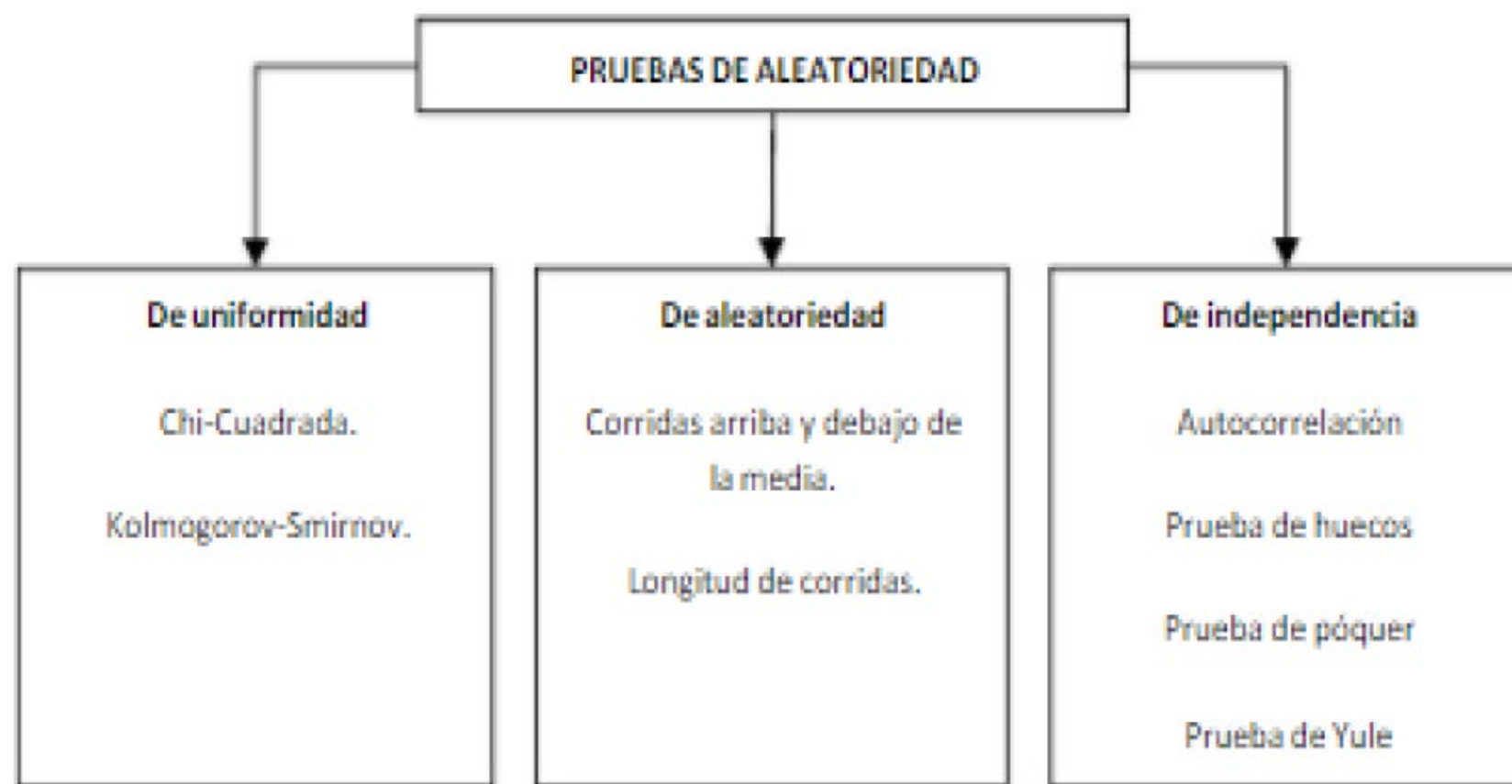




CHI-CUADRADA





PRUEBAS DE UNIFORMIDAD

Una prueba básica que siempre será desarrollada para validar un nuevo generador es la prueba de uniformidad. Dos métodos de pruebas disponibles. Estas son las pruebas Kolmogorov-Smirnov y la prueba Chi-Cuadrada. Ambas de estas pruebas miden el grado de ajuste entre la distribución de una muestra de números aleatorios generados y la distribución uniforme teórica.

Ambas de estas pruebas están basadas en la Hipótesis Nula de que no existe diferencia entre la distribución de la muestra y la distribución teórica.

La prueba de uniformidad es la propiedad más importante que deben cumplir un conjunto de números aleatorios, y para comprobar su acatamiento se han desarrollado pruebas estadísticas como la prueba chi-cuadrada y de kolmogorov-smirnov.

PRUEBAS DE UNIFORMIDAD

Una de las propiedades mas importantes que debe cumplir un conjunto de números r_i es la uniformidad.

Para comprobar esto se ha desarrollado pruebas estadísticas tales como:

Prueba de Chi-Cuadrada

Prueba Kolmogoroc- Smirnov

Donde:

$H_0 = r_i \sim U(0,1)$

$H_1: \underline{r_i \text{ no son uniformes}}$

Prueba Chi-Cuadrada

Esta prueba es una de las más útiles y ampliamente utilizadas en la estadística, para determinar qué tan significativa es la diferencia entre las frecuencias observadas y esperadas de uno o más categorías.

La diferencia entre las frecuencias esperadas y observadas, son consideradas como el error muestral. Las frecuencias observadas son calculadas a partir de un conteo de los números que coinciden en un subintervalo determinado, y las frecuencias esperadas están en función a una distribución de probabilidad teórica.

Una prueba de chi-cuadrada es una prueba de hipótesis que compara la distribución observada de los datos con una distribución esperada de los datos.

Existen varios tipos de pruebas de chi-cuadrada:

Prueba de bondad de ajuste de chi-cuadrada

Utilice este análisis para probar qué tan bien una muestra de datos categóricos ~~se ajusta a una distribución teórica~~. Por ejemplo, usted puede comprobar si un dado es justo, lanzando el dado muchas veces y utilizando una prueba de bondad de ajuste de chi-cuadrada para determinar si los resultados siguen una distribución uniforme. En este caso, el estadístico de chi-cuadrada cuantifica qué tanto varía la distribución observada de los conteos con respecto a la distribución hipotética.

Pruebas de chi-cuadrada de asociación e independencia

Los cálculos para estas pruebas son iguales, pero la pregunta que se está tratando de contestar puede ser diferente.

Prueba de asociación: Utilice una prueba de asociación para determinar si una variable está asociada a otra variable. Por ejemplo, determine si las ventas de diferentes colores de automóviles dependen de la ciudad donde se venden.

Prueba de independencia: Utilice una prueba de independencia para determinar si el valor observado de una variable depende del valor observado de otra variable. Por ejemplo, determine si el hecho de que una persona vote por un candidato no depende del sexo del elector.

Esta prueba busca determinar si los números del conjunto r_i se distribuyen uniformemente en el intervalo (0,1).

Procedimiento:

1. Generar la muestra de números aleatorios de tamaño N.

2. Subdividir el intervalo [0,1] en n subintervalos. Para esto se lleva a cabo una división del intervalo en m subintervalos, en donde es recomendable $m = \sqrt{N}$.

3. Para cada subintervalo contar la frecuencia observada FO y calcular la frecuencia esperada FE de números aleatorios, la cual se obtiene dividiendo N/m .

4. Calcular el estadístico de prueba.

$$X_0^2 = \sum_{i=1}^n \frac{(FE_i - FO_i)^2}{FE_i}$$

5. Comparar el valor calculado X_0^2 contra el valor tabulado de la distribución X^2 , con (grados de libertad y una significancia? Si X_0^2 es menor que $X^2(n-1)$? entonces no se puede rechazar la uniformidad de los números aleatorios

Nota: Calcular los grados de libertad (gl) en función de número de categorías [K]:

$$gl = K - 1 = m - 1$$

La ji cuadrada se utiliza cuando:

- Cuando los datos puntualizan a las escalas nominal u ordinal.
- Se utiliza solo la frecuencia.
- Poblaciones pequeñas.
- Cuando se desconocen los parámetros media, moda, etc.
- Cuando los datos son independientes.
- Cuando se quiere contrastar o comparar hipótesis.
- Investigaciones de tipo social - muestras pequeñas no representativas >5.
- Cuando se requiere de establecer el nivel de confianza o significatividad en las diferencias.
- Cuando la muestra es seleccionada no probabilísticamente.
- χ^2 permite establecer diferencias entre f y se utiliza solo en escala nominal.
- Población > a 5 y < a 20.

Realizar la prueba Chi-cuadrada a los siguientes 100 números de un conjunto r_i , con un nivel de confianza de 95 %.

$$\alpha = 5\%$$

$$N = 100$$

0.347	0.832	0.966	0.472	0.797	0.101	0.696	0.966	0.404	0.603
0.993	0.371	0.729	0.007	0.189	0.977	0.843	0.562	0.549	0.992
0.674	0.628	0.055	0.494	0.494	0.235	0.178	0.775	0.797	0.252
0.426	0.004	0.022	0.742	0.674	0.898	0.641	0.674	0.821	0.19
0.46	0.224	0.99	0.786	0.393	0.461	0.011	0.977	0.246	0.881
0.189	0.753	0.73	0.797	0.292	0.876	0.707	0.562	0.562	0.821
0.112	0.191	0.584	0.347	0.426	0.057	0.819	0.303	0.404	0.64
0.37	0.314	0.731	0.742	0.213	0.472	0.641	0.944	0.28	0.663
0.909	0.764	0.999	0.303	0.718	0.933	0.050	0.415	0.819	0.444
0.178	0.516	0.437	0.393	0.268	0.123	0.945	0.527	0.459	0.652

$$m = \sqrt{100} = \boxed{10}$$

Antes de proceder, es recomendable crear una tabla similar a la tabla 2.1, en donde se resumen los pasos que deben llevarse a cabo en la prueba Chi-cuadrada.

Tabla 2.1 Cálculos para la prueba Chi-cuadrada.

Intervalo	O_i	$E_i = \frac{n}{m}$	$\frac{(E_i - O_i)^2}{E_i}$
● [0.00-0.10)	7	10	0.9
● [0.10-0.20)	9	10	0.1
● [0.20-0.30)	8	10	0.4
● [0.30-0.40)	9	10	0.1
● [0.40-0.50)	14	10	1.6
● [0.50-0.60)	7	10	0.9
● [0.60-0.70)	11	10	0.1
● [0.70-0.80)	14	10	1.6
● [0.80-0.90)	9	10	0.1
● [0.90-1.00)	12	10	0.4

El estadístico $\chi_0^2 = \sum_{i=1}^{10} \frac{(E_i - O_i)^2}{E_i} = 6.2$ es menor al estadístico correspondiente de la

Chi-cuadrada $\chi_{0.05,9}^2 = 16.9$. En consecuencia, no se puede rechazar que los números r_i siguen una distribución uniforme.