PRUEBA DE KOLMOGOROV-SMIRNOV

La prueba de Kolmogorov-Smirnov es una prueba estadística utilizada para determinar si una muestra de datos sigue una distribución específica. Fue desarrollada por los matemáticos Andrey Kolmogorov y Nikolai Smirnov en la década de 1930. Esta prueba es ampliamente utilizada para evaluar la normalidad de los datos, pero también se aplica a otros tipos de distribuciones.

La prueba de Kolmogorov-Smirnov se basa en la comparación de la función de distribución acumulada (FDA) empírica de los datos con la FDA teórica de la distribución en estudio. El procedimiento general de la prueba se puede resumir en los siguientes pasos:

- **Hipótesis nula (H0):** La hipótesis nula establece que la muestra de datos proviene de una distribución específica (por ejemplo, una distribución normal).
- **Hipótesis alternativa (H1):** La hipótesis alternativa afirma que la muestra de datos no sigue la distribución específica.
- **Cálculo de la función de distribución acumulada empírica:** Se calcula la función de distribución acumulada empírica de los datos, que es una estimación de la FDA de la muestra.
- **Comparación con la FDA teórica:** Se compara la FDA empírica con la FDA teórica de la distribución en estudio. Se busca la mayor diferencia absoluta entre ambas distribuciones, conocida como estadístico de prueba (D).
- **Cálculo del valor de prueba:** Se calcula el valor de prueba (KS) utilizando una fórmula que depende del tamaño de la muestra y el nivel de significancia deseado.
- **Valor crítico:** Se compara el valor de prueba con un valor crítico, obtenido de tablas de referencia o mediante software estadístico. Si el valor de prueba es mayor que el valor crítico, se rechaza la hipótesis nula y se concluye que los datos no siguen la distribución específica.

Es importante destacar que la prueba de Kolmogorov-Smirnov es no paramétrica, lo que significa que no hace suposiciones sobre los parámetros de la distribución. Por lo tanto, puede ser utilizada para evaluar la normalidad de los datos incluso cuando la media y la varianza son desconocidas o cuando los datos presentan asimetría o sesgo.

Además, la prueba de Kolmogorov-Smirnov es sensible al tamaño de la muestra. Para muestras grandes, la prueba tiene más poder para detectar desviaciones de la distribución específica. Sin embargo, en muestras pequeñas, la prueba puede tener menos poder y puede ser más propensa a errores de tipo II.

La prueba de Kolmogorov-Smirnov es una prueba estadística utilizada para evaluar si una muestra de datos sigue una distribución específica. Se basa en la comparación de la función de distribución acumulada empírica de los datos con la función de distribución acumulada teórica de la distribución en estudio. La interpretación adecuada de los resultados debe considerar el tamaño de la muestra y el nivel de significancia deseado.

Estadístico de prueba de Kolmogorov-Smirnov

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \begin{cases} 1 & \text{si } y_i \leq x, \\ 0 & \text{alternativa.} \end{cases}$$

Para dos colas el estadístico viene dado por

$$D_n^+ = \max(F_n(x) - F(x))$$

$$D_n^- = \max(F(x) - F_n(x))$$

donde $F(x)$ es la distribución presentada como hipótesis.

Referencias:

Box, G. E., Hunter, W. G., & Hunter, J. S. (2005). Estadística para investigadores: Diseño, innovación y descubrimiento (2a ed.). Wiley.

Cochran, W. G. (1977). Técnicas de muestreo (3a ed.). Ediciones Díaz de Santos.

Devore, J. L., & Peck, R. (2018). Estadística para ingenieros (9a ed.). Cengage Learning.

Johnson, R. A., & Wichern, D. W. (2007). Applied multivariate statistical analysis (6th ed.). Pearson.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introducción al análisis de regresión lineal (7a ed.). Cengage Learning.

Montgomery, D. C., Runger, G. C., & Hubele, N. F. (2019). Estadística aplicada y probabilidad para ingenieros (7a ed.). Cengage Learning.

Rice, J. A. (2006). Mathematical statistics and data analysis (3rd ed.). Cengage Learning.

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2018). Probabilidad y estadística para ingeniería y ciencias (9a ed.). Pearson.