indica que la recta tiene una pendiente positiva aumentando a razón de 0,864
- El valor de $a_0 = -75{,}191$ indica el punto en donde la recta interseca al eje Y cuanto X = 0

2) Con los datos de la tabla anterior se substituye valores en las siguientes ecuaciones:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N\sum X^2 - (\sum X)^2} = \frac{573 \cdot 231967 - 1359 \cdot 98295}{8 \cdot 231967 - (1359)^2} = \frac{-665814}{8855} - 75{,}191$$

$$a_1 = \frac{N\sum XY - \sum X \cdot \sum Y}{N\sum X^2 - (\sum X)^2} = \frac{8 \cdot 98295 - 1359 \cdot 573}{8 \cdot 231967 - (1359)^2} = \frac{7653}{8855} = 0{,}864$$

Remplazando valores en la ecuación respectiva se obtiene:
$Y = a_0 + a_1X \Rightarrow Y = -75{,}191 + 0{,}864X$

3) Se calcula las medias aritméticas de X y Y para llenar la siguiente tabla:

$$\bar{X} = \frac{1359}{8} = 169{,}875\ ;\ \bar{Y} = \frac{573}{8} = 71{,}625$$

X	Y	$x = X - \bar{X}$	$y = Y - \bar{Y}$	xy	x^2	y^2
152	56	-17,88	-15,625	279,297	319,516	244,141
157	61	-12,88	-10,625	136,797	165,766	112,891
162	67	-7,875	-4,625	36,422	62,016	21,391
167	72	-2,875	0,375	-1,078	8,266	0,141
173	70	3,125	-1,625	-5,078	9,766	2,641
178	72	8,125	0,375	3,047	66,016	0,141
182	83	12,125	11,375	137,922	147,016	129,391
188	92	18,125	20,375	369,297	328,516	415,141
$\Sigma X = 1359$	$\Sigma Y = 573$			$\Sigma xy = 956{,}625$	$\Sigma x^2 = 1106{,}875$	$\Sigma y^2 = 925{,}875$

Remplazando valores en la fórmula respectiva se obtiene:

$$y = \left(\frac{\sum xy}{\sum x^2}\right)x \Rightarrow y = \frac{956{,}625}{1106{,}875}x \Rightarrow Y - \bar{Y} = \frac{956{,}625}{1106{,}875}(X - \bar{X})$$
$$Y - 71{,}625 = \frac{956{,}625}{1106{,}875}(X - 169{,}875) \Rightarrow 1106{,}875(Y - 71{,}625) = 956{,}625(X - 169{,}875)$$
$$1106{,}875Y - 79280{,}20838 = 956{,}625X - 162510{,}4984$$
$$1106{,}875Y = 956{,}625X - 162510{,}4984 + 79280{,}20838$$
$$1106{,}875Y = 956{,}625X - 83230{,}29$$
$$Y = \frac{956{,}625X - 83230{,}29}{1106{,}875} \Rightarrow Y = \frac{956{,}625X}{1106{,}875} - \frac{83230{,}29}{1106{,}875} \Rightarrow Y = 0{,}864X - 75{,}19$$
$$Y = -75{,}19 + 0{,}864X$$

4) Remplazando valores en sistema respectivo se obtiene:

$$\begin{cases} \Sigma X = b_0 N + b_1 \Sigma Y \\ \Sigma XY = b_0 \Sigma Y + b_1 \Sigma Y^2 \end{cases} \Rightarrow \begin{matrix} 1359 = b_0 \cdot 8 + b_1 \cdot 573 \\ 98295 = b_0 \cdot 573 + b_1 \cdot 41967 \end{matrix} \Rightarrow \begin{matrix} 8b_0 + 573b_1 = 1359 \\ 573b_0 + 41967b_1 = 98295 \end{matrix}$$

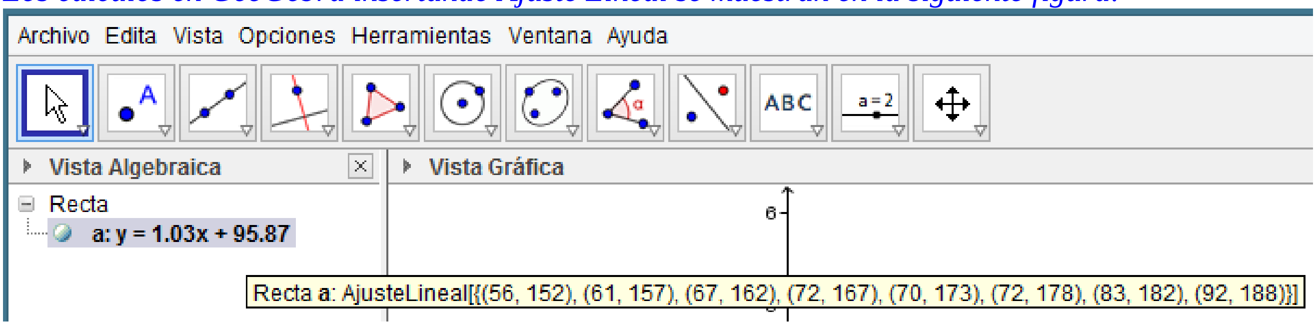
Resolviendo el sistema se obtiene:

$b_0 = 95{,}871$; $b_1 = 1{,}033$

Remplazando valores en la ecuación de la recta de mínimos cuadrados se obtiene:

$X = b_0 + b_1 Y \Longrightarrow X = 95{,}871 + 1{,}033Y$

Los cálculos en GeoGebra insertando Ajuste Lineal se muestran en la siguiente figura:



Interpretación:

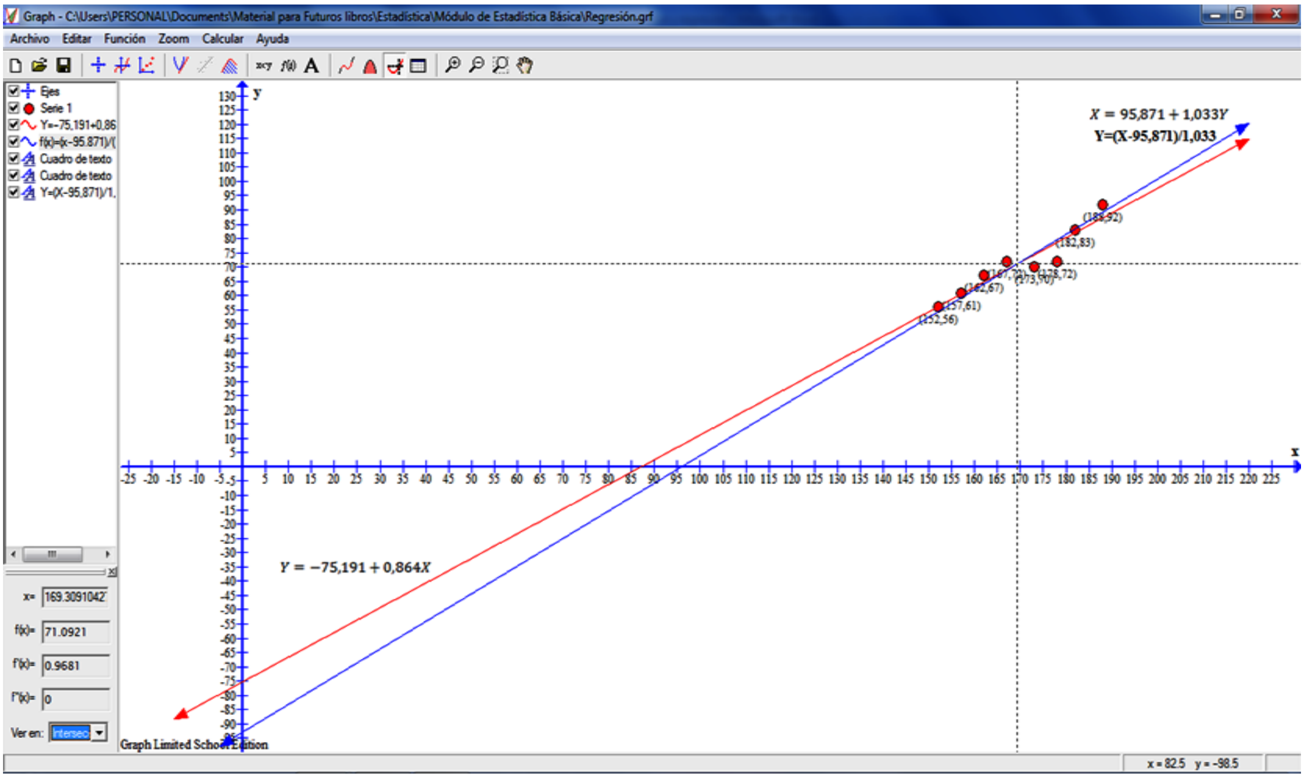
- El valor $b_1 = 1{,}033$ indica que la recta tiene una pendiente positiva aumentando a razón de 1,033
- El valor de $b_0 = 95{,}871$ indica el punto en donde la recta interseca al eje X cuanto Y = 0

5) Para calcular el centroide $(\bar{X}, \bar{Y})$ se resuelve el sistema formado por las dos rectas de los mínimos cuadrados en donde X es $\bar{X}$ y Y es $\bar{Y}$.

$$\begin{cases} Y = -75{,}191 + 0{,}864X \\ X = 95{,}871 + 1{,}033Y \end{cases}$$

Al resolver el sistema se obtiene el centroide: X = 169,3 y Y = 71,092

6) *Empleando el programa Graph se obtiene la siguiente figura:*



7) Remplazando X = 200 en la ecuación solicitada se obtiene:

$Y = -75{,}191 + 0{,}864X = -75{,}191 + 0{,}864 \cdot 200 = -75{,}191 + 172{,}8 = 97{,}609$

8) Remplazando Y = 100 en la ecuación solicitada se obtiene:

$X = 95{,}871 + 1{,}033Y = X = 95{,}871 + 1{,}033 \cdot 100 = X = 95{,}871 + 103{,}3 = 199{,}171$

b) LA PARÁBOLA DE LOS MÍNIMOS CUADRADOS

La parábola de mínimos cuadrados que aproxima el conjunto de puntos $(X_1,Y_1),(X_2,Y_2),(X_3,Y_3),\ldots(Y_N,Y_N)$ tiene ecuación dada por $Y = a_0 + a_1X + a_2X^2$, donde las constantes a_0, a_1 y a_2 se determinan al resolver simultáneamente el sistema de ecuaciones que se forma al multiplicar la ecuación $Y = a_0 + a_1X + a_2X^2$ por $1, X, Y$ sucesivamente, y sumando después.

$$\begin{cases} \Sigma Y = a_0N + a_1\Sigma X + a_2\Sigma X^2 \\ \Sigma XY = a_0\Sigma X + a_1\Sigma X^2 + a_2\Sigma X^3 \\ \Sigma X^2Y = a_0\Sigma X^2 + a_1\Sigma X^3 + a_2\Sigma X^4 \end{cases}$$

Ejemplo ilustrativo: La siguiente tabla muestra la población de un país en los años 1960-2010 en intervalos de 5 años.

Año	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Población (millones)	4,52	5,18	6,25	7,42	8,16	9,12	10,92	11,62	12,68	13,12	13,97

1) Ajustar una parábola de mínimos cuadrados de la forma $Y = a_0 + a_1X + a_2X^2$

2) Calcular los valores de tendencia para los años dados.

3) Estimar la población para los años 2015 y 2020.

4) Elaborar un diagrama de dispersión, y en el mismo diagrama graficar la parábola de los mínimos cuadrados.

Nota: Se recomienda codificar o cambiar la numeración de los años, tratando que X = 0 esté ubicado en lo posible en el centro.

Solución: Para ajustar una parábola de mínimos cuadrados se llena la siguiente tabla:

Año	X	Y	X^2	X^3	X^4	XY	X^2Y
1960	-5	4,52	25	-125	625	-22,6	113
1965	-4	5,18	16	-64	256	-20,72	82,88
1970	-3	6,25	9	-27	81	-18,75	56,25
1975	-2	7,42	4	-8	16	-14,84	29,68
1980	-1	8,16	1	-1	1	-8,16	8,16
1985	0	9,12	0	0	0	0	0
1990	1	10,92	1	1	1	10,92	10,92
1995	2	11,62	4	8	16	23,24	46,48
2000	3	12,68	9	27	81	38,04	114,12
2005	4	13,12	16	64	256	52,48	209,92
2010	5	13,97	25	125	625	69,85	349,25
Σ	0	102,96	110	0	1958	109,46	1020,66

Se remplaza valores en el sistema y se obtiene:

$$\begin{cases} \Sigma Y = a_0 N + a_1 \Sigma X + a_2 \Sigma X^2 \\ \Sigma XY = a_0 \Sigma X + a_1 \Sigma X^2 + a_2 \Sigma X^3 \\ \Sigma X^2 Y = a_0 \Sigma X^2 + a_1 \Sigma X^3 + a_2 \Sigma X^4 \end{cases}$$

$$\begin{cases} 102{,}96 = a_0 \cdot 11 + a_1 \cdot 0 + a_2 \cdot 110 \\ 109{,}46 = a_0 \cdot 0 + a_1 \cdot 110 + a_2 \cdot 0 \\ 1020{,}66 = a_0 \cdot 110 + a_1 \cdot 0 + a_2 \cdot 1958 \end{cases} \Rightarrow \begin{cases} 11a_0 + 0a_1 + 110a_2 = 102{,}96 \\ 0a_0 + 110a_1 + 0a_2 = 109{,}46 \\ 110a_0 + 0a_1 + 1958a_2 = 1020{,}66 \end{cases}$$

Resolviendo el sistema empleando determinantes (regla de Cramer) se obtiene:

$$a_o = \frac{\Delta a_0}{\Delta} = \frac{\begin{vmatrix} 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \\ 1020{,}66 & 0 & 1958 \end{vmatrix}}{\begin{vmatrix} 11 & 0 & 110 \\ 0 & 110 & 0 \\ 110 & 0 & 1958 \end{vmatrix}} = \frac{\begin{vmatrix} 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \\ 1020{,}66 & 0 & 1958 \\ 102{,}96 & 0 & 110 \\ 109{,}46 & 110 & 0 \end{vmatrix}}{\begin{vmatrix} 11 & 0 & 110 \\ 0 & 110 & 0 \\ 110 & 0 & 1958 \\ 11 & 0 & 110 \\ 0 & 110 & 0 \end{vmatrix}}$$

$$a_0 = \frac{22175524{,}8 + 0 + 0 - 12349986 - 0 - 0}{2369180 + 0 + 0 - 1331000 - 0 - 0} = \frac{9825538{,}8}{1038180} = 9{,}464$$

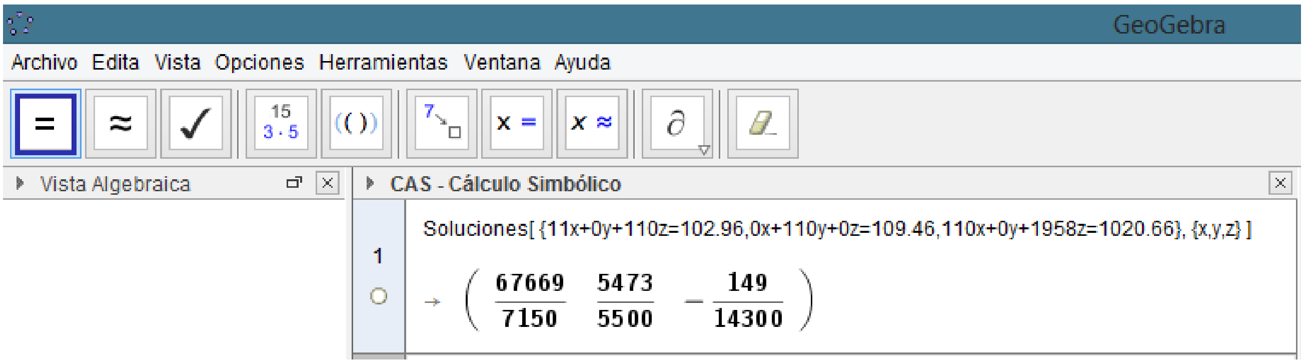
$$a_1 = \frac{\Delta a_1}{\Delta} = \frac{\begin{vmatrix} 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \\ 110 & 1020{,}66 & 1958 \end{vmatrix}}{1038180} = \frac{\begin{vmatrix} 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \\ 110 & 1020{,}66 & 1958 \\ 11 & 102{,}96 & 110 \\ 0 & 109{,}46 & 0 \end{vmatrix}}{1038180}$$

$$a_1 = \frac{23577549{,}48 + 0 + 0 - 1324466 - 0 - 0}{1038180} = \frac{2357549{,}48}{1038180} = 0{,}995$$

$$a_2 = \frac{\Delta a_2}{\Delta} = \frac{\begin{vmatrix} 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \\ 110 & 0 & 1020{,}66 \end{vmatrix}}{1038180} = \frac{\begin{vmatrix} 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \\ 110 & 0 & 1020{,}66 \\ 11 & 0 & 102{,}96 \\ 0 & 110 & 109{,}46 \end{vmatrix}}{1038180}$$

$$a_2 = \frac{1234998{,}6 + 0 + 0 - 1245816 - 0 - 0}{1038180} = \frac{-10817{,}4}{1038180} = -0{,}01$$

El sistema de ecuaciones resuelto en GeoGebra:



$$\frac{67669}{7150} = 9{,}464 \,; \frac{5473}{5500} = 0{,}995 \,; -\frac{149}{14300} = -0{,}01$$

Remplazando los valores encontrados se obtiene la ecuación de la parábola de mínimos cuadrados:
$Y = a_0 + a_1X + a_2X^2 \Rightarrow Y = 9{,}464 + 0{,}995X - 0{,}01X^2$

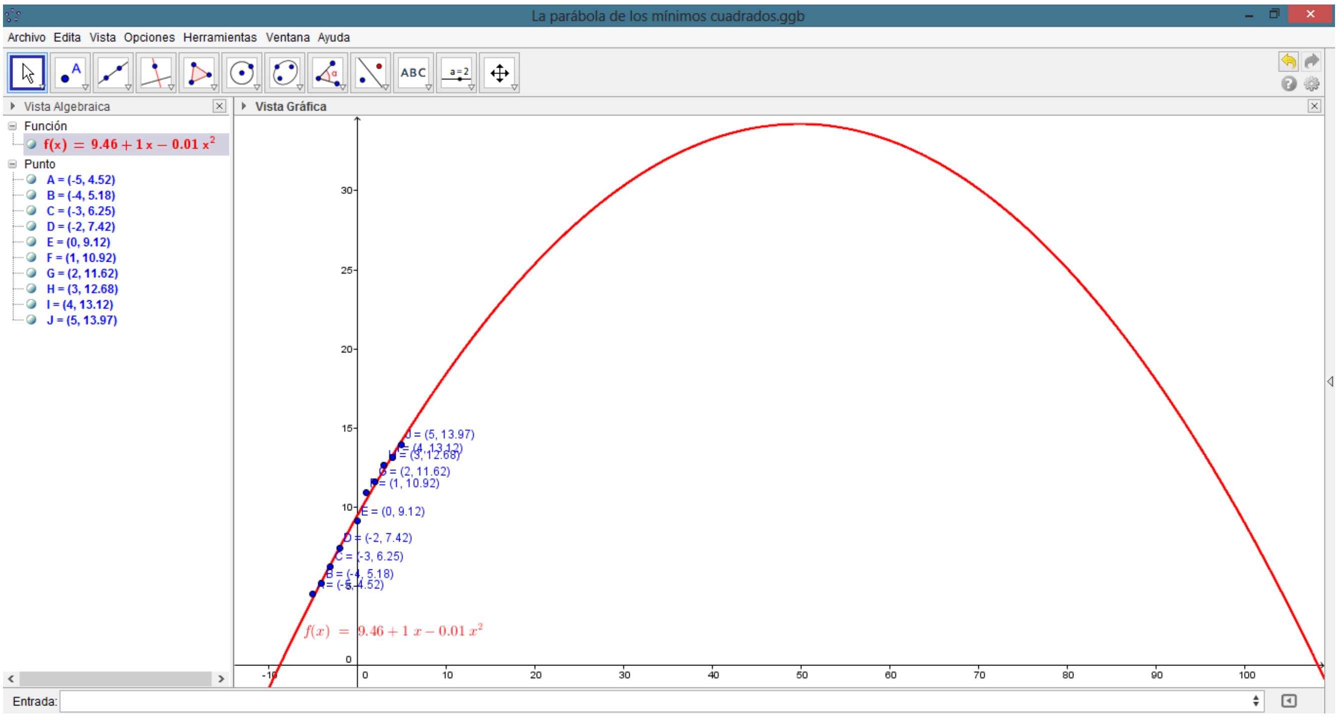
2) Los valores de tendencia se obtienen al remplazar los valores de X en la ecuación de la parábola de mínimos cuadrados, los cuales se presenta en la siguiente tabla:

Año	X	Y	Valores de tendencia $Y = 9{,}464 + 0{,}995X - 0{,}01X^2$
1960	-5	4,52	4,24
1965	-4	5,18	5,32
1970	-3	6,25	6,39
1975	-2	7,42	7,43
1980	-1	8,16	8,46
1985	0	9,12	9,46
1990	1	10,92	10,45
1995	2	11,62	11,41
2000	3	12,68	12,36
2005	4	13,12	13,28
2010	5	13,97	14,19

3) Para estimar la población de los años 2015 y 2020 se transforma estos años a X siguiendo la secuencia de la tabla anterior, siendo X = 6 para el año 2015 y X= 7 para el 2020

Entonces para el 2015 se tiene:
$Y = 9{,}464 + 0{,}995X - 0{,}01X^2 = 9{,}464 + 0{,}995(6) - 0{,}01(6)^2 = 9{,}464 + 5{,}97 - 0{,}36 = 15{,}074$
Para el 2020 se tiene:
$Y = 9{,}464 + 0{,}995X - 0{,}01X^2 = 9{,}464 + 0{,}995(7) - 0{,}01(7)^2 = 9{,}464 + 6{,}965 - 0{,}49 = 15{,}939$

4) *El diagrama de dispersión y la parábola de los mínimos cuadrados en GeoGebra:*



REGRESIÓN EXPONENCIAL

Cuando la curva de regresión de y sobre x es exponencial, es decir para cualquier x considerada, la media de la distribución está dada por la siguiente ecuación predictora:

$$Y = \alpha \cdot \beta^{X}$$

Tomando logaritmos en ambos miembros:

$$logY = log\alpha + X \cdot log\beta$$

Y se puede estimar ahora log Y y log β, y de ahí obtener α y β, aplicando *los métodos de los mínimos cuadrados*.

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + log\beta \cdot \Sigma X \\ \Sigma X \cdot \log Y = \log\alpha \cdot \Sigma X + \log\beta \cdot \Sigma X^{2} \end{cases}$$

Ejemplo ilustrativo: Las cifras siguientes son datos sobre el porcentaje de llantas radiales producidas por cierto fabricante que aún pueden usarse después de recorrer cierto número de millas:

Miles de Millas recorridas (X)	1	2	5	15	25	30	35	40
Porcentaje útil (Y)	99	95	85	55	30	24	20	15

1) Ajustar una curva exponencial aplicando el método de mínimos cuadrados.
2) Calcular la ecuación predictora.
3) Graficar la ecuación predictora.
4) Estimar qué porcentaje de las llantas radiales del fabricante durarán 50000 millas.

Solución:
1) Se llena la siguiente tabla:

X	Y	$logY$	X^2	$X \cdot logY$
1	99	1,996	1	1,996
2	95	1,978	4	3,955
5	85	1,929	25	9,647
15	55	1,740	225	26,105
25	30	1,477	625	36,928
30	24	1,380	900	41,406
35	20	1,301	1225	45,536
40	15	1,176	1600	47,044
$\Sigma X = 153$		$\Sigma logY = 12{,}97759$	$\Sigma X^2 = 4605$	$\Sigma X \cdot logY = 212{,}61769$

Remplazando valores en el sistema se obtiene:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + log\beta \cdot \Sigma X \\ \Sigma X \cdot \mathrm{logY} = \mathrm{log}\alpha \cdot \Sigma X + \mathrm{log}\beta \cdot \Sigma X^2 \end{cases}$$

$$\begin{cases} 12{,}97759 = \mathrm{log}\alpha \cdot 8 + \mathrm{log}\beta \cdot 153 \\ 212{,}61769 = \mathrm{log}\alpha \cdot 153 + \mathrm{log}\beta \cdot 4605 \end{cases} \Rightarrow \begin{cases} 8\mathrm{log}\alpha + 153\mathrm{log}\beta = 12{,}97759 \\ 153\mathrm{log}\alpha + 4605\mathrm{log}\beta = 212{,}61769 \end{cases}$$

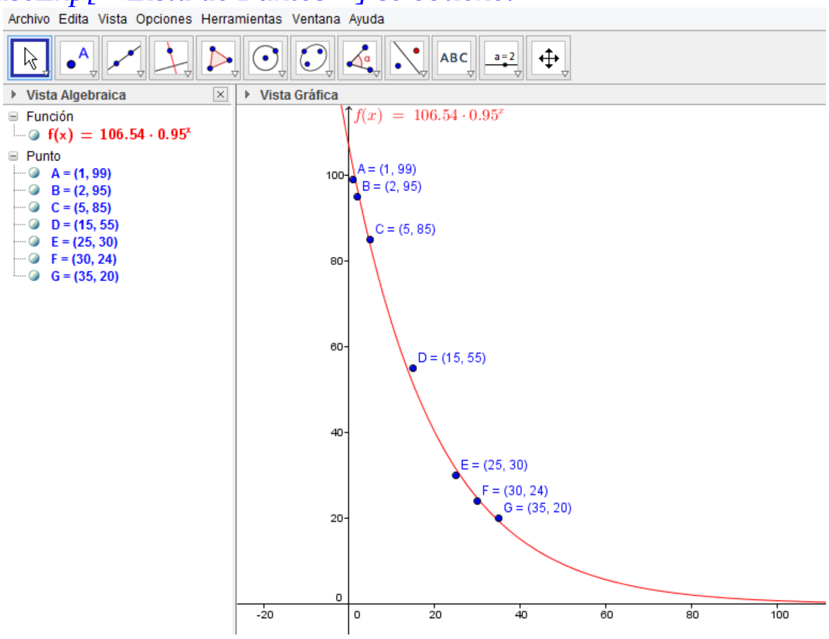
Al resolver el sistema se obtiene:
$log\alpha = 2{,}027495747; log\beta = -0{,}02119180389$

Remplazando valores se obtiene:
$logY = log\alpha + X \cdot log\beta \Longrightarrow logY = 2{,}027496 - 0{,}02119X$

Aplicando el antilogaritmo se obtiene:
$\alpha = \mathrm{anti\ log}\ \ 2{,}027495747 = 106{,}536; \ \ \beta = \mathrm{anti\ log}(-0{,}02119180389) = 0{,}952$

2) Remplazando en la ecuación predictora se obtiene:
$Y = \alpha \cdot \beta^X \Longrightarrow Y = 106{,}536 \cdot 0{,}952^X$

3) Realizando el diagrama de dispersión y los cálculos de la ecuación predictora de GeoGebra insertando AjusteBaseExp[<Lista de Puntos>] se obtiene:



4) La estimación del porcentaje de llantas radiales que durarán 50000 millas se obtiene remplazando en la ecuación predictora el valor de X = 50
$Y = 106{,}536 \cdot 0{,}952^X \Longrightarrow Y = 106{,}536 \cdot 0{,}952^{50} = 9{,}106$
Entonces el porcentaje sería de 9,106%

REGRESIÓN POTENCIAL

La regresión potencial tiene por ecuación predictora:

$$Y = \alpha \cdot X^{\beta}$$

Y la regresión recíproca es:

$$Y = \frac{1}{\alpha + \beta \cdot X}$$

Para el primer caso los valores siguen una ley potencial. Si la ecuación predictora está dada por: $Y = \alpha \cdot X^{\beta}$ tomando logaritmos en ambos miembros, queda:

$$logY = log\alpha + \beta \cdot logX$$

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + \beta \cdot \Sigma \mathrm{logX} \\ \Sigma \mathrm{logX} \cdot \mathrm{logY} = \mathrm{log\alpha} \cdot \Sigma \mathrm{logX} + \beta \cdot \Sigma (logX)^2 \end{cases}$$

Para el segundo caso, si la ecuación predictora está dada por $Y = \frac{1}{\alpha+\beta \cdot X}$, entonces invirtiendo, la misma expresión se puede escribir $\frac{1}{Y} = \frac{\alpha+\beta \cdot X}{1}$, o sea:

$$Y = \frac{1}{\alpha + \beta \cdot X} \Rightarrow \frac{1}{Y} = \alpha + \beta \cdot X$$

Donde las constantes α y β quedan fijadas al resolver simultáneamente las ecuaciones:

$$\begin{cases} \Sigma \frac{1}{Y} = \alpha \cdot N + \beta \cdot \Sigma \mathrm{X} \\ \Sigma \mathrm{X} \cdot \frac{1}{Y} = \alpha \cdot \Sigma \mathrm{X} + \beta \cdot \Sigma X^2 \end{cases}$$

Ejemplos ilustrativo N° 1: Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es el volumen (variable independiente) e Y es la presión de una masa dada de gas (variable resultante).

X	1	2	3	4	5	6	7
Y	7	30	90	170	290	450	650

1.1) Ajustar una curva exponencial aplicando el método de mínimos cuadrados.
1.2) Calcular la ecuación predictora.
1.3) Graficar la ecuación predictora.
1.4) Estimar la presión de la masa de gas de volumen 9.

Solución:

1.1) Para ajustar una curva exponencial aplicando el método de mínimos cuadrados:

X	Y	$logX$	$logY$	$logX \cdot logY$	$(logX)^2$
1	7	0,0000	0,8451	0,0000	0,0000
2	30	0,3010	1,4771	0,4447	0,0906
3	90	0,4771	1,9542	0,9324	0,2276
4	170	0,6021	2,2304	1,3429	0,3625
5	290	0,6990	2,4624	1,7211	0,4886
6	450	0,7782	2,6532	2,0646	0,6055
7	650	0,8451	2,8129	2,3772	0,7142
$\Sigma X = 28$		$\Sigma logX = 3,7024$	$\Sigma logY = 14,4354$	$\Sigma logX \cdot logY = 8,8829$	$\Sigma (logX)^2 = 2,4890$

Remplazando valores en el sistema de ecuaciones se obtiene:

$$\begin{cases} \Sigma logY = log\alpha \cdot N + \beta \cdot \Sigma \mathrm{logX} \\ \Sigma \mathrm{logX} \cdot \mathrm{logY} = \mathrm{log\alpha} \cdot \Sigma \mathrm{logX} + \beta \cdot \Sigma (logX)^2 \end{cases}$$

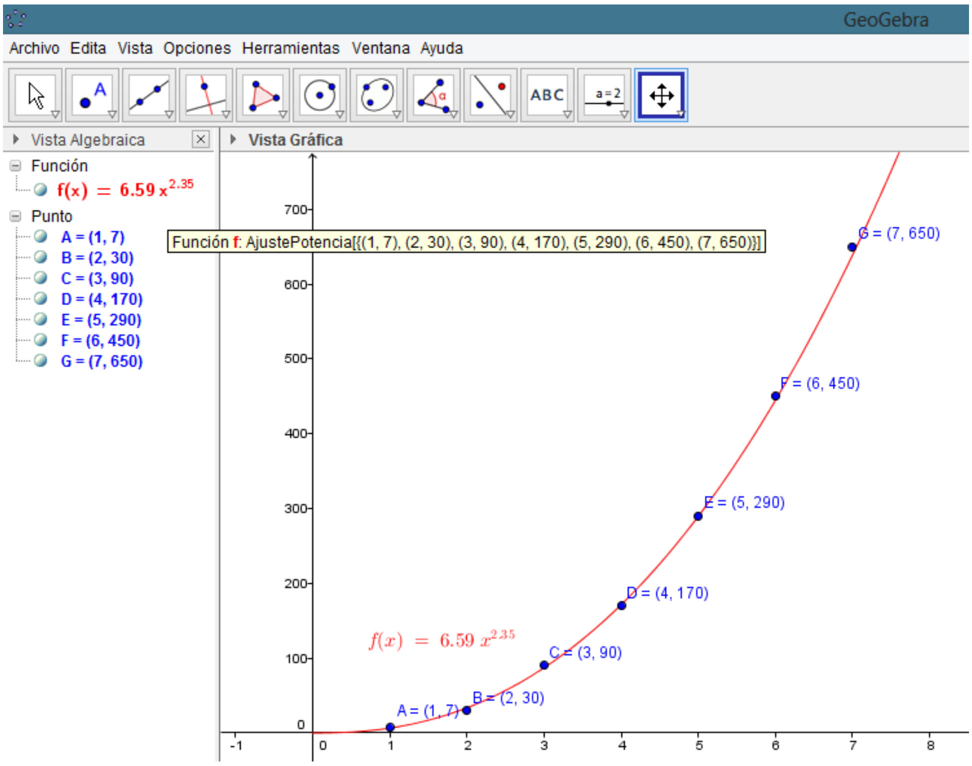
$$\begin{cases} 14,4354 = log\alpha \cdot 7 + \beta \cdot 3,7024 \\ 8,8829 = log\alpha \cdot 3,7024 + \beta \cdot 2,4890 \end{cases} \Rightarrow \begin{cases} 7log\alpha + 3,7024\beta = 14,4354 \\ 3,7024log\alpha + 2,4890\beta = 8,8829 \end{cases}$$

Al resolver el sistema se obtiene: log α = 0,819; β = 2,351

Remplazando valores en la ecuación predictora expresada en logaritmos se tiene:
$logY = log\alpha + \beta \cdot logX \Longrightarrow logY = 0,819 + 2,351 \cdot logX$

1.2) Para calcular la ecuación predictora, primero se calcula el valor de α de la siguiente manera:
$log\alpha = 0,819 \Rightarrow \alpha = antilog\ 0,819 = 6,592$
Remplazando en la ecuación predictora se obtiene:
$Y = \alpha \cdot X^{\beta} \Longrightarrow Y = 6,592 \cdot X^{2,351}$

1.3) Realizando el diagrama de dispersión y calculando la ecuación predictora en GeoGebra:



1.4) Para estimar la presión de la masa de gas de volumen 9 se reemplaza el valor X = 9 en la ecuación predictora
$Y = 6,592 \cdot X^{2,351} \Longrightarrow Y = 6,592 \cdot 9^{2,351} = 1154,63$

Ejemplo ilustrativo N° 2: Sea el siguiente conjunto de valores, las lecturas de un experimento donde X es la variable independiente e Y la variable resultante.

X	1	2	3	4	5	6	7
Y	1,4	1	0,9	0,7	0,6	0,55	0,5

2.1) Calcular las constantes α y β, aplicando el método de mínimos cuadrados.
2.2) Calcular la ecuación predictora.
2.3) Graficar la ecuación predictora.
2.4) Estimar el valor de Y para X = 9

Solución:

2.1) Para calcular las constantes α y β, aplicando el método de mínimos cuadrados se llena la siguiente tabla:

X	Y	$1/Y$	$X(1/Y)$	X^2
1	1,4	0,7143	0,7143	1
2	1	1,0000	2,0000	4
3	0,9	1,1111	3,3333	9
4	0,7	1,4286	5,7143	16
5	0,6	1,6667	8,3333	25
6	0,55	1,8182	10,9091	36
7	0,5	2,0000	14,0000	49
$\Sigma X = 28$		$\Sigma\,(1/Y) = 9{,}7388$	$\Sigma\,X(1/Y) = 45{,}0043$	$\Sigma X^2 = 140$

Remplazando valores en el siguiente sistema se obtiene:

$$\begin{cases} \Sigma\dfrac{1}{Y} = \alpha\cdot N + \beta\cdot\Sigma X \\ \Sigma X\cdot\dfrac{1}{Y} = \alpha\cdot\Sigma X + \beta\cdot\Sigma X^2 \end{cases} \Rightarrow \begin{cases} 9{,}7388 = \alpha\cdot 7 + \beta\cdot 28 \\ 45{,}0043 = \alpha\cdot 28 + \beta\cdot 140 \end{cases} \Rightarrow \begin{cases} 7\alpha + 28\beta = 9{,}7388 \\ 28\alpha + 140\beta = 45{,}0043 \end{cases}$$

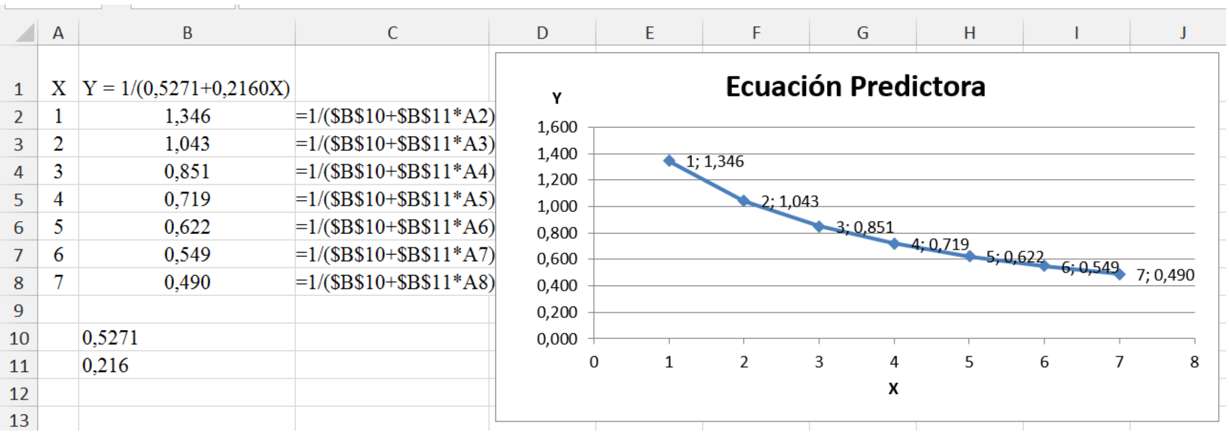
Al resolver el sistema se obtiene:

$\alpha = 0{,}5271$; $\beta = 0{,}2160$

2.2) Para calcular la ecuación predictora se remplaza los valores encontrados de α y β, y se obtiene:

$$Y = \frac{1}{\alpha + \beta\cdot X} \Rightarrow Y = \frac{1}{0{,}5271 + 0{,}2160X}$$

2.3) *La gráfica la ecuación predictora elaborada en Excel:*



2.4) Para estimar el valor de Y para X = 9 se reemplaza el valor de X en la ecuación predictora.

$$Y = \frac{1}{0{,}5271 + 0{,}2160X} \Rightarrow Y = \frac{1}{0{,}5271 + 0{,}2160\cdot 9} = 0{,}405$$

ERROR ESTÁNDAR DE ESTIMACIÓN

Es el grado de dispersión de los datos con respecto a la recta de regresión $Y = a_0 + a_1X$

El error estándar de estimación se calcula con la fórmula:

$$s_e = \sqrt{\frac{\sum(Y_i - Y_{est})^2}{N-2}}$$

Donde:
Y_i = cada valor de Y
Y_{est}= valor estimado de Y a partir de la recta de regresión
N = número de datos
Otras ecuaciones para calcular el error estándar de estimación son:

$$s_e = \sqrt{\frac{\sum Y^2 - a_0 \sum Y - a_1 \sum XY}{N-2}} \qquad s_e = \sqrt{\frac{\sum y^2 - a_1 \sum xy}{N-2}}$$

Donde:
a_0 = ordenada en el origen (punto de intersección de la recta con el eje y)
a_1 = pendiente de la recta (tangente del ángulo de inclinación de la recta)
$x = X - \bar{X}$
$y = Y - \bar{Y}$

Ejemplo ilustrativo: Calcular error estándar de estimación empleando las 3 fórmulas dadas, utilizando los datos de la tabla del ejemplo para ajustar la recta de mínimos cuadrados para Y como variable dependiente.

X	152	157	162	167	173	178	182	188
Y	56	61	67	72	70	72	83	92

Solución:
Para comenzar a resolver este ejemplo recordemos que ya se obtuvo los valores respectivos al resolver el ejemplo para ajustar la recta de mínimos cuadrados, los cuales fueron:
$\Sigma X = 1359; \Sigma Y = 573; \Sigma XY = 98295; \Sigma X^2 = 231967; \Sigma Y^2 = 41967; \Sigma xy = 956{,}625$
$\Sigma x^2 = 1106{,}875; \Sigma y^2 = 925{,}875; a_0 = -75{,}191; a_1 = 0{,}864; Y = -75{,}191 + 0{,}864X$

1) Para emplear la primera fórmula se llena la siguiente tabla:

X	Y	$Y_{est} = 75{,}191 + 0{,}86X$	Y_{est}	$(Y - Y_{est})^2$
152	56	-75,191+0,86(152)	55,529	0,222
157	61	-75,191+0,86(157)	59,829	1,371
162	67	-75,191+0,86(162)	64,129	8,243
167	72	-75,191+0,86(167)	68,429	12,752
173	70	-75,191+0,86(173)	73,589	12,881
178	72	-75,191+0,86(178)	77,889	34,680
182	83	-75,191+0,86(182)	81,329	2,792
188	92	-75,191+0,86(188)	86,489	30,371
Σ				103,312

Se remplaza valores en la primera fórmula se obtiene:

$$s_e = \sqrt{\frac{\sum (Y_i - Y_{est})^2}{N-2}} = \sqrt{\frac{103{,}312}{8-2}} = 3{,}842$$

2) Remplazando valores en la segunda fórmula se obtiene:

$$s_e = \sqrt{\frac{\sum Y^2 - a_0 \sum Y - a_1 \sum XY}{N-2}}$$

$$s_e = \sqrt{\frac{41967 - (-75{,}191)(573) - 0{,}864(98295)}{8-2}} = \sqrt{\frac{41967 + 43084{,}443 - 84926{,}88}{6}} = 4{,}556$$

3) Remplazando valores en la tercera fórmula se obtiene:

$$s_e = \sqrt{\frac{\sum y^2 - a_1 \sum xy}{N-2}} = \sqrt{\frac{925{,}875 - 0{,}864(956{,}625)}{8-2}} = \sqrt{\frac{99{,}351}{6}} = 4{,}069$$

Empleando exclusivamente Excel para calcular el error estándar de estimación se procede de la siguiente manera:

Se inserta la función ERROR.TIPICO.XY. Se selecciona las celdas respectivas. Pulsar en Aceptar.

	A	B	C	D
1	X	Y		
2	152	56		
3	157	61		
4	162	67		
5	167	72		
6	173	70		
7	178	72		
8	182	83		
9	188	92		
10	4,06417	=ERROR.TIPICO.XY(B2:B9;A2:A9)		

Interpretación: El valor de $s_e = 4{,}064$, significa que los puntos están dispersos a una distancia de 4,064 de la recta de regresión.

SERIES DE TIEMPO O SERIES CRONOLÓGICAS

Matemáticamente, una serie de tiempo se define por los valores $Y_1, Y_2, Y_{3,}$ .... de una variable Y (ventas mensuales, producción total, etc.) en tiempos $t_1, t_2, t_{3,}$ .... Si se reemplaza a X por la variable tiempo, estas series se definen como distribuciones de pares ordenados (X,Y) en el plano cartesiano, siendo Y una función de X; esto se denota por:

$$Y = f(t) \rightarrow Y = f(X)$$

El principal objetivo de las series de tiempo es hacer proyecciones o pronósticos sobre una actividad futura, suponiendo estables las condiciones y variaciones registradas hasta la fecha, lo cual permite planear y tomar decisiones a corto o largo plazo.

MÉTODOS DE SUAVIZAMIENTO Y PRONÓSTICO

Estos métodos eliminan las fluctuaciones aleatorias de la serie de tiempo, proporcionando datos menos distorsionados del comportamiento real de misma.

a) MÉTODO DE LOS PROMEDIOS MÓVILES

El movimiento medio de orden N de una serie de valores Y_1, Y_2, Y_3, Y_n se define por la sucesión de valores correspondientes a las medias aritméticas:

$$\frac{Y_1 + Y_2 + \cdots Y_N}{N}; \frac{Y_2 + Y_3 + \cdots Y_{N+1}}{N}; \frac{Y_3 + Y_4 + \cdots Y_{N+2}}{N}; \ldots \ldots$$

Nota:
Utilizando adecuadamente estos movimientos medios se eliminan los movimientos o variaciones estacionales, cíclicas e irregulares, quedando sólo el movimiento de tendencia. Este método presenta el inconveniente de que se pierden datos iniciales y finales de la serie original. También se puede observar que a medida que N crece, la cantidad de nuevos datos se reduce.

Si se emplean medias aritméticas ponderadas en el método de los promedios móviles, el método toma de nombre ***Promedios Móviles Ponderados de Orden N.***

Ejemplo ilustrativo: Con los siguientes datos acerca de la ventas en miles de dólares de la Empresa D & M durante los últimos 3 años tomados en períodos de trimestres:

Trimestre	Ventas
1	12
2	16
3	20
4	34
5	23
6	19
7	20
8	35
9	11
10	19
11	24
12	36

1) Suavizar los datos empleando el método de los promedios móviles de orden 3 (longitud de 3 períodos).
2) Pronosticar las ventas para el trimestre número 13.
3) Suponga que para el Gerente de Ventas la última venta realizada es el doble de importante que la penúltima, y la antepenúltima venta tiene la mitad de importancia que la penúltima. Realizar el pronóstico de ventas para el trimestre número 13 empleando el método de los promedios móviles ponderados de orden 3.
4) Elaborar un gráfico en el que consten las ventas y los promedios móviles (ventas suavizadas).

Solución:
1) El cálculo de los promedios móviles de orden 3 se presentan en la siguiente tabla:

Trimestre	Ventas	Pronóstico (Promedios móviles)
1	12	
2	16	(12+16+20)/3 = 16,00
3	20	(16+20+34)/3 = 23,33
4	34	(20+34+23)/3 = 25,67
5	23	(34+23+19)/3 = 25,33
6	19	(23+19+20)/3 = 20,67
7	20	(19+20+35)/3 = 24,67
8	35	(20+35+11)/3 = 22,00
9	11	(35+11+19)/3 = 21,67
10	19	(11+19+24)/3 = 18,00
11	24	(19+24+36)/3 = 26,33
12	36	

2) El último valor del promedio móvil, que en este ejemplo es 26,33, representa el pronóstico de las ventas para el trimestre número 13, y teóricamente para todo trimestre futuro.

3) Para resolver lo planteado se toma en cuenta las 3 últimas ventas con sus respectivos pesos o ponderaciones. Estos datos se presentan en la siguiente tabla:

Trimestre	Ventas	Pesos (w)
10	19	0,5
11	24	1
12	36	2

Remplazando valores en la fórmula de la media aritmética ponderada se obtiene:

$$Pronóstico = \bar{x} = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + \cdots w_k \cdot x_k}{w_1 + w_2 + w_3 + \cdots w_k} = \frac{\sum w \cdot x}{\sum w}$$

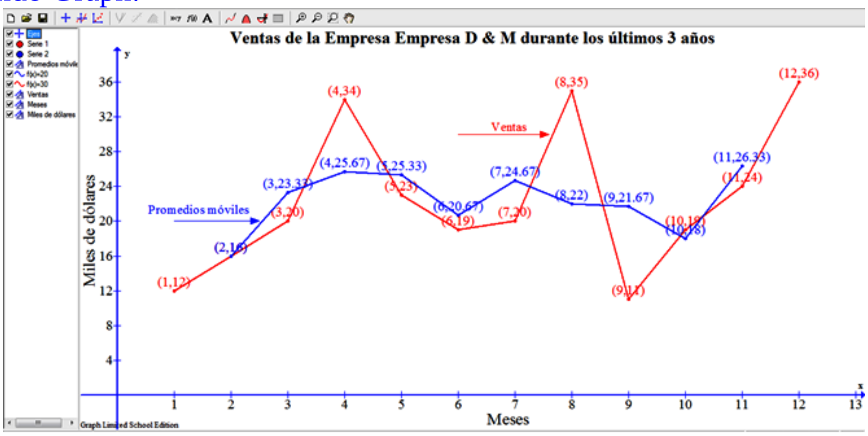
$$Pronóstico = \frac{0,5 \cdot 19 + 1 \cdot 24 + 2 \cdot 36}{0,5 + 1 + 2} = \frac{105,5}{3,5} = 30,14$$

El valor 30,14 es el pronóstico de ventas para el trimestre número 13.

En Excel:

	A	B	C	D	E	F
1	Trimestre	Ventas	Pesos (W)			
2	10	19	0,5			
3	11	24	1			
4	12	36	2			
5	Pronóstico	30,1429	=SUMAPRODUCTO(C2:C4;B2:B4)/SUMA(C2:C4)			

4) El gráfico en el que constan las ventas y los promedios móviles se muestra en la siguiente figura elaborado empleando Graph:



b) SUAVIZACIÓN EXPONENCIAL

Este método contiene un mecanismo de autocorrección que ajusta los pronósticos en dirección opuesta a los errores pasados. Es un caso particular de promedios móviles ponderados de los valores actuales y anteriores en el cual las ponderaciones disminuyen exponencialmente. Se emplea tanto para suavizar como para realizar pronósticos. Se emplea la siguiente fórmula:

$$Y_{t+1} = \alpha \cdot X_t + (1 - \alpha) \cdot Y_t$$

Donde:
Y_{t+1} = pronóstico para cualquier período futuro.
α = constante de suavización, a la cual se le da un valor entre 0 y 1.
X_t = valor real para el período de tiempo.
Y_t = pronóstico hecho previamente para el período de tiempo

Cuando exista menos dispersión en los datos reales respecto a los datos pronosticados entonces será más confiable el método empleado. Para saber cuan preciso es el método empleado en la realización del pronóstico se utiliza la siguiente fórmula del **cuadrado medio del error (CME)** como *indicador de precisión del pronóstico:*

$$CME = \frac{\sum(Y_t - X_t)^2}{n}$$

Siendo n el número de errores

Ejemplo ilustrativo: Con los siguientes datos acerca de la ventas en miles de dólares de la Empresa D & M durante los últimos 12 meses:

Meses	Ventas
Septiembre	6
Octubre	7
Noviembre	6
Diciembre	12
Enero	7
Febrero	10
Marzo	6
Abril	4
Mayo	9
Junio	7
Julio	8
Agosto	6

1) Suavizar los datos empleando el método de suavización exponencial con α = 0,5**.** Pronosticar las ventas para el mes de septiembre. Calcular el cuadrado medio del error. Elaborar un gráfico en el que consten las ventas y los pronósticos.
2) Suavizar los datos empleando el método de los promedios móviles de orden 3. Pronosticar las ventas para mes de septiembre. Calcular el cuadrado medio del error. Elaborar un gráfico en el que consten las ventas y los promedios móviles.
3) ¿Qué método es el más preciso?

Solución: 1) *En Excel:*

	A	B	C	D	E	F
1	Meses	Ventas (X_t)	Pronóstico con α=0,5	$Y_{t+1} = \alpha \cdot X_t + (1-\alpha) \cdot Y_t$	Error	$(Y_t - X_t)^2$
2	Septiembre	6				
3	Octubre	7	6,000	=B2	1,000	=(C3-B3)^2
4	Noviembre	6	6,500	=$B$15*B3+(1-$B$15)*C3	0,250	=(C4-B4)^2
5	Diciembre	12	6,250	=$B$15*B4+(1-$B$15)*C4	33,063	=(C5-B5)^2
6	Enero	7	9,125	=$B$15*B5+(1-$B$15)*C5	4,516	=(C6-B6)^2
7	Febrero	10	8,063	=$B$15*B6+(1-$B$15)*C6	3,754	=(C7-B7)^2
8	Marzo	6	9,031	=$B$15*B7+(1-$B$15)*C7	9,188	=(C8-B8)^2
9	Abril	4	7,516	=$B$15*B8+(1-$B$15)*C8	12,360	=(C9-B9)^2
10	Mayo	9	5,758	=$B$15*B9+(1-$B$15)*C9	10,512	=(C10-B10)^2
11	Junio	7	7,379	=$B$15*B10+(1-$B$15)*C10	0,144	=(C11-B11)^2
12	Julio	8	7,189	=$B$15*B11+(1-$B$15)*C11	0,657	=(C12-B12)^2
13	Agosto	6	7,595	=$B$15*B12+(1-$B$15)*C12	2,543	=(C13-B13)^2
14		Y_{t+1}	6,797	=$B$15*B13+(1-$B$15)*C13		
15	α	0,5		CME	7,090	=PROMEDIO(E3:E13)

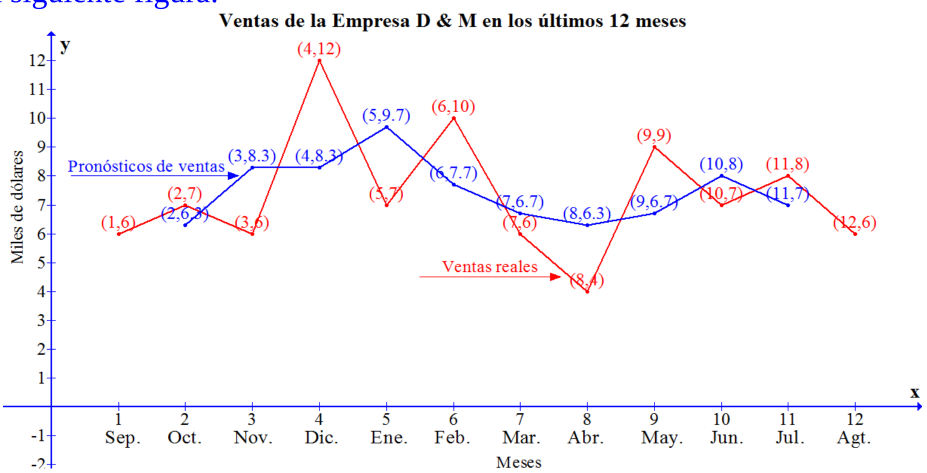
2) Suavizando los datos empleando el método de los promedios móviles de orden 3 elaborado en Excel se muestra en la siguiente figura:

	A	B	C	D	E	F
1	Meses	Ventas (X_t)	Pronósticos (promedios móviles)		Error	$(Y_t - X_t)^2$
2	Septiembre	6				
3	Octubre	7	6,333	=PROMEDIO(B2:B4)	0,444	=(C3-B3)^2
4	Noviembre	6	8,333	=PROMEDIO(B3:B5)	5,444	=(C4-B4)^2
5	Diciembre	12	8,333	=PROMEDIO(B4:B6)	13,444	=(C5-B5)^2
6	Enero	7	9,667	=PROMEDIO(B5:B7)	7,111	=(C6-B6)^2
7	Febrero	10	7,667	=PROMEDIO(B6:B8)	5,444	=(C7-B7)^2
8	Marzo	6	6,667	=PROMEDIO(B7:B9)	0,444	=(C8-B8)^2
9	Abril	4	6,333	=PROMEDIO(B8:B10)	5,444	=(C9-B9)^2
10	Mayo	9	6,667	=PROMEDIO(B9:B11)	5,444	=(C10-B10)^2
11	Junio	7	8	=PROMEDIO(B10:B12)	1,000	=(C11-B11)^2
12	Julio	8	7	=PROMEDIO(B11:B13)	1,000	=(C12-B12)^2
13	Agosto	6				
14				CME	4,522	=PROMEDIO(E3:E12)

Observando el gráfico anterior se tiene que el último pronóstico calculado es de 7, por lo que el pronóstico para septiembre es de 7.

Observando el gráfico anterior se tiene que el cuadrado medio del error es de 4,522.

La gráfica de las ventas y los pronósticos con el método de los promedios móviles elaborada en Graph se muestra en la siguiente figura:



3) Como CME en el método de suavización exponencial es de 7,09 y con el método de los promedios móviles es de 4,52, se concluye que el método de los promedios móviles es el más preciso para este ejemplo ilustrativo.

c) ANÁLISIS DE TENDENCIA

Es necesario describir la tendencia ascendente o descendente a largo plazo de una serie cronológica por medio de alguna línea, y la más adecuada será la que mejor represente los datos y sea útil para desarrollar pronósticos. Se utilizan con más frecuencia los siguientes métodos:

1) MÉTODO DE LOS MÍNIMOS CUADRADOS

Ejemplo ilustrativo: Con los siguientes datos acerca de las ventas en millones de dólares de la Empresa M & M:

Año(X)	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Ventas (Y)	3,4	3,1	3,9	3,3	3,2	4,3	3,9	3,5	3,6	3,7	4	3,6	4,1	4,7	4,2	4,5

1) Hallar la ecuación de tendencia por el método de los mínimos cuadrados**.**

2) Pronosticar la tendencia de exportación para el 2011.

3) Elaborar la gráfica para los datos y la recta de tendencia.

Solución:

1) Para hallar la ecuación de tendencia por el método de los mínimos cuadrados se llena la siguiente tabla, codificando la numeración de los años 1995 como 1, 1996 como 2, y así consecutivamente para facilitar los cálculos.

