

Práctica 3 de Estadística para Data Science

Problema 1 *Estamos esperando un tren que llegará en los próximos 100 minutos pero desconocemos el momento exacto. Definimos la variable aleatoria $X = \text{“tiempo de espera en minutