

INTERVALOS DE CONFIANZA CON t DE STUDENT EXCEL, WINSTATS Y GEOGEBRA

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 $\alpha/2$ 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.

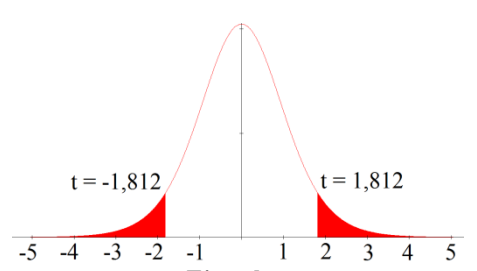
Los grados de libertad de esta distribución se calculan con la siguiente fórmula

$$n - 1$$

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:

<p style="text-align: center;">TABLA N° 4 DISTRIBUCIÓN t DE STUDENT</p>  <p style="text-align: center;">Ejemplos: Para n-1 = 10 grados de libertad $P(t > 1,812) = 0,05$ $P(t < -1,812) = 0,05$</p>									
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

$$\text{Si } 1 - \alpha = 0,95 \Rightarrow \alpha = 1 - 0,95 = 0,05$$

Para leer en la tabla se necesita calcular el área de una cola, la cual es:

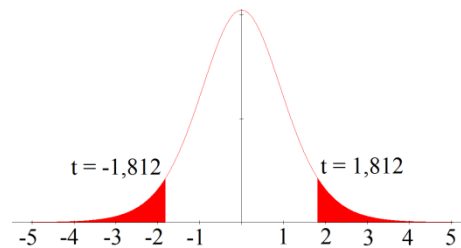
$$\frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

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:

$$n - 1 = 13 - 1 = 12$$

<p style="text-align: center;">TABLA N° 4 DISTRIBUCIÓN t DE STUDENT</p>  <p style="text-align: center;">Ejemplos: Para $n-1 = 10$ grados de libertad $P(t \geq 1,812) = 0,05$ $P(t \leq -1,812) = 0,05$</p>									
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.

The screenshot shows an Excel spreadsheet with the following data:

	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	=	
7	t_2		

The 'Insertar función' (Insert Function) dialog box is open, showing the 'Estadísticas' (Statistics) category selected. The function 'INV.T' is chosen from the list. The description states: 'Devuelve el inverso de cola izquierda de la distribución t de Student.'

b) Clic en Aceptar. En la ventana Argumentos de la función, en Probabilidad seleccionar B3, y en Grados de libertad seleccionar B6.

The screenshot shows the 'Argumentos de función' (Function Arguments) dialog box for the 'INV.T' function. The arguments are set as follows:

- Probabilidad:** B3 (0,025)
- Grados_de_libertad:** B5 (12)

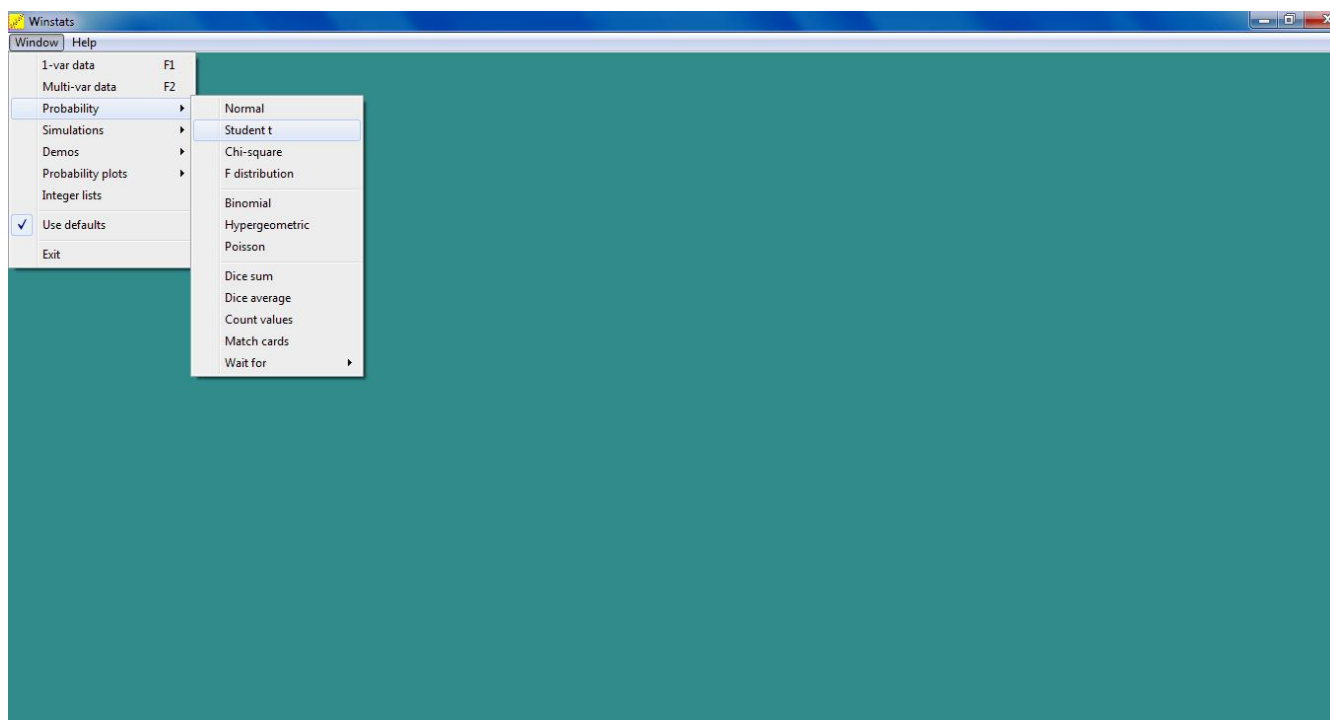
The result of the formula is displayed as -2,17881283. The description states: 'Devuelve el inverso de cola izquierda de la distribución t de Student.'

