
Introducción a Estadística usando Wolfram Mathematica

Algunas Funciones Importantes!

RandomReal: Se utiliza para generar números aleatorios con punto flotante (números reales) en un rango especificado. Podemos especificar el *rango* y la cantidad de números reales que queremos que nos entregue. Otras funciones similares a ésta son:

- ✓ RandomInteger: Entrega un número entero aleatorio
- ✓ RandomComplex: Entrega un número complejo aleatorio
- ✓ RandomChoice: Escoge una variable aleatoria de una lista entregada
- ✓ RandomVariate: Toma una distribución como argumento y produce valores aleatorios que siguen esa distribución.

⋮
etc

SeedRandom: Es como una especie de “inicio” para el generador de números aleatorios en Mathematica. Cuando utilizamos `SeedRandom``, le damos al generador de números aleatorios un *número inicial*, llamado “semilla”. Este número inicial es como un punto de partida para generar números aleatorios.

La idea importante aquí es que si utilizamos la misma semilla, obtendremos los mismos números aleatorios cada vez que ejecutes nuestro programa. Esta función nos permite hacer que nuestros resultados aleatorios sean predecibles. Notemos que la diferencia entre diferentes semillas en la función `SeedRandom` es que cada semilla genera una secuencia diferente de números aleatorios. En otras palabras, cambiar la semilla cambiará la secuencia de números aleatorios generados, por ejemplo, `SeedRandom[123]`, `SeedRandom[12]` y `SeedRandom[245]` generan tres secuencias diferentes de números aleatorios.

AbsoluteTime: Devuelve un número que representa la marca de tiempo en segundos desde la “Época” (que es un momento de referencia en el tiempo, a menudo el 1 de enero de 1900 o el 1 de enero de 1970, dependiendo del sistema) .

Histogram: Función utilizada para crear gráficos de histogramas a partir de una lista de datos, es decir, crea una representación gráfica de la distribución de frecuencia de los datos.

SmoothHistogram: Se utiliza para crear un histograma *suavizado* que muestra una estimación de la densidad de probabilidad de los datos. A diferencia de un histograma convencional, que utiliza barras, el histograma suavizado utiliza una curva suave para representar la *distribución* de los datos.

Plot: Se utiliza para crear gráficos de funciones matemáticas o representaciones de datos en general. Podemos utilizar esta función para visualizar funciones matemáticas, relaciones, curvas de datos, etc. Otras funciones similares a ésta son:

✓ **ListPlot & ListLinePlot:** Se utiliza para crear gráficos de dispersión a partir de listas de datos. Muestra puntos en un plano cartesiano donde el eje x representa una variable y el eje y representa otra. Es ideal para visualizar relaciones entre dos conjuntos de datos.

✓ **ContourPlot:** Se utiliza para crear gráficos de contorno que representan superficies tridimensionales en un plano bidimensional. Es útil para visualizar curvas de nivel de funciones bidimensionales.

✓ **DensityPlot:** Se utiliza para visualizar la densidad de una función bidimensional. Crea un mapa de colores donde los colores representan los valores de la función en diferentes partes del plano.

✓ **ArrayPlot:** Se utiliza para crear gráficos de matrices bidimensionales. Toma una matriz de datos numéricos y representa cada valor como un cuadro de color en una cuadrícula bidimensional. Cada color en el gráfico corresponde al valor de la matriz en esa ubicación. Esto es útil para visualizar patrones en datos matriciales, como mapas de calor, imágenes, datos geoespaciales y más.

✓ **Show:** Se utiliza para combinar gráficos o representaciones visuales en una sola figura. Permite superponer, agrupar y personalizar gráficos previamente creados, lo que facilita la creación de visualizaciones complejas combinando múltiples elementos gráficos en un solo output.

PDF: Se utiliza para calcular la *función de densidad de probabilidad* (PDF, por sus siglas en inglés) de una distribución de probabilidad. Básicamente describe cómo se distribuyen las probabilidades de los valores en una variable aleatoria continua o discreta. La sintaxis básica es `PDF[distribución, valor]`, donde *distribución* es el nombre de la distribución de probabilidad que estamos analizando y *valor* es el valor en el que deseamos evaluar la PDF para la distribución. Es útil para comprender la forma de la distribución de probabilidad y calcular probabilidades asociadas a eventos específicos en la variable aleatoria.

CDF: Se utiliza para calcular la función de distribución acumulativa (CDF, por sus siglas en inglés) de una distribución de probabilidad. La CDF describe cómo se acumulan las probabilidades de que la variable aleatoria asociada a una distribución tome valores menores o iguales a un valor específico.

