Prueba de Ji-Cuadrada

Se trata de una prueba de hipótesis a partir de datos, basada en un cálculo de valor llamado estadístico de prueba, al que se compara con un valor conocido como valor crítico (obtenido de tablas estadísticas).

$$\lambda^2 = \sum_{1}^{k} \frac{(f_0 - f_e)^2}{f_e}$$

$$v = k - 1 - m$$

f_o: frecuencia observada
f_e: frecuencia esperada
k: cantidad de clases o intervalos
m: cantidad de datos empíricos
v: grados de libertad

Para llevar a cabo la prueba de Ji-cuadrada se debe tener en cuenta:

- Las frecuencias esperadas para cada intervalo de clase deben ser de 5 o más. De no alcanzar esta cifra, se deberá agrupar clases o intervalos adyacentes.
- La cantidad de datos empíricos m se refiere a la cantidad de datos obtenidos en base a la observación, que fueron utilizados para calcular las frecuencias esperadas.
- El método es más adecuado para muestras mayores o iguales a 30 elementos.

Si el λ^2 calculado es menor o igual al λ^2 tabulado, entonces no se rechaza la hipótesis planteada.

El método en pasos:

1. Obtener al menos 30 datos de la variable aleatoria a analizar.
2. Calcular la media y la varianza de los datos (cuando corresponda).
3. Crear un histograma de k intervalos (se sugiere $k = \sqrt{n}$), y obtener la frecuencia observada en cada intervalo fo_i.
4. Establecer la hipótesis nula, proponiendo una distribución de probabilidad que se ajuste a la forma obtenida en el histograma.
5. Calcular la frecuencia esperada fe_i, a partir de la función de densidad de la distribución estadística propuesta.
6. Calcular el *estadístico de prueba*:
$$c = \sum_{i=1}^{k} \frac{(fe_i - fo_i)^2}{fe_i}$$
7. Definir el nivel de significancia de la prueba, α, y los grados de libertad υ=k-1-m (m es la cantidad de parámetros o datos empíricos estimados en la distribución propuesta), para determinar el valor crítico de la prueba $\chi^2_{\alpha,\upsilon}$.
8. Comparar el *estadístico de prueba* con el *valor crítico*. Si el *estadístico de prueba* es menor o igual que el *valor crítico*, no se puede rechazar la hipótesis nula.

Ejercicio:

0,15 – 0,22 – 0,41 – 0,65 – 0,84 – 0,81 – 0,62 – 0,45 – 0,32 – 0,07 – 0,11 – 0,29 – 0,58 – 0,73 – 0,93 – 0,97 – 0,79 – 0,55 – 0,35 – 0,09 – 0,99 – 0,51 – 0,35 – 0,02 – 0,19 – 0,24 – 0,98 – 0,10 – 0,31 – 0,17

Intervalo	fo	fe	C	C (AC)
0,0 - 0,2				
0,2 - 0,4				
0,4 - 0,6				
0,6 - 0,8				
0,8 - 1,0				

Prueba de Kolmogorov-Smirnov

Esta prueba también permite determinar la distribución de probabilidad de una serie de datos. Una limitante de la prueba es que solamente puede ser aplicada al análisis de variables continuas.

$$KS = \max_{1-K}(|P(f_o)_{AC} - P(f_e)_{AC}|)$$

f_o: frecuencia observada
f_e: frecuencia esperada
k: cantidad de clases o intervalos
$P()_{AC}$: probabilidad acumulada

Para llevar a cabo esta prueba se debe tener en cuenta:

- En esta prueba los grados de libertad están dados por el tamaño de muestra.
- La prueba es más adecuada para muestras pequeñas, entre 10 y 30 elementos.
- Es aplicable solamente a variables aleatorias continuas.

Si el KS calculado es menor al KS tabulado, entonces no se rechaza la hipótesis planteada.

El método en pasos:

1. Obtener hasta 30 datos de la variable aleatoria a analizar (esta prueba es más adecuada para muestras pequeñas, de 10 a 30 datos).
2. Calcular la media y la varianza de los datos (cuando corresponda).
3. Crear un histograma de k intervalos (se sugiere k= $\sqrt{n}$), y obtener la frecuencia observada en cada intervalo fo_i.
4. Calcular la probabilidad observada Po_i en cada intervalo. Esto es, dividir la frecuencia observada fo_i entre el número total de datos de la muestra n. ($Po_i = fo_i/n$).
5. Acumular las probabilidades Po_i para obtener la probabilidad observada hasta el i-ésimo intervalo (POA_i).
6. Establecer la hipótesis nula, proponiendo una distribución de probabilidad que se ajuste a la forma obtenida en el histograma.
7. Calcular la probabilidad esperada acumulada para cada intervalo, PEA_i, a partir de la función de densidad de la distribución estadística propuesta.
8. Calcular el estadístico de prueba: $c = máx|PEA_i - POA_i|$ con i=1,2,3,…,k
9. Definir el nivel de significancia de la prueba, α, y determinar el valor crítico de la prueba, $D_{\alpha,n}$.
10. Comparar el *estadístico de prueba* con el *valor crítico*. Si el *estadístico de prueba* es menor o igual que el *valor crítico*, no se puede rechazar la hipótesis nula.

Ejercicio:

0,15 – 0,22 – 0,41 – 0,65 – 0,84 – 0,81 – 0,62 – 0,45 – 0,32 – 0,07 – 0,11 – 0,29 – 0,58 – 0,73 – 0,93 – 0,97 – 0,79 – 0,55 – 0,35 – 0,09 – 0,99 – 0,51 – 0,35 – 0,02 – 0,19 – 0,24 – 0,98 – 0,10 – 0,31 – 0,17

Intervalo	**fo**	**fe**	**Po**	**Pe**	**Po_{AC}**	**Pe_{AC}**	**$\vert Po_{AC} - Pe_{AC}\vert$**	**max($\vert Po_{AC} - Pe_{AC}\vert$)**
0,0 - 0,2								
0,2 - 0,4								
0,4 - 0,6								
0,6 - 0,8								
0,8 - 1,0								