c) Clic en Aceptar. Los demás cálculos se muestran en la siguiente figura:

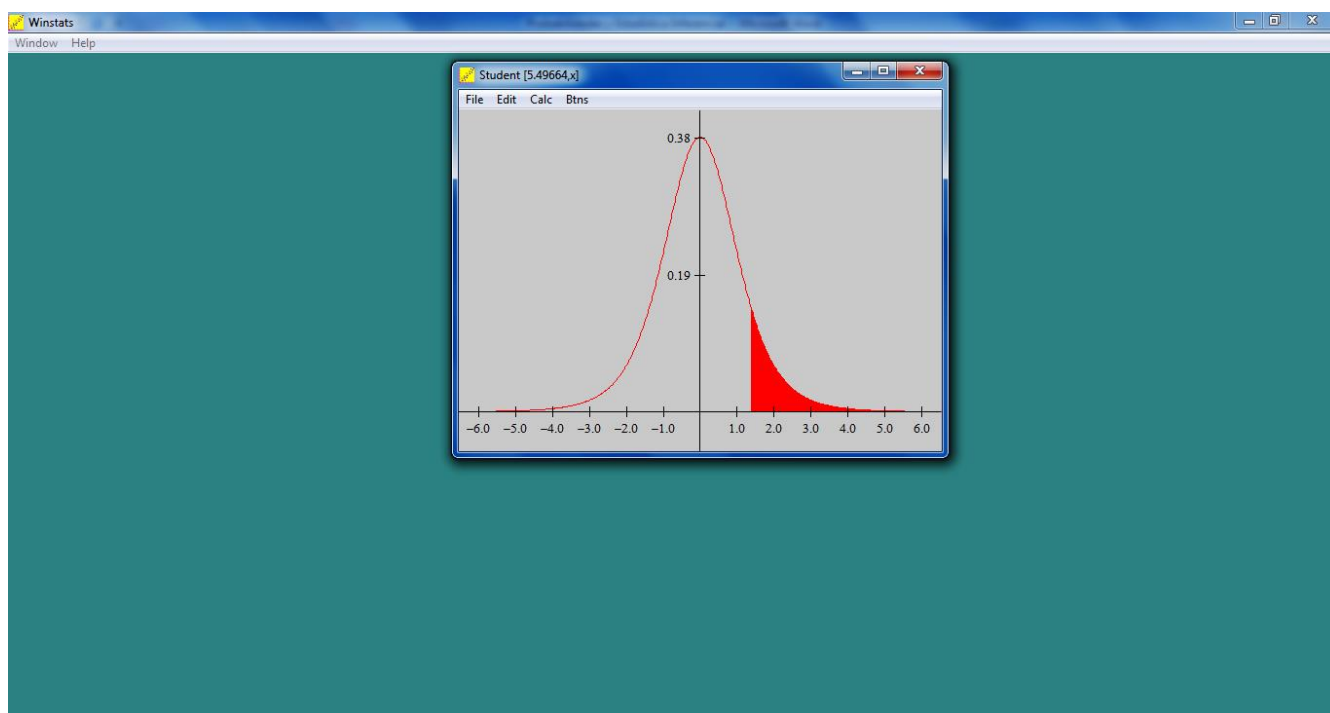
	A	B	C	D
2	n	13		
3	α	0,025	$=(1-B1)/2$	
4	$\frac{\alpha}{2}$			
5	$n - 1$	12	$=B2-1$	
6	t_1	-2,1788	$=INV.T(B3;B5)$	
7	t_2	2,17881	$=B6*-1$	

Para resolver con Winstats se procede de la siguiente manera:

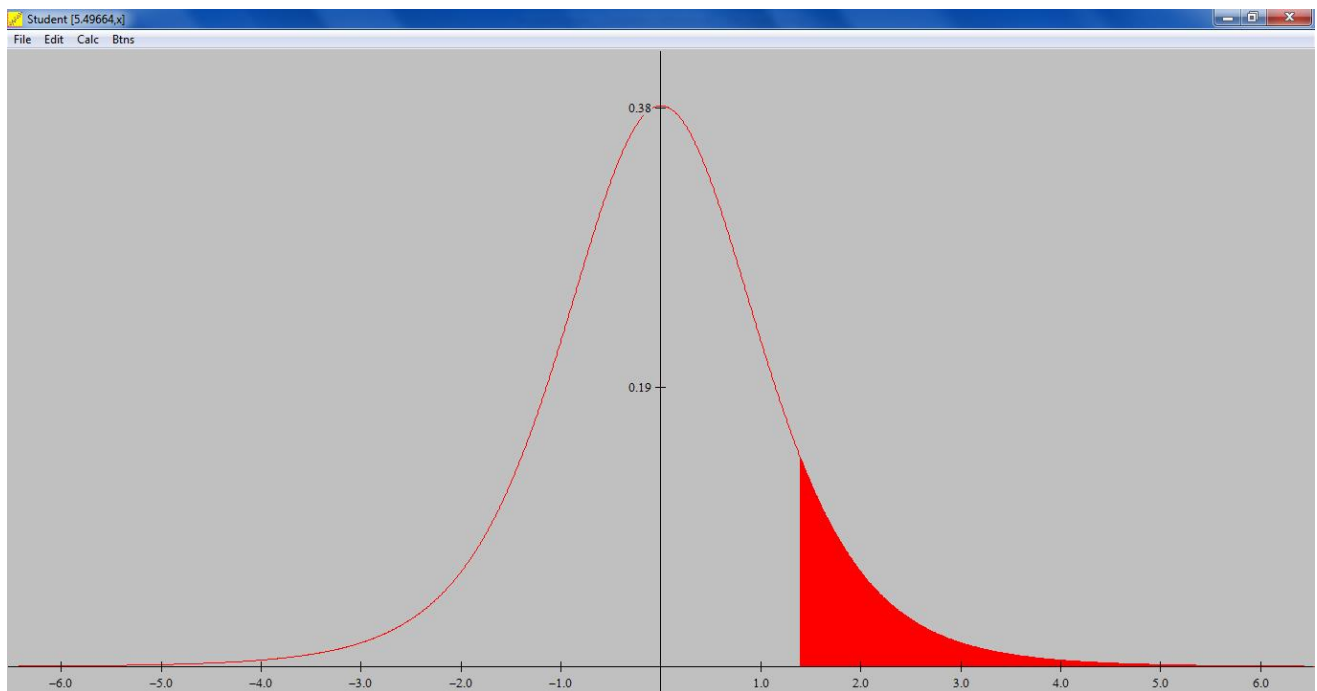
a) Clic en Window y luego en Probability seleccionar Student t



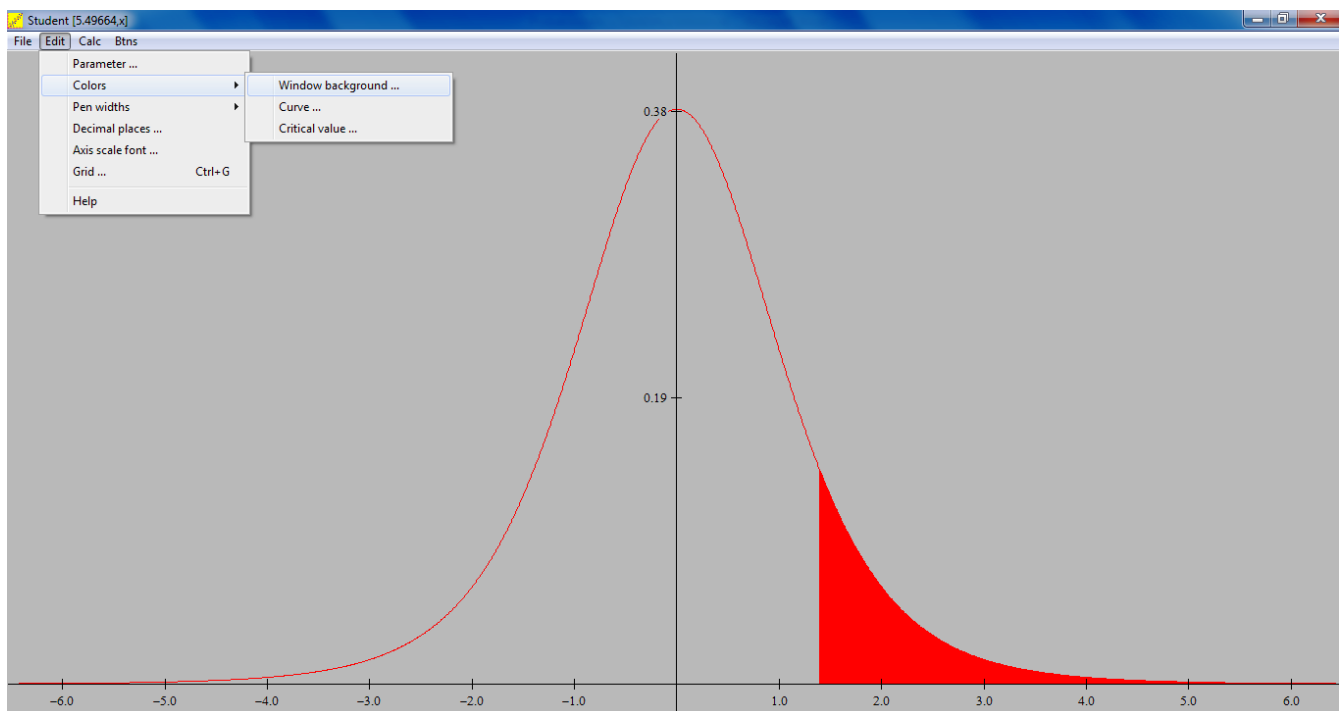
b) Clic en Student t



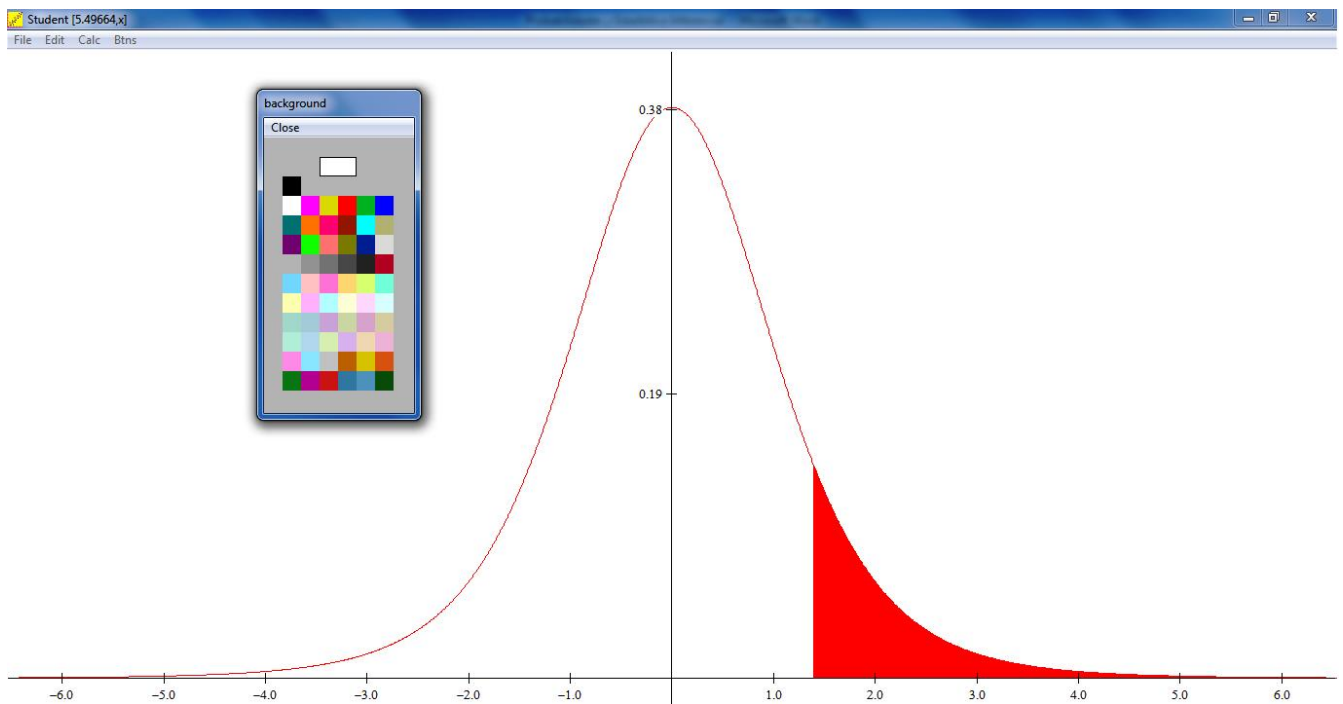
c) Maximizar la ventana de la distribución



