



Lección #4

Tipo actividad: (Taller)

Tiempo de ejecución: 2 horas

PLANTEAMIENTO DE LA SESIÓN

Materiales

Como ejercicio de aplicación se propone un taller de estadística para aplicar los conocimientos adquiridos. En este taller se realizará el análisis de algunas variables aleatorias del dataset propuesto y se revisarán sus funciones de probabilidad.

Utilizando el dataset Iris (disponible en <https://archive.ics.uci.edu/dataset/53/iris>) realice las siguientes actividades:

1. Cargue el conjunto de datos a una hoja de Excel o Google sheets.
2. Utilizando filtros, separe los tres tipos de iris (setosa, versicolor y virgínica), Guarde en hojas diferentes los datos asociados a cada especie.
3. Para cada uno de los subconjuntos, grafique el histograma de la longitud del sépalo. También calcule la media y la desviación estándar.
4. Utilizando la ecuación que define la curva gaussiana, reemplace los valores de la media y de la desviación estándar y compare si la distribución se parece al histograma encontrado.



5. Repita los puntos 3 y 4 pero con el dato de ancho del sépalo. ¿Se asemejan las distribuciones? ¿Por qué?

6. Repita los puntos 3 y 4 para las variables longitud del pétalo y ancho del pétalo. Compare los resultados de las variables aleatorias con las curvas gaussianas encontradas.