PRUEBA JI CUADRADO CON EXCEL, WINSTAS Y GEOGEBRA

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: $e = p \cdot o_{total}$
p = proporción muestral
o_{total} = frecuencia total observada

El estadístico de prueba es

$$\chi^2_{prueba} = \frac{(o_1 - e_1)^2}{e_1} + \frac{(o_2 - e_2)^2}{e_2} + \frac{(o_3 - e_3)^2}{e_3} + \cdots \frac{(o_n - e_n)^2}{e_n}$$

$$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$$

Donde:
χ es la letra griega ji
χ^2 se lee ji cuadrado

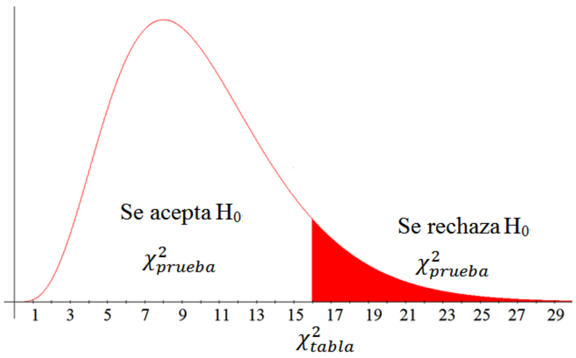
Por lo tanto el valor estadístico de prueba para este caso es la prueba *ji cuadrado* o conocida también como *chi cuadrado*

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.

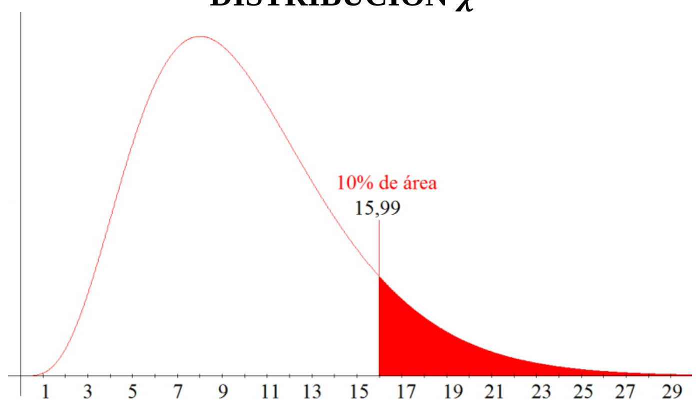
Solución:

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
$\boldsymbol{P(\chi^2 > 15,99) = 0,10 = 10\%}$

	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
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}	21,026	=PRUEBA.CHI.INV(B6;B4)	
8				
9	α	0,01		
10	χ^2_{tabla}	26,2170	=PRUEBA.CHI.INV(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:

$$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$$

$$\chi^2_{prueba} = \frac{(6-10)^2}{10} + \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(15-10)^2}{10} + \frac{(14-10)^2}{10} + \frac{(8-10)^2}{10}$$

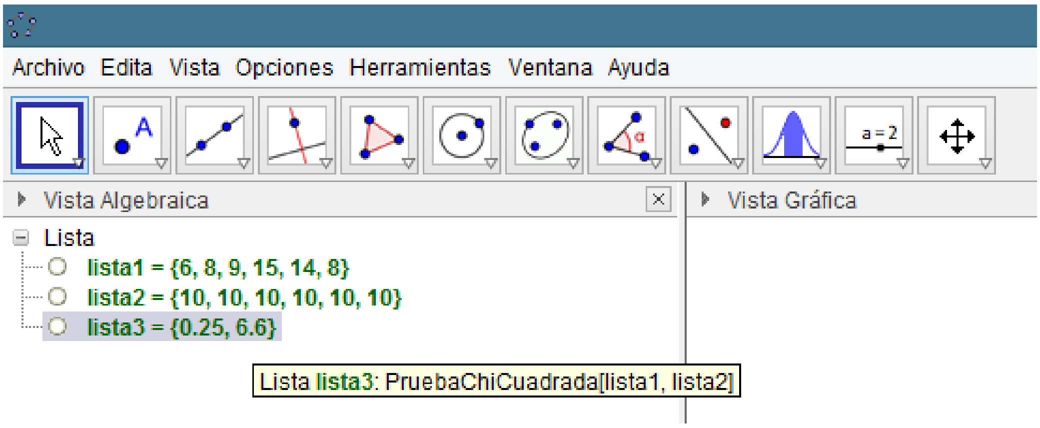
$$\chi^2_{prueba} = 1{,}6 + 0{,}4 + 0{,}1 + 2{,}5 + 1{,}6 + 0{,}4$$

$$\chi^2_{prueba} = 6{,}6$$

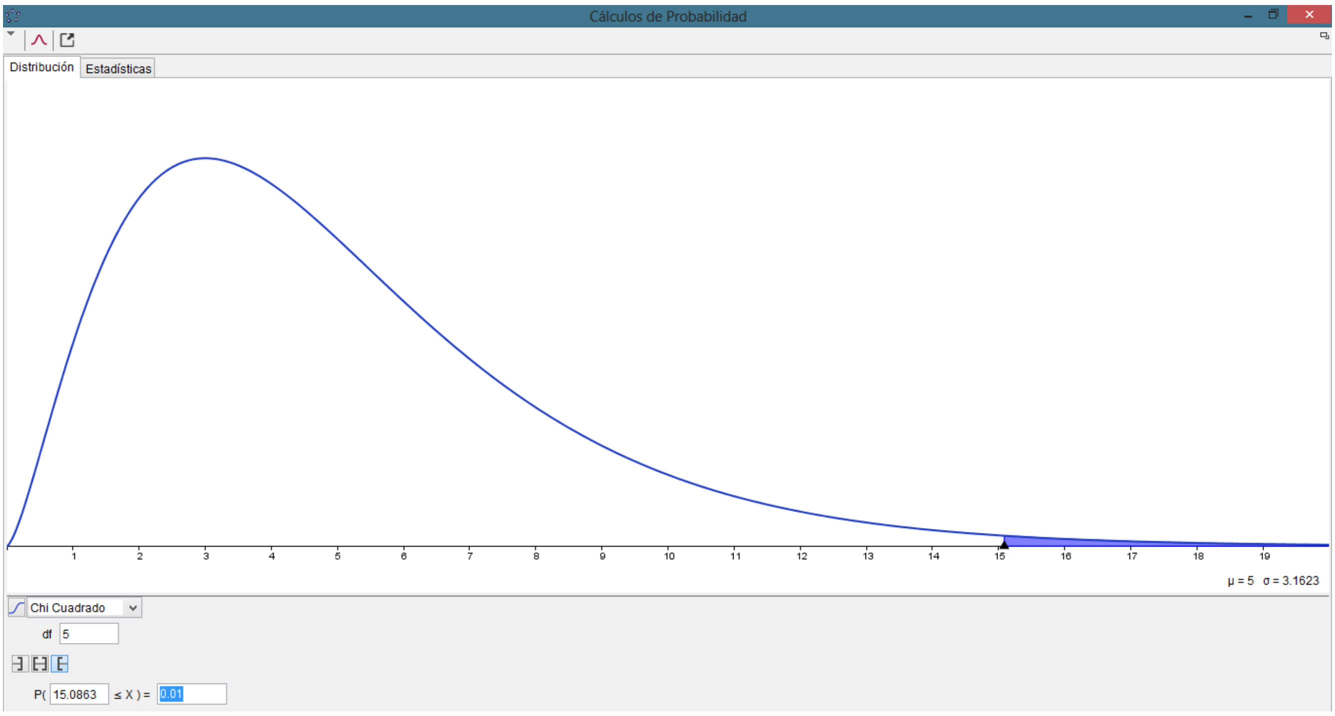
Los cálculos en Excel se muestran en la siguiente figura:

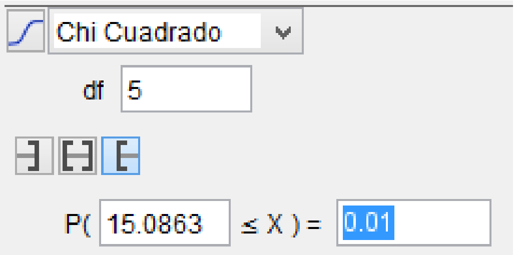
	A	B	C	D	E	F	G
1	Cara del dado	1	2	3	4	5	6
2	Frecuencia observada	6	8	9	15	14	8
3	Frecuencia esperada	10	10	10	10	10	10
4							
5	α	0,01					
6	r	2	=CONTAR(B2:B3)				
7	k	6	=CONTAR(B2:G2)				
8	(r-1)(k-1)	5	=(B6-1)*(B7-1)				
9	χ^2_{tabla}	15,086	=PRUEBA.CHI.INV(B5;B8)				
10							
11	Probabilidad de χ^2_{prueba}	0,2521	=PRUEBA.CHI(B2:G2;B3:G3)				
12	$\chi^2_{prueba} = \sum \frac{(o_i - e_i)^2}{e_i}$	6,6	=PRUEBA.CHI.INV(B11;B8)				
13							

Los cálculos en GeoGebra se muestran en la siguiente figura:

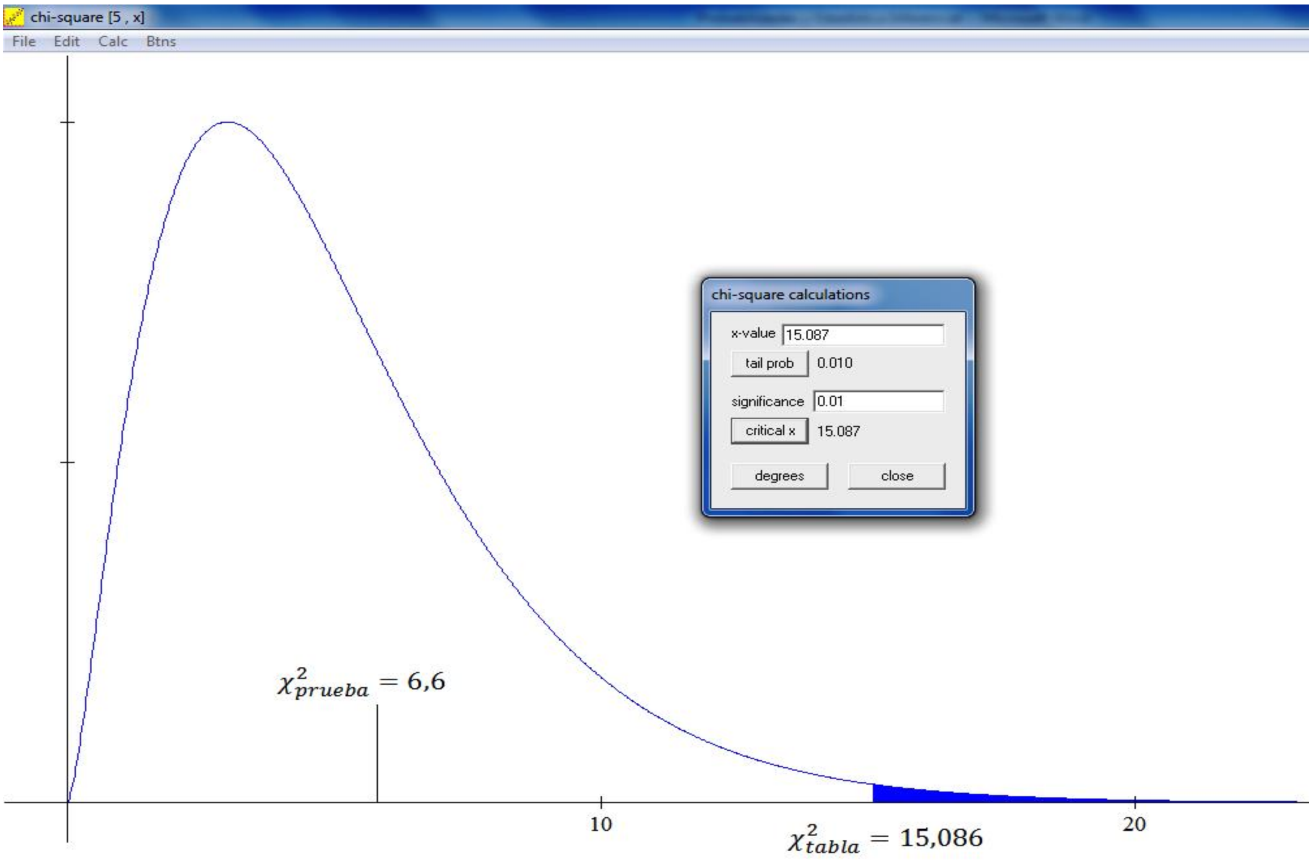


El gráfico elaborado con GeoGebra se muestra a continuación:





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.