Año (X)	X	Y	XY	X^2	Y^2
1995	1	3,4	3,40	1	11,56
1996	2	3,1	6,20	4	9,61
1997	3	3,9	11,70	9	15,21
1998	4	3,3	13,20	16	10,89
1999	5	3,2	16,00	25	10,24
2000	6	4,3	25,80	36	18,49
2001	7	3,9	27,30	49	15,21
2002	8	3,5	28,00	64	12,25
2003	9	3,6	32,40	81	12,96
2004	10	3,7	37,00	100	13,69
2005	11	4	44,00	121	16,00
2006	12	3,6	43,20	144	12,96
2007	13	4,1	53,30	169	16,81
2008	14	4,7	65,80	196	22,09
2009	15	4,2	63,00	225	17,64
2010	16	4,5	72,00	256	20,25
Total	136	61	542,3	1496	235,86

Remplazando valores en las siguientes fórmulas se obtiene los valores de a_0 y a_1:

$$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N \sum X^2 - (\sum X)^2} = \frac{61 \cdot 1496 - 136 \cdot 542{,}3}{16 \cdot 1496 - (136)^2} = \frac{17503{,}2}{5440} = 3{,}2175 = 3{,}22$$

$$a_1 = \frac{N \sum XY - \sum X \cdot \sum Y}{N \sum X^2 - (\sum X)^2} = \frac{16 \cdot 542{,}3 - 136 \cdot 61}{16 \cdot 1496 - (136)^2} = \frac{380{,}8}{5440} = 0{,}07$$

Interpretación:

- El valor $a_1 = 0{,}07$ al ser positiva indica que existe una tendencia ascendente de las exportaciones aumentando a un cambio o razón promedio de 0,07 millones de dólares por cada año**.**
- El valor de $a_0 = 3{,}22$ indica el punto en donde la recta interseca al eje Y cuando X = 0, es decir indica las exportaciones estimadas para el año 1996 igual a 3,22.

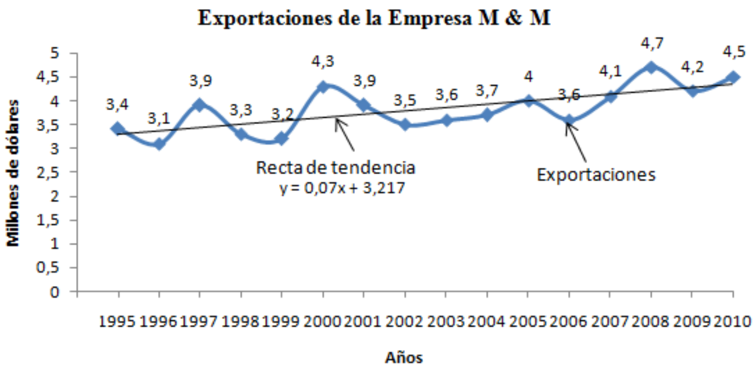
Remplazado los valores anteriores en la recta de tendencia se obtiene:

$Y = a_0 + a_1 X \Longrightarrow Y = 3{,}22 + 0{,}07X$

2) Para pronosticar la tendencia de exportación para el 2011 se reemplaza X = 17 en la recta de tendencia, obteniendo el siguiente resultado:

$Y = 3{,}22 + 0{,}07X \Longrightarrow Y = 3{,}22 + 0{,}07 \cdot 17 = 4{,}41$

3) La gráfica de los datos y la recta de tendencia elaborada en Excel se muestran en la siguiente figura:



2) MÉTODO DE LOS SEMIPROMEDIOS

Este método se aplica con el objeto de simplificar los cálculos y consiste en:

a) Agrupar los datos en dos grupos iguales
b) Obtener el valor central (mediana) de los tiempos y la media aritmética de los datos de cada grupo, consiguiéndose así dos puntos de la recta de tendencia (X_1, Y_1) y (X_2, Y_2).
c) Estos valores se reemplazan en el siguiente sistema:

$$\begin{cases} Y_1 = a_0 + a_1 X_1 \\ Y_2 = a_0 + a_1 X_2 \end{cases}$$

d) Resolviendo el sistema se encuentran los valores de a_0 y a_1, los cuales se reemplazan en la ecuación de la recta de tendencia, la cual es:

$$Y = a_0 + a_1 X$$

Con esta recta de tendencia se puede realizar pronósticos, los cuales son menos exactos que los obtenidos con el método de los mínimos cuadrados, sin embargo, su diferencia es mínima.

Ejemplo ilustrativo : Con los siguientes datos sobre las ventas en millones de dólares de la Empresa D & M

Año(X)	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Ventas(Y)	1,5	1,8	2	1,5	2,2	2	3	2,8	2,4	2,9	3

1) Hallar la ecuación de tendencia por el método de los semipromedios**.**
2) Pronosticar la tendencia de ventas para el 2011.
3) Elaborar la gráfica para los datos y la recta de tendencia.

Solución:

1) Se codifica la numeración de los años 2000 como 1, 2001 como 2, y así consecutivamente para facilitar los cálculos. Se agrupa en dos grupos iguales.

Año	X	Y	Valor central X	Semipromedio Y
2000	1	1,5		
2001	2	1,8		
2002	3	2	3	1,8
2003	4	1,5		
2004	5	2,2		
2005	6	2		
2006	7	3		
2007	8	2,8		
2008	9	2,4	9	2,82
2009	10	2,9		
2010	11	3		

El año 2005 se dejó por fuera para tener grupos con el mismo número de años. El valor central de 3 corresponde a la mediana del primer grupo 1, 2, 3, 4 y 5. El valor central de 9 corresponde a la mediana del segundo grupo 7, 8, 9, 10 y 11. El semipromedio 1,8 corresponden a la media aritmética del primer grupo. El semipromedio 2,82 corresponden a la media aritmética del segundo grupo. De esta manera se obtienen dos puntos (3, 1.8) y (9, 2.82) de la recta de tendencia.

Remplazando los puntos en el siguiente sistema se obtiene:

$$\begin{cases} Y_1 = a_0 + a_1 X_1 \\ Y_2 = a_0 + a_1 X_2 \end{cases} \Rightarrow \begin{cases} 1{,}8 = a_0 + 3a_1 \\ 2{,}82 = a_0 + 9a_1 \end{cases}$$

Resolviendo el sistema empleando la regla de Cramer se obtiene:

$$a_0 = \frac{\Delta_{a_0}}{\Delta} = \frac{\begin{vmatrix} 1,8 & 3 \\ 2,82 & 9 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 1 & 9 \end{vmatrix}} = \frac{7,74}{6} = 1,29\ ;\ a_1 = \frac{\Delta_{a_1}}{\Delta} = \frac{\begin{vmatrix} 1 & 1,8 \\ 1 & 2,82 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 1 & 9 \end{vmatrix}} = \frac{1,02}{6} = 0,17$$

Como a_1 es positiva, la recta tiene una tendencia ascendente (pendiente positiva).

Remplazando los valores calculados se tiene la recta de tendencia, la cual es:

$Y = a_0 + a_1 X \Longrightarrow Y = 1,29 + 0,17X$

2) Para pronosticar la tendencia de exportación para el 2011 se reemplaza X = 12 en la recta de tendencia, obteniendo el siguiente resultado:

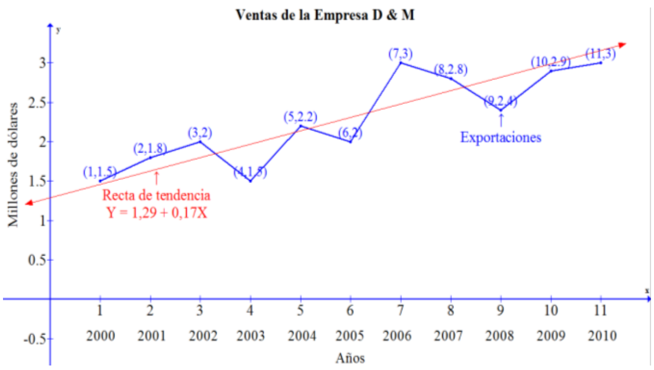
$Y = 1,29 + 0,17X \Longrightarrow Y = 1,29 + 0,17 \cdot 12 \Longrightarrow Y = 3,33$

Interpretación: Existe una tendencia ascendente a un cambio promedio de 0,17 millones de dólares por cada año, por lo que el Gerente de ventas de la empresa debe seguir aplicando las políticas necesarias para mantener la tendencia ascendente y mejorar la tasa de crecimiento.

Los cálculos realizados en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I
1	Año (X)	X	Y	Valor central X		Semipomedio Y		**Pronóstico**	
2	2000	1	1,5						
3	2001	2	1,8						
4	2002	3	2	3	=MEDIANA(B2:B6)	2	=PROMEDIO(C2:C6)		
5	2003	4	1,5						
6	2004	5	2,2						
7	2005	6	2						
8	2006	7	3						
9	2007	8	2,8						
10	2008	9	2,4	9	=MEDIANA(B8:B12)	2,82	=PROMEDIO(C8:C12)		
11	2009	10	2,9						
12	2010	11	3						
13	**2011**	**12**						**3,33**	=H19+H22*B13
14									
15	$Y = a_0 + a_1 X$								
16	1,8	1	3		6	=MDETERM(B16:C17)			
17	2,82	1	9						
18									
19		1,80	3		7,74	=MDETERM(B19:C20)	a_0	1,29	=E19/E16
20		2,82	9						
21									
22		1	1,80		1,02	=MDETERM(B22:C23)	a_1	0,17	=E22/E16
23		1	2,82						

3) La gráfica de los datos y la recta de tendencia elaborada en Graph se muestran en la siguiente figura:



d) ANÁLISIS DE MOVIMIENTOS ESTACIONALES

Para analizar el movimiento estacional debemos estimar cómo varían los datos de la serie cronológica en el período de tiempo. Un conjunto de números que muestra los valores relativos de una variable durante los períodos de tiempo se llama un ***índice estacional*** para la variable. El índice estacional medio del año ha de ser 100%; esto es, la suma de los números índice de los 12 meses suman 1200%, o de los cuatro trimestres suman el 400%, en caso contrario ha se corregirse multiplicado por el factor de ajuste, el mismo que es:

$$Factor\ de\ ajuste\ mensual = \frac{120}{suma\ de\ medias\ mensuales}$$

$$Factor\ de\ ajuste\ trimestral = \frac{400}{suma\ de\ medias\ trimestrales}$$

1) CÁLCULO DEL ÍNDICE ESTACIONAL POR EL MÉTODO DEL PORCENTAJE MEDIO

Este método consiste en calcular los índices estacionales como porcentajes de los períodos de tiempo (mensual o trimestral). Para lo cual se calcula de cada año la media mensual o trimestral, según sea el caso, luego se divide el dato de cada mes o trimestre por la media mensual o trimestral del correspondiente año y se multiplica por 100, y luego se calcular la media de cada mes o trimestre, obteniéndose el índice estacional.

2) DESESTACIONALIZACIÓN DE LOS DATOS O AJUSTE DE LOS DATOS A LA VARIACIÓN ESTACIONAL

Una vez obtenidos los índices estacionales es posible eliminar el movimiento estacional de los datos, para lo cual se divide todos los datos originales por el índice estacional del período de tiempo (mes o trimestre) correspondiente. Los valores desestacionalizados reflejan cómo sería la variable si se corrigiera la influencia estacional.

Ejemplo ilustrativo: Con los datos de la siguiente tabla que muestra las exportaciones en millones de dólares de la Empresa D & M.

Trimestre / Año	I	II	III	IV
2008	20	32	22	40
2009	25	35	30	45
2010	28	38	36	44

Calcular el índice estacional y Desestacionalizar los datos

Solución: *Los cálculos en Excel se muestran en la siguiente figura:*

	A	B	C	D	E	F	G	H
1	Trimestre	I	II	III	IV	Media trimestral		
2	Año							
3	2008	20	32	22	40	28,50		
4	2009	25	35	30	45	33,75		
5	2010	28	38	36	44	36,50		
6	Total	73	105	88	129			
7	Media	24,333	35	29,333	43,000			
8								
9	Trimestre	I	II	III	IV			
10	Año							
11	2008	70,175	112,281	77,193	140,351	=(E3/$F$3)*$A$17		
12	2009	74,074	103,704	88,889	133,333	=(E4/$F$4)*$A$17		
13	2010	76,712	104,110	98,630	120,548	=(E5/$F$5)*$A$17		
14	Media	73,654	106,698	88,237	131,411	=PROMEDIO(E11:E13)	400	=SUMA(B14:E14)
15	**Índice**	73,654	106,698	88,237	131,411	=E14*$C$18		
16								
17	100							
18	Factor de ajuste		1	=400/G14				
19								
20			Datos desestacionalizados					
21	Trimestre	I	II	III	IV			
22	Año							
23	2008	27,154	29,991	24,933	30,439	=(E3/$E$15)*$A$17		
24	2009	33,943	32,803	33,999	34,244	=(E4/$E$15)*$A$17		
25	2010	38,016	35,615	40,799	33,483	=(E5/$E$15)*$A$17		

e) ANÁLISIS DE MOVIMIENTOS CÍCLICOS E IRREGULARES

Los movimientos cíclicos son de tipo periódico y presentan más de un año de duración. Comúnmente, tales movimientos o variaciones no se pueden apartar de la naturaleza irregular, por lo que se analizarán juntas. Recordemos que Y = T·C·E·I de donde C·I = Y/T·E. Por lo que los movimientos cíclicos e irregulares se obtienen dividiendo los datos originales entre el valor de tendencia estimado, y este cociente multiplicando por 100% de la siguiente manera:

$$CI = \frac{Y}{Y_{est.}} \cdot 100\%$$

Donde:

Y = Variable Y; $Y_{est.}$ = Valor de tendencia estimado; CI = Movimientos cíclicos e irregulares

El cociente se *multiplica por 100* a fin de que la media cíclica sea 100. Un valor cíclico relativo de 100 indicará la ausencia de toda influencia cíclica en el valor de la serie de tiempo anual. Para facilitar la interpretación de relativos ciclos suele elaborarse una *gráfica de ciclos*, en el que se describen los ciclos relativos según el año correspondiente. Las cumbres y valles asociados con el componente cíclico de las series de tiempo pueden resultar más evidentes por medio de la elaboración de una gráfica de este tipo.

Ejemplo ilustrativo

Determinar el componente cíclico de cada uno de los valores de la serie cronológica usando la ecuación de tendencia con los siguientes datos acerca de las ventas en millones de dólares de la Empresa M & M:

Año (X)	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Venta(Y)	3,4	3,1	3,9	3,3	3,2	4,3	3,9	3,5	3,6	3,7	4	3,6	4,1	4,7	4,2	4,5

Solución:

1) La ecuación de tendencia lineal obtenida empleando el método de los mínimos cuadrados es:

$$Y = 3{,}22 + 0{,}07X$$

Con esta ecuación se calcula los valores estimados de Y reemplazando los valores de X en la recta de tendencia. Luego se divide los datos originales Y entre el valor de tendencia estimado, y este cociente se multiplica por 100%. Los cálculos en Excel se muestran en la siguiente figura:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Año (X)	X	Y	XY		X^2		Y^2		$Y_{est.} = 3{,}22 + 0{,}07X$	$C·I = (Y/Y_{est.})·100$	
2	1995	1	3,4	3,40	=B2*C2	1	=B2^2	11,56	=C2^2	3,29	103,42	=(C2/J2)*100
3	1996	2	3,1	6,20	=B3*C3	4	=B3^2	9,61	=C3^2	3,36	92,33	=(C2/J2)*101
4	1997	3	3,9	11,70	=B4*C4	9	=B4^2	15,21	=C4^2	3,43	113,79	=(C2/J2)*102
5	1998	4	3,3	13,20	=B5*C5	16	=B5^2	10,89	=C5^2	3,50	94,35	=(C2/J2)*103
6	1999	5	3,2	16,00	=B6*C6	25	=B6^2	10,24	=C6^2	3,57	89,70	=(C2/J2)*104
7	2000	6	4,3	25,80	=B7*C7	36	=B7^2	18,49	=C7^2	3,64	118,21	=(C2/J2)*105
8	2001	7	3,9	27,30	=B8*C8	49	=B8^2	15,21	=C8^2	3,71	105,19	=(C2/J2)*106
9	2002	8	3,5	28,00	=B9*C9	64	=B9^2	12,25	=C9^2	3,78	92,65	=(C2/J2)*107
10	2003	9	3,6	32,40	=B10*C10	81	=B10^2	12,96	=C10^2	3,85	93,57	=(C2/J2)*108
11	2004	10	3,7	37,00	=B11*C11	100	=B11^2	13,69	=C11^2	3,92	94,45	=(C2/J2)*109
12	2005	11	4	44,00	=B12*C12	121	=B12^2	16,00	=C12^2	3,99	100,31	=(C2/J2)*110
13	2006	12	3,6	43,20	=B13*C13	144	=B13^2	12,96	=C13^2	4,06	88,72	=(C2/J2)*111
14	2007	13	4,1	53,30	=B14*C14	169	=B14^2	16,81	=C14^2	4,13	99,33	=(C2/J2)*112
15	2008	14	4,7	65,80	=B15*C15	196	=B15^2	22,09	=C15^2	4,20	111,97	=(C2/J2)*113
16	2009	15	4,2	63,00	=B16*C16	225	=B16^2	17,64	=C16^2	4,27	98,42	=(C2/J2)*114
17	2010	16	4,5	72,00	=B17*C17	256	=B17^2	20,25	=C17^2	4,34	103,75	=(C2/J2)*115
18												
19	Total	136	61	542,30		1496		235,86				
20		=SUMA(B2:B18)	=SUMA(C2:C17)	=SUMA(D2:D17)		=SUMA(F2:F18)		=SUMA(H2:H17)				
21	N	16	CONTAR(B2:B17)									
22												
23	$a_0 = \frac{\sum Y \cdot \sum X^2 - \sum X \cdot \sum XY}{N\sum X^2 - (\sum X)^2}$			3,22		=(C19*F19-B19*D19)/(B21*F19-(B19)^2)						
24												
25												
26	$a_1 = \frac{N\sum YX - \sum X \cdot \sum Y}{N\sum X^2 - (\sum X)^2}$			0,070		=(B21*D19-B19*C19)/(B21*F19-(B19)^2)						
27												

REFERENCIAS BIBLIOGRÁFICAS

Daza, Jorge. (2006). *Estadística Aplicada con Microsoft Excel*. Lima, Perú: Grupo Editorial Megabyte

Govinden, Lincoyán. (1985). *Introducción a la Estadística*. Bogotá, Colombia: Ed. McGraw Hill. Interamericana Editores S.A.

Shao, Stephen. (1980). *Estadística para Economistas y Administradores de Empresas*. México DF: Ed. Herrero Hnos.

Spiegel, Murray. (2000). *Estadística*. Serie de Compendios Schaum. México: Ed. McGraw-Hill

Suárez, Mario. y Tapia, Fausto. (2012). *Interaprendizaje de Estadística Básica*. Ibarra, Ecuador: Universidad Técnica de Norte

Suárez, Mario. (2012). *Interaprendizaje de Probabilidades y Estadística Inferencial con Excel, Winstats y Graph. Ibarra,* Ecuador: Imprenta M&V Gráfic.

Suárez, Mario. (2011). *Análisis Combinatorio*. http://www.monografias.com/trabajos89/analisis-combinatorio/analisis-combinatorio.shtml

Suárez, Mario. (2011). *Probabilidad Teórica*. http://www.monografias.com/trabajos88/probabilidad-teorica/probabilidad-teorica.shtml

Suárez, Mario. (2011). *Regla Particular de la Adición de Probabilidades para Eventos Mutuamente Excluyentes.* http://www.monografias.com/trabajos89/adicion-probabilidades-eventos-mutuamente/adicion-probabilidades-eventos-mutuamente.shtml

Suárez, Mario. (2011). *Regla General de la Adición de Probabilidades para Eventos No Mutuamente Excluyentes*. http://www.monografias.com/trabajos88/regla-general-adicion-probabilidades/regla-general-adicion-probabilidades.shtml

Suárez, Mario. (2011). *Regla General y Particular de la Multiplicación de Probabilidades*. http://es.scribd.com/doc/63206328/REGLA-GENERAL-Y-PARTICULAR-DE-LA-MULTIPLICACION-DE-PROBABILIDADES

Suárez, Mario. (2011). *Probabilidad Total y Teorema de Bayes*. http://www.monografias.com/trabajos89/probabilidad-total-y-teorema-bayes/probabilidad-total-y-teorema-bayes.shtml

Suárez, Mario. (2011), *Distribución Binomial*. http://www.monografias.com/trabajos85/distribucion-binomial/distribucion-binomial.shtml

Suárez, Mario. (2011). *Ejercicios de la Distribución Binomial resueltos con Excel*. http://www.monografias.com/trabajos89/ejercicios-distribucion-binomial-resueltos-excel/ejercicios-distribucion-binomial-resueltos-excel.shtml

Suárez, Mario. (2011). *Introducción a las distribuciones de probabilidad discretas*. http://www.monografias.com/trabajos85/distribuciones-probabilidad-discretas/distribuciones-probabilidad-discretas.shtml

Suárez, Mario. (2011). *Distribución Normal con Excel*. http://www.monografias.com/trabajos89/distribucion-normal-excel/distribucion-normal-excel.shtml

Suárez, Mario. (2011). *Cálculo del tamaño de la muestra*. http://www.monografias.com/trabajos87/calculo-del-tamano-muestra/calculo-del-tamano-muestra.shtml

Suárez, Mario. (2011). *Ejercicios resueltos de prueba de hipótesis*. http://www.monografias.com/trabajos89/ejercicios-resueltos-prueba-hipotesis/ejercicios-resueltos-prueba-hipotesis.shtml

Suárez, Mario. (2011). *Distribución de frecuencias para datos agrupados en intervalos*. http://www.monografias.com/trabajos87/distribucion-frecuencias-datos-agrupados intervalos/distribucion-frecuencias-datos-agrupados-intervalos.shtml

Suárez, Mario. (2011). *Cálculo del tamaño de la muestra*. http://www.monografias.com/trabajos87/calculo-del-tamano-muestra/calculo-del-tamano-muestra.shtml

Suárez, Mario. (2011). *Guía didáctica para el interaprendizaje de medidas de tendencia central.* http://www.monografias.com/trabajos85/interaprendizaje-medidas-tendencia-central/interaprendizaje-medidas-tendencia-central.shtml

Suárez, Mario. (2011). *Media aritmética.* http://www.monografias.com/trabajos85/media-aritmetica/media-aritmetica.shtml

Suárez, Mario. (2011). *Ejemplos ilustrativos resueltos de la Moda*. http://www.monografias.com/trabajos85/ejemplos-ilistrativos-resueltos-moda/ejemplos-ilistrativos-resueltos-moda.shtml

Suárez, Mario. (2011). *La mediana para datos no agrupados y agrupados*. http://www.monografias.com/trabajos85/ejecicios-mediana/ejecicios-mediana.shtml

Suárez, Mario. (2011). *Medidas de posición.* http://www.monografias.com/trabajos87/medidas-posicion/medidas-posicion.shtml

Suárez, Mario. (2011). *Medidas de dispersión.* http://www.monografias.com/trabajos89/medidas-de-dispersion/medidas-de-dispersion.shtml

Suárez, Mario. (2011). *Medidas de forma: asimetría y curtosis.* http://www.monografias.com/trabajos87/medidas-forma-asimetria-curtosis/medidas-forma-asimetria-curtosis.shtml

Suárez, Mario. (2011). *Coeficiente de Correlación por Rangos de Spearman.* http://www.monografias.com/trabajos85/coeficiente-correlacion-rangos-spearman/coeficiente-correlacion-rangos-spearman.shtml

Suárez, Mario. (2011). *Dispersión relativa o coeficiente de variación.* http://www.monografias.com/trabajos88/dispersion-relativa/dispersion-relativa.shtml

Suárez, Mario. (2011). *Regresión potencial mediante el método de los mínimos cuadrados.* http://www.monografias.com/trabajos89/regresion-potencial-metodo-minimos-cuadrados/regresion-potencial-metodo-minimos-cuadrados.shtml

Suárez, Mario. (2011). *Regresión exponencial mediante el método de los mínimos cuadrados.* http://www.monografias.com/trabajos89/regresion-exponencial-metodo-minimos-cuadrados/regresion-exponencial-metodo-minimos-cuadrados.shtml

Suárez, Mario. (2011). *Análisis de tendencia para series de tiempo.* http://www.monografias.com/trabajos87/analisis-tendencia-series-tiempo/analisis-tendencia-series-tiempo.shtml

Suárez, Mario. (2011), *Métodos de suavizamiento y pronóstico para series de tiempo*, http://www.monografias.com/trabajos87/metodos-suavizamiento-y-pronostico-series-tiempo/metodos-suavizamiento-y-pronostico-series-tiempo.shtml

Suárez, Mario. (2012). Interaprendizaje de Probabilidades y Estadística Inferencial con Excel, Winstats y Graph. http://repositorio.utn.edu.ec/handle/123456789/940

Suárez, Mario. (2012). *Introducción a las probabilidades empleando Excel*. http://docentesinnovadores.net/Contenidos/Ver/5744

Suárez, Mario. (2012). *Introducción al Interaprendizaje de Estadística empleando TICs* http://docentesinnovadores.net/Contenidos/Ver/5627

Suárez, Mario. (2012). *Distribuciones Continuas*. http://es.scribd.com/doc/86958218/DISTRIBUCIONES-CONTINUAS

Suárez, Mario. (2012). *Distribución t de Student con Excel* y Winstats. http://docentesinnovadores.net/Contenidos/Ver/5907

Suárez, Mario. (2012). *Análisis de varianza y la F de Fisher con Excel y Winstats*. http://docentesinnovadores.net/Contenidos/Ver/5916

Suárez, Mario. (2012). *Prueba de Hipótesis Chi Cuadrado Emplenado Excel y Winstats*. http://docentesinnovadores.net/Contenidos/Ver/5942

Suárez, Mario. (2012). *Gráficas de control de la calidad empleando Excel.* http://docentesinnovadores.net/Contenidos/Ver/5976

Suárez, Mario. (2012). *Tablas de Probabilidades con ejemplos de lectura*. http://es.scribd.com/doc/86227209/Tablas-de-Probabilidades-Con-Ejemplos-de-Lectura

Suárez, Mario y Tapia, Fausto. (2012). *Interaprendizaje de Estadística Básica. 1ra Edición.* Ecuador, Ibarra. Universidad Técnica del Norte.

Suárez, Mario. (2012). *Análisis de correlación empleando Excel y Graph.* http://www.monografias.com/trabajos93/analisis-correlacion-empleando-excel-y-graph/analisis-correlacion-empleando-excely-graph.shtml

Suárez, Mario. (2013). *Distribuciones de Poisson con Excel, Winstats y Geogebra*. http://es.scribd.com/doc/158851173/DISTRIBUCION-DE-POISSON-CON-EXCEL-WINSTATS-Y-GEOGEBRA-pdf

Suárez, Mario. (2013). *Conceptos básicos de probabilidades y Estadística Inferencial.* http://es.scribd.com/doc/129480693/Conceptos-basicos-de-Probabilidades-y-Estadistica-Inferencial

Suárez, Mario. (2013). *Conceptos básicos de estadística descriptiva e inferencial.* http://www.monografias.com/trabajos96/conceptos-basicos-estadistica-descriptiva-e-inferencial/conceptos-basicos-estadistica-descriptiva-e-inferencial.shtml

Suárez, Mario. (2014). *Coeficiente de correlación de Karl Pearson con Excel, Graph y GeoGebra*. http://www.monografias.com/trabajos-pdf5/coeficiente-correlacion-pearson-excel-graph-y-geogebra/coeficiente-correlacion-pearson-excel-graph-y-geogebra.shtml

Suárez, Mario. (2014). *La recta de los mínimos cuadrados con Excel y GeoGebra.* http://www.monografias.com/trabajos-pdf5/recta-minimos-cuadrados-excel-y-geogebra/recta-minimos-cuadrados-excel-y-geogebra.shtml

Suárez, Mario. (2014). *La Parábola de los mínimos cuadrados con Excel, Graph y Geogebra.* http://www.monografias.com/trabajos-pdf5/parabola-minimos-cuadrados-excel-graph-y-geogebra/parabola-minimos-cuadrados-excel-graph-y-geogebra.shtml

Suárez, Mario. (2014). *Cuartiles, diagrama de caja y bigotes, deciles y percentiles con Excel y con Geogebra.* http://www.monografias.com/trabajos-pdf5/cuartiles-diagrama-caja-y-bigotes-deciles-y-percentiles-excel-y-geogebra/cuartiles-diagrama-caja-y-bigotes-deciles-y-percentiles-excel-y-geogebra.shtml

Suárez, Mario. (2014). *Diagrama de tallo y hojas con GeoGebra.* http://www.monografias.com/trabajos100/diagrama-tallo-y-hojas-geogebra/diagrama-tallo-y-hojas-geogebra.shtml

DATOS BIOGRÁFICOS DEL AUTOR

Mario Orlando Suárez Ibujés nació el 24 de marzo de 1978 en la ciudad de Ibarra, Imbabura, Ecuador. Sus padres Bertha Ibujés y Segundo Suárez. Su esposa Dyanita Elizabeth Rivera Paredes. Sus hijos Emily Monserrath y Mathías Josué.

Sus primeros estudios los realizó en la Escuela Fiscal Mixta "Alejandro Pasquel Monge", del Barrio "La Florida" de la ciudad de Ibarra, en la cual fue Abanderado del Estandarte de la Escuela, Abanderado del Pabellón Nacional y Mejor Egresado.

Sus estudios secundarios los realizó en el Colegio Nacional "Teodoro Gómez de la Torre" de la ciudad de Ibarra, en el cual fue el Mejor Alumno en Matemática durante los tres últimos años, Abanderado del Estandarte del Colegio y Mejor Egresado.

Sus estudios de tercer nivel los realizó en la Universidad Técnica del Norte (UTN), en la cual siendo el Mejor Egresado obtiene el título de Licenciado en Física y Matemática.

Sus estudios de cuarto nivel los realizó en la UTN en convenio con la Asociación de Facultades Ecuatorianas de Filosofía y Ciencias de la Educación (AFEFCE), en la cual obtiene el título de Magíster en Gerencia de Proyectos Educativos y Sociales a los 26 años de edad.

Sus publicaciones registradas en el Instituto Ecuatoriano de Propiedad Intelectual (IEPI) y en la Cámara Ecuatoriana del Libro: Unidades para Producir Medios Instruccionales en Educación (coautor a los 24 años), Interaprendizaje Holístico de Matemática (autor), Hacia un Interaprendizaje Holístico de Álgebra y Geometría (autor), Matemática Recreativa (coautor) , Interaprendizaje de Probabilidades y Estadística Inferencial con Excel, Winstats y Graph (autor), Interaprendizaje de Estadística Básica (coautor), y Probabilidades y Estadística empleando las TIC (autor).

Sus obras artísticas inéditas registradas en el IEPI: Poliprisma 3.0, Poliprisma 4.0, Poliprisma 7.0 y Poliprisma 9.0. Se trata de rompecabezas tridimensionales bicolores integrados por partes prismáticas. Los Poliprismas se encuentran publicados en http://www.monografias.com/trabajos-pdf5/poliprismas/poliprismas.shtml

Sus publicaciones en Internet: Sus tesis de licenciatura y maestría, así como más de 100 artículos sobre Aritmética, Álgebra, Geometría, Trigonometría, Lógica Matemática, Probabilidades, Estadística Descriptiva, Estadística Inferencial, Cálculo Diferencial, Cálculo Integral, y Planificaciones Didácticas se encuentran publicados http://www.monografias.com/usuario/perfiles/mario_suarez_7/monografias ,http://repositorio.utn.edu.ec/handle/123456789/24, http://es.scribd.com/mariosuarezibujes, https://wwwdocentesinnovadores.net/Usuarios/Ver/29591,

Su experiencia docente: Escuela "Alejandro Pasquel Monge" (1998-2001), Colegio Universitario "UTN" (2003-2004), Academia Militar "San Diego" (2004-2008), Unidad Educativa Experimental "Teodoro Gómez de la Torre" (1998-2011), siendo los dos últimos años docente del Bachillerato Internacional. Colegio Nacional Técnico "Mariano Suárez Veintimilla" (2011-2013). Actual profesor de la Unidad Educativa "Ibarra", y de la UTN en la Facultad de Ciencias Administrativas y Económicas.

Sus principales reconocimientos profesionales: Ganador del VI Concurso Nacional y I Internacional de Excelencia Educativa con el proyecto Interaprendizaje de Matemática empleando el Poliprisma 7.0, premio otorgando por la Fundación para la Integración y Desarrollo de América latina (FIDAL) y la Revista Edu@news. Ecuador-Quito en el año 2014, se encuentra publicado en https://www.youtube.com/watch?v=hilX-jZUM8g. Premio a la Excelencia Docente "Rita Lecumberri" en la categoría Educador Innovador, Premio Nacional otorgando por el Ministerio de Educación del Ecuador en el año 2013, se encuentra publicado en http://www.youtube.com/watch?v=fN614do_3lI. Diploma y placa de reconocimiento por la excelente trayectoria como docente investigador y destacado profesor universitario entregado por la Asociación de Profesores de la Facultad de Ciencias Administrativas y Económicas de la UTN en el año 2013. Estatuilla "El Pensador" al Mérito Académico entregado por la Asociación General de Profesores de la UTN en el año 2013. Diploma como Profesor tutor de estudiantes ganadores de Concursos Intercolegiales de Matemática otorgado por la Academia Militar "San Diego" en el año 2008. Diploma y medalla de Honor por haber aportado positivamente al desarrollo académico de Academia Militar "San Diego" en el año 2005. Diploma como Asesor de proyectos ganadores en la Primera Feria Binacional de Ciencia y Tecnología Ecuador-Colombia otorgado por la Unidad Educativa Experimental "Teodoro Gómez de la Torre en el año 2005. Diploma como Profesor tutor de estudiantes ganadores de Concursos Intercolegiales de Matemática entregado por el Colegio Nacional "Ibarra" en el año 2003. Diploma al Mejor Trabajo de Investigación otorgado por la UTN y el Centro Universitario de Investigación Científica y Tecnológica por su trabajo de Tesis "Interaprendizaje de poliedros irregulares de bases paralelas empleando al Multiprisma" en el año 2003. Mención Especial en Ciencias Básicas (Matemática), Premio Nacional otorgado por la VI Feria de Ciencia y Tecnología por haber triunfado con el Proyecto Multiprisma (Un rompecabezas tridimensional bicolor integrado por partes prismáticas) en el año 2001.