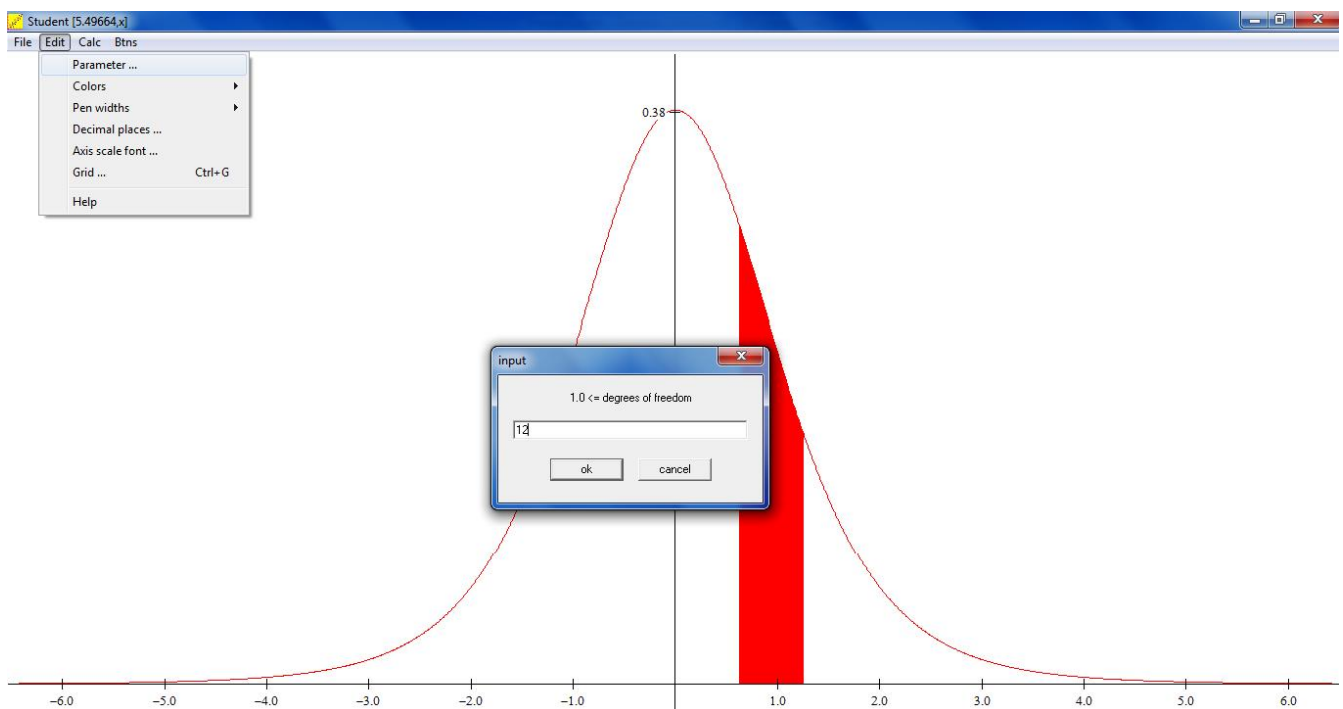
d) Para cambiar el color del fondo, clic en Edit + Colors + Window background



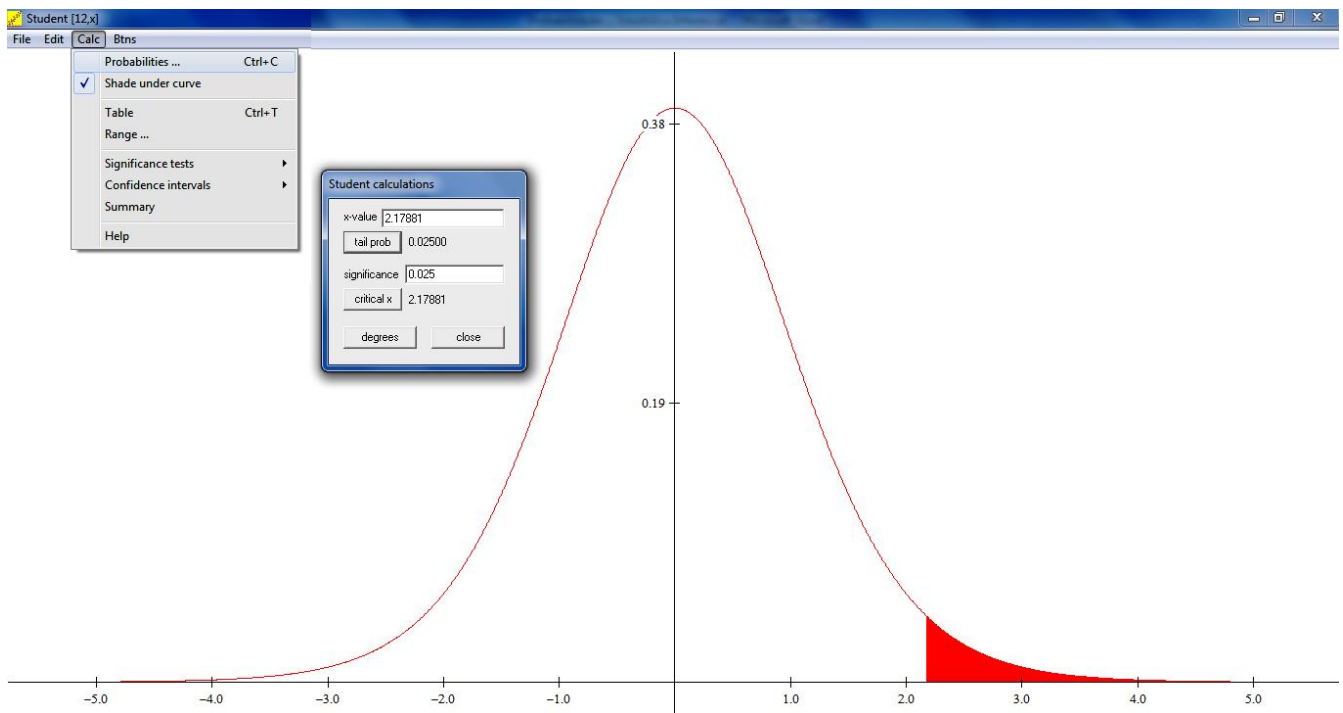
e) Clic Window background. En la venta de background seleccionar el color deseado, que este caso se seleccionó el color blanco. Luego clic en Close para cerrar la venta background.



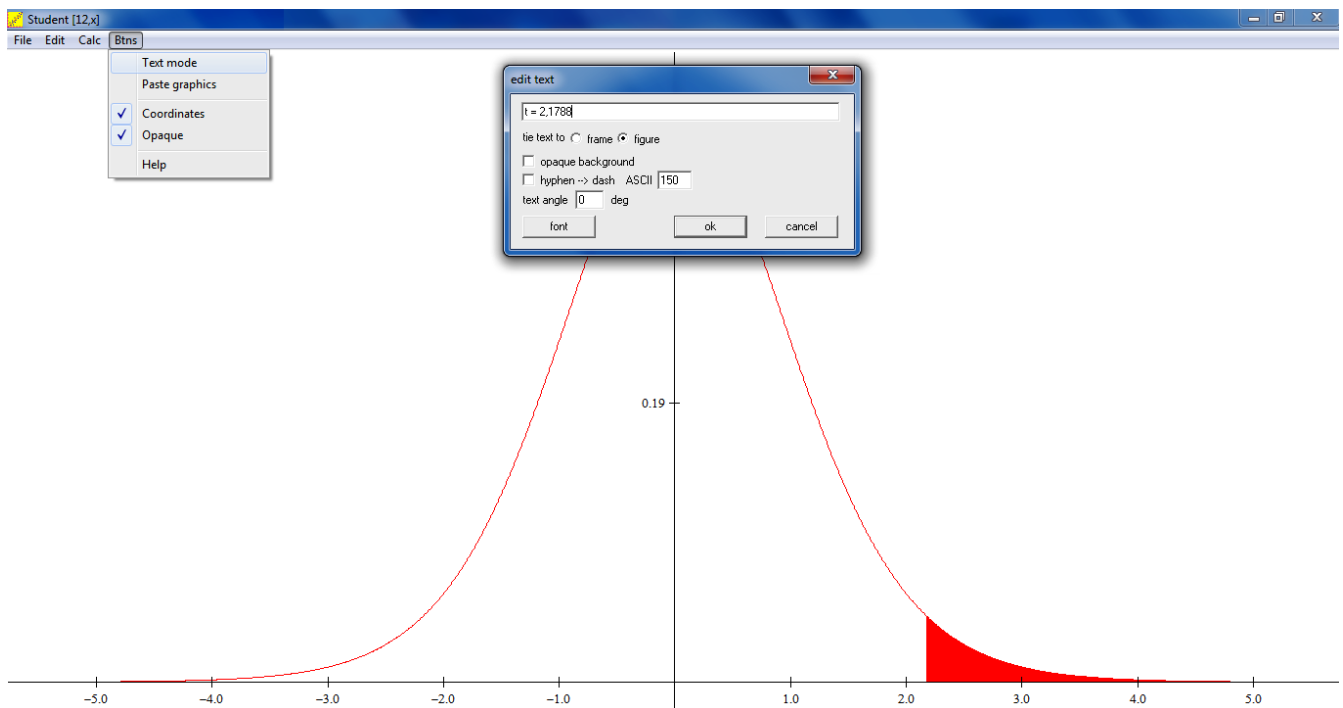
f) Para editar lo grados de libertad, clic en Edit + Parameter...(Parámetros). Clic en Parámetros. En la casilla de la ventana input escribir 12. Clic en ok



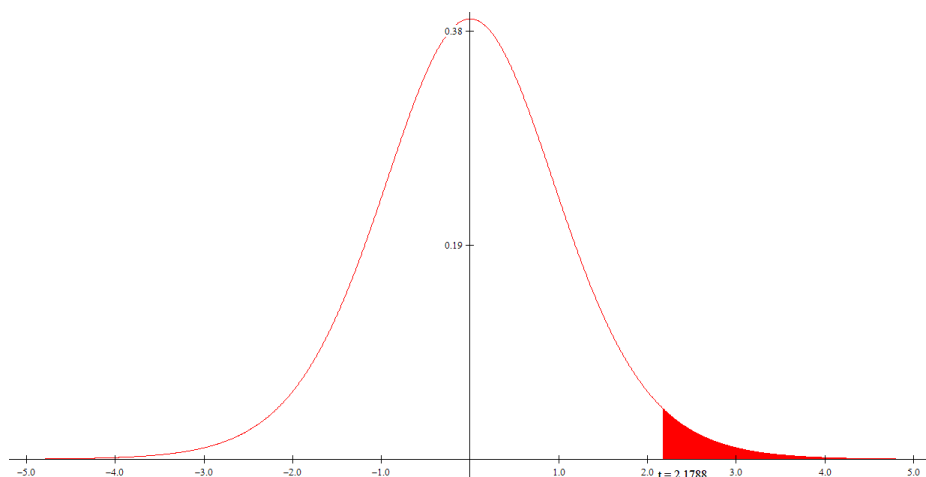
g) Para calcular el valor crítico de t, clic en Calc + Probabilities. En la ventana Student calculations, en significance escribir 0,025 y luego clic en critical x. Clic en close para cerrar la ventana Student calculations.



h) Para escribir textos, clic en Btms. 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.

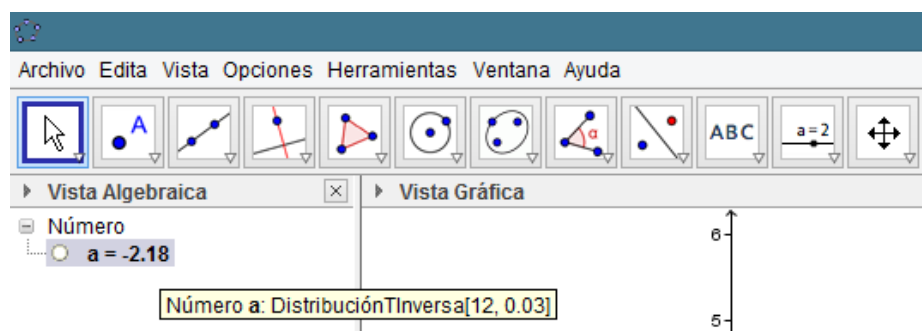


i) Clic en ok de la ventana edit text. Luego arrastar con el mause el texto al lugar deseado



En GeoGebra

Seleccione DistribuciónTInversa[<Grados de Libertad>, <Probabilidad>]



2) Sea $X = t_{(10)}$ hallar el valor de $P(X \leq -1,3722) + P(X \geq 2,7638)$ con lectura en la tabla, Excel y Winstats

Solución:

<p style="text-align: center;">TABLA N° 4 DISTRIBUCIÓN t DE STUDENT</p>									
<p style="text-align: center;">Ejemplos: Para $n-1 = 10$ grados de libertad $P(t \geq 1,812) = 0,05$ $P(t \leq -1,812) = 0,05$</p>									
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:

$$P(X \leq -1,3722) = 0,1 \text{ y } P(X \geq 2,7638) = 0,01$$

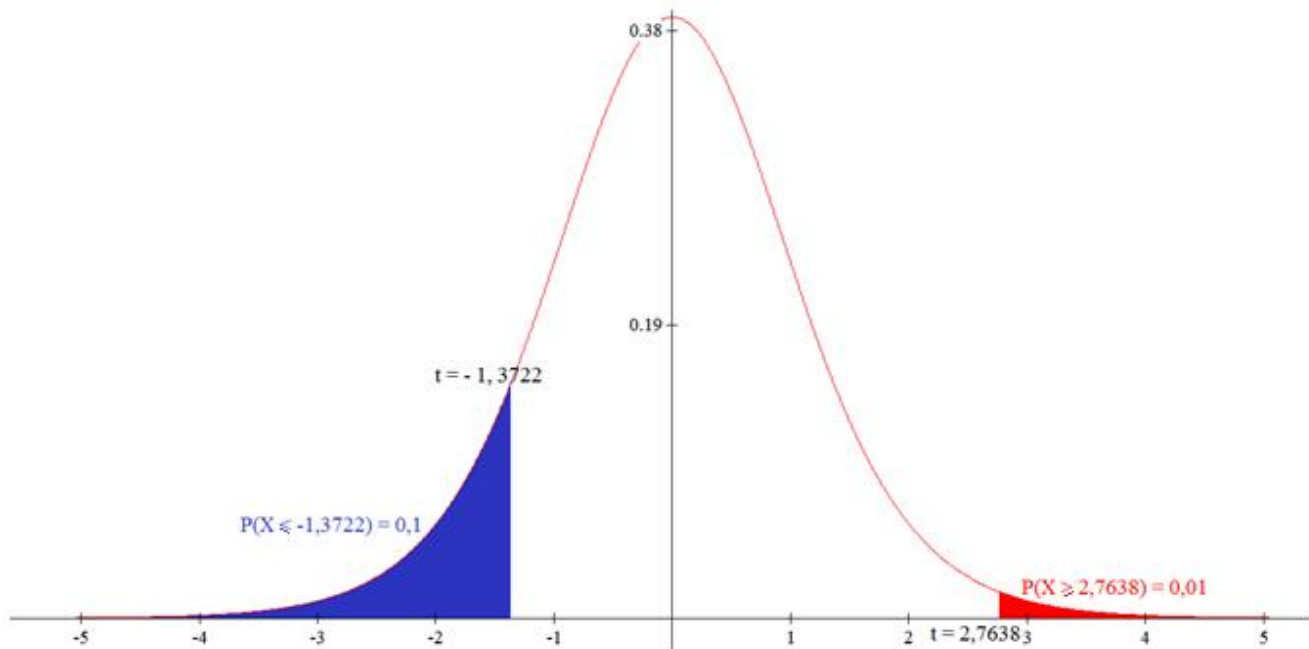
Entonces:

$$P(X < -1,3722) + P(X > 2,7638) = 0,1 + 0,01 = 0,11$$

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	$P(t_1 \leq -1,3722)$	0,1	=DISTR.T.N(B1;B3;VERDADERO)		
5	$P(t_2 \geq 2,7638)$	0,01	=DISTR.T.CD(B2;B3)		
6	$P(X \leq t_1) + P(X \geq t_2)$	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:

$$\mu = 11$$

$$N = 500$$

$$n = 29$$

$$\bar{x} = 10,998$$

$$S = 0,02$$

$$\text{Confianza} = 99\%$$

Como en los datos aparece el tamaño de la población, se debe verificar si el tamaño de la muestra es mayor que el 5% para emplear la fórmula con el factor finito de corrección. Se reemplaza valores en la siguiente fórmula:

$$\frac{n}{N} \cdot 100\% > 5\%$$

$$\frac{29}{500} \cdot 100\% = 5,8\%$$

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:

$$\text{Nivel de confianza} = (1 - \alpha) \cdot 100\%$$

$$\frac{\alpha}{2} = \frac{100\% - \text{Nivel de confianza}}{2}$$

$$\frac{\alpha}{2} = \frac{100\% - 99\%}{2} = 0,005$$

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 $t = \pm 2,7633$

Remplazando valores y realizando los cálculos se obtiene:

$$\bar{x} - t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{x} + t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$10,998 - 2,7633 \frac{0,02}{\sqrt{29}} \sqrt{\frac{500-29}{500-1}} \leq \mu \leq 10,998 + 2,7633 \frac{0,02}{\sqrt{29}} \sqrt{\frac{500-29}{500-1}}$$

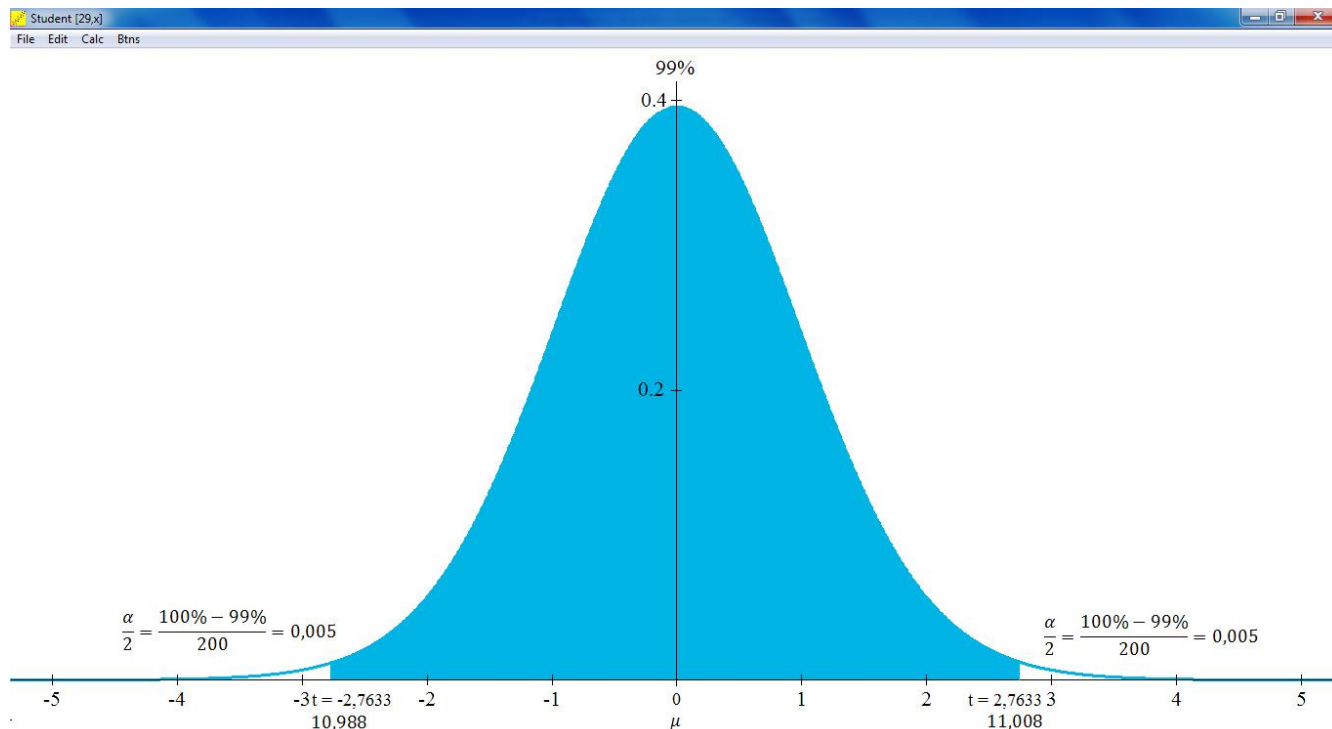
$$10,988 \leq \mu \leq 11,008$$

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	$= \text{INTERVALO.CONFIANZA.T}(B9;B5;B3)$			
11						
12						
13	$\bar{X} - t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \leq \mu \leq \bar{X} + t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$					
14						
15	10,988 $\leq \mu \leq$ 11,008					
16	$= B4 - B10 * \text{RCUAD}((B2 - B3) / (B2 - 1))$		$= B4 + B10 * \text{RCUAD}((B2 - B3) / (B2 - 1))$			

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:

