

La Estadística es la ciencia de la recopilación y análisis de los datos. Datos que en su mayoría provienen de experimentos que se realizan para simular el mundo real.

La probabilidad ayuda a saber que tan cercanos están estos datos a la realidad o que tan probables son los resultados comparados con los experimentos planteados.

Ligada en sus inicios a la resolución de problemas de los juegos de azar, la probabilidad es la ciencia que mide la certidumbre/incertidumbre de que ocurra o no un evento.

La probabilidad permite comprender y analizar fenómenos aleatorios, hallando un valor real dentro de un rango generalizado de $[0,1]$ 0 a 100%.

La Probabilidad condicionada nos muestra cual es la probabilidad de que ocurra un evento cuando es afectado por otro evento previo.

Estos eventos pueden ser independientes, al ser realizados de manera sucesiva, no afectan a la acción siguiente.

Los eventos dependientes, son aquellos que al ser realizados, afectan los eventos subsecuentes.

El Teorema de Bayes, que genera discordia con los frequentistas, permite actualizar probabilidades a medida que obtenemos más evidencia e información de los sucesos u ocurrencias aleatorias. En términos generales, actualiza o modifica o mejora la creencia inicial sobre las probabilidades de un evento aleatorio. Corrigiendo así las posibles grandes diferencias entre probabilidades teóricas y reales

Le componen cuatro elementos fundamentales:

- Probabilidades a priori $P(A)$, que reflejan la creencia inicial que se tiene de la variable aleatoria.
- Evidencia $P(B)$, condicionada por las probabilidades iniciales, es la probabilidad obtenida por los experimentos en la vida real.
- Verosimilitud $P(B|A)$, es la probabilidad de las evidencias condicionada por las probabilidades a priori. La probabilidad de obtener la evidencia dada una creencia inicial
- Probabilidad a posteriori $P(A|B)$: Probabilidad de que se observen los eventos A, dada la evidencia B

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)}$$

Una variable aleatoria es el resultado numérico de un experimento aleatorio.

Existen las variables aleatorias discretas, que son aquellas que se pueden describir con valores finitos y enteros. Se pueden contar, como por ejemplo en el lanzamiento de un dado, las caras de 1 a 6.

La variable aleatoria continua puede tomar valores infinitos dentro de un rango de números reales no numerables. Por ejemplo, el tiempo o la temperatura.

Las variables aleatorias se pueden caracterizar mediante la densidad o función de probabilidad, la cual describe la probabilidad de que la variable aleatoria tome un valor particular dentro de un rango definido. (Masa o Densidad de probabilidad según el tipo)

También las caracteriza el espacio muestral, que es el conjunto de todos los resultados posibles de un experimento aleatorio. Otros parámetros para caracterizar pueden incluir las medidas de tendencia central o las medidas de dispersión de la variable aleatoria.

El Valor esperado es un promedio ponderado de todos los posibles valores de la una variable aleatoria, donde los pesos corresponden a la probabilidad asociada a esos valores.

La Media es una medida de tendencia central; es la suma de todos los valores en un conjunto de datos dividida por el número total de elementos del conjunto.

La desviación estándar (en las mismas unidades de la media), es la raíz cuadrada de la varianza, la cual mide qué tan propagados o que tan dispersos se encuentran un conjunto de valores aleatorios respecto a su media.

Ley de los grandes números es teorema fundamental que enuncia que, si un experimento es repetido muchas veces, tantas como tendiendo al infinito, la frecuencia de que suceda un cierto evento tiende a ser una constante.

El Teorema del límite central enuncia que la suma de n variables aleatorias independientes, con un $n > 30$ por ejemplo, tiende a una distribución normal, incluso si las variables aleatorias originales no tienen una distribución normal