

Tarea7. [Prob. y Estadística 1 \(jupyter nb\)](#), [Estad. básica \(jupyter nb\)](#), [Estad. tutorial \(jupyter nb\)](#), [Estad. descriptiva \(jupyter nb\)](#), [Estadística con scipy \(jupyter nb\)](#); [Linear regression \(jupyter nb\)](#).

Formalismo

El objetivo de la tarea es realizar un análisis básico de estadísticas y calcular la **media**, **varianza** y la **distribución de probabilidad de masa (PMF)** a partir de una serie de datos.

1. **Media:** La **media** es el valor promedio de un conjunto de datos.
2. **Varianza:** La **varianza** mide la dispersión de los datos respecto a la media.
3. **Distribución de Probabilidad de Masa (PMF):** La **PMF** es una función que asigna una probabilidad a cada valor de un conjunto discreto de datos.

Algoritmos

Para llevar a cabo este análisis, se usan los siguientes métodos numéricos:

1. **Cálculo de la media:** Utilizamos la fórmula de la media mencionada anteriormente, implementada a través de una función que suma todos los elementos del conjunto y los divide entre el número total de elementos.
2. **Cálculo de la varianza:** Después de obtener la media, se calcula la diferencia cuadrada de cada elemento respecto a la media. Luego, se calcula la media de esas diferencias al cuadrado, lo que nos da la varianza.
3. **Cálculo de la PMF:** Primero, se cuenta la frecuencia de cada valor en el conjunto de datos. Luego, se divide la frecuencia de cada valor por el número total de elementos para obtener la probabilidad de cada valor.

Análisis Crítico

Lo que hemos aprendido:

La media nos proporciona una idea clara del centro de los datos. En este caso, la media de los datos es 3.5, lo que refleja que los valores están distribuidos alrededor de este número.

La varianza nos da una medida de dispersión. Con una varianza de 6.4, sabemos que los datos no están demasiado concentrados alrededor de la media.

La PMF nos da una idea de la frecuencia de cada valor en el conjunto de datos, lo cual es útil para entender cómo se distribuyen los datos.

¿Cómo se podría hacer mejor?:

Se podría usar un enfoque más eficiente para el cálculo de la varianza utilizando herramientas de bibliotecas como numpy o scipy que optimizan los cálculos.

Además, la visualización gráfica de la PMF, la distribución de frecuencias y los histogramas podría ofrecer una mejor comprensión visual de los resultados. Se podrían agregar gráficos de barras o histogramas para ilustrar la distribución de los datos.