Operadores Útiles:

? Function: Se usa para buscar y listar todas las funciones, símbolos y expresiones que tienen nombres que coinciden con un patrón específico. Por ejemplo, si ejecutas `?*Distribution``, Mathematica mostrará una lista de todas las funciones, distribuciones y expresiones cuyos nombres contienen dicha

palabra .

?* Function: Se utiliza para buscar información y ayuda sobre funciones, símbolos y expresiones en el sistema. Al escribir ``?`` seguido de un nombre de función, símbolo o expresión, obtendremos información detallada sobre esa entidad.

@: Se utiliza como una forma abreviada de aplicar una función a un argumento. Es una notación alternativa a la notación funcional tradicional de `f[expr]`. Por ejemplo, en lugar de escribir `Sin[x]`, podemos escribir equivalentemente `Sin @ x`.

//: El símbolo `//` al final de una función en Wolfram Mathematica se utiliza para aplicar una función a un argumento particular. Por ejemplo, si tenemos la función `Sin` y deseamos aplicarla al valor `x`, `x//Sin`, es equivalente a `Sin[x]`

@/: El operador se utiliza para mapear la función a través de una lista de elementos. En otras palabras, es una forma concisa de aplicar una función a todos los elementos de una lista. Por ejemplo, si tenemos una función `f[x_]:=x^2` y una lista de valores `lista={1,2,3,4}`, escribir `f @/ lista` aplicará la función `f` a cada elemento de la lista, generando una nueva lista con los resultados de `{1, 4, 9, 16}`.

[[]]: Se utiliza para acceder a elementos individuales de una lista, matriz u otra estructura de datos indexada. Por ejemplo, si tenemos una lista llamada `miLista`, puedes acceder a su primer elemento utilizando `miLista[[1]]`, al segundo elemento con `miLista[[2]]`, y así sucesivamente.

Clear: Se utiliza para eliminar definiciones previas de símbolos y funciones. Por ejemplo, si previamente has definido una variable `x` con un valor específico y luego deseas borrar esa definición, podemos hacerlo con `Clear[x]`.

Ejercitación & Aplicación

Ejercicio 1:

(a) Utiliza la función `SeedRandom` para establecer una semilla para la generación de números aleatorios. Luego, utiliza la función `RandomReal` para generar una lista de 20 números aleatorios entre 0 y 1. Guarda esta lista en una variable llamada `datos`.

(b) Utiliza la función `ListPlot` para crear un gráfico de dispersión de los datos generados. Etiqueta los ejes y dale un título al gráfico.

(c) Cambia la semilla y ejecuta todo el código nuevamente. Observa cómo cambian los datos generados y el gráfico resultante.

Ejercicio 2:

- (a) Utiliza *AbsoluteTime* para obtener la marca de tiempo actual. Luego, usa *RandomReal* para generar un número aleatorio entre 0 y 1. Combina estos dos valores para crear un dato de tiempo aleatorio. Repite este proceso 10 veces y almacena los resultados en una lista llamada *datosTemporales*.
- (b) Crea un gráfico de líneas que muestre cómo cambian los datos temporales con el tiempo, etiquetando los ejes y dándole un título al gráfico.
- (c) Genera un nuevo conjunto de datos temporales con el tiempo actual y observa cómo cambian los datos y el gráfico resultante .

Ejercicio 3:

- (a) Genera una lista de 1000 datos aleatorios desde una distribución normal con media (μ) 5 y desviación estándar (σ) 2 utilizando la función *RandomVariate*.
- (b) Crea un histograma de los datos utilizando la función *Histogram*. Personaliza el gráfico, agregando etiquetas de ejes y un título.
- (c) Calcula la función de densidad de probabilidad (PDF) de una distribución normal con $\mu = 5$ y $\sigma = 2$ utilizando la función *PDF*
- (d) Grafica la PDF junto con el histograma en un solo. Asegúrate de que el rango del gráfico muestre adecuadamente tanto el histograma como la PDF.

Ejercicio 4:

- (a) Genera dos listas de datos aleatorios: una con 1000 datos desde una distribución normal con $\mu = 0$ y $\sigma = 1$, y la otra con 1000 datos desde una distribución uniforme en el intervalo $[0, 1]$.
- (b) Crea un gráfico que muestre ambas distribuciones en un mismo gráfico utilizando *SmoothHistogram* para suavizar los datos.
- (c) Calcula la función de distribución acumulativa (CDF) de ambas distribuciones utilizando *CDF*.
- (d) Grafica las CDF de ambas distribuciones en un solo gráfico utilizando *Plot*. Asegúrate de que el rango del gráfico sea adecuado para mostrar ambas CDF.