

Test Kolmogorov-Smirnov

Definición

La prueba de Kolmogorov-Smirnov (KS) de una muestra es una prueba estadística no paramétrica que se utiliza para determinar si una muestra proviene de una distribución teórica específica. Es una herramienta útil para comparar la distribución observada de una muestra con una distribución teórica como la normal, exponencial, entre otras.

La prueba KS de una muestra compara la función de distribución acumulada (CDF) observada de una muestra con la CDF de una distribución teórica especificada. El objetivo es identificar si hay una diferencia significativa entre la distribución empírica de la muestra y la distribución teórica.

Hipótesis

- **Hipótesis Nula (H_0):** La muestra proviene de la distribución teórica especificada.
- **Hipótesis Alternativa (H_1):** La muestra no proviene de la distribución teórica especificada.

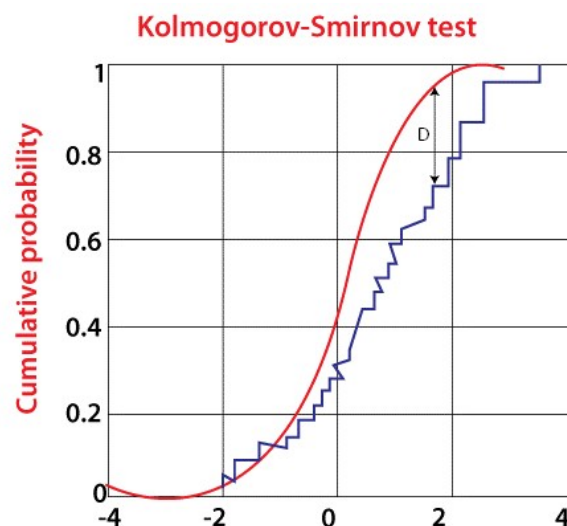
Cálculo del Estadístico KS

El estadístico KS se calcula como la máxima diferencia absoluta entre la CDF empírica de la muestra y la CDF de la distribución teórica:

$$D = \sup_x |F_n(x) - F(x)|$$

Donde:

- $F_n(x)$ es la CDF observada de la muestra.
- $F(x)$ es la CDF de la distribución teórica.



Supuestos

Para que la prueba KS sea válida se deben cumplir los siguientes supuestos:

- Las observaciones en la muestra son independientes e idénticamente distribuidas (i.i.d.).
- La distribución teórica es continua.

Interpretación

En el caso de la prueba KS, la distribución del estadístico D se deriva de la teoría del proceso estocástico y se relaciona con el proceso de Brown, es decir, La distribución exacta de D es compleja y no tiene una forma cerrada simple pero depende de n , el tamaño de la muestra, por lo que para tamaños de muestras grandes el valor crítico se puede aproximar usando la siguiente fórmula:

$$D_{crit} = \sqrt{-\frac{1}{2n} \ln\left(\frac{\alpha}{2}\right)}$$

Una vez calculado D_{crit} podemos compararlo con D con el fin de aceptar o rechazar la hipótesis nula de la siguiente manera:

- Si $D > D_{crit}$ Se rechaza la hipótesis nula, es decir, no se tiene suficiente evidencia para decir que la muestra proviene de la distribución teórica.
- Si $D < D_{crit}$ Se acepta la hipótesis nula, es decir, se tiene suficiente evidencia para decir que la muestra proviene de la distribución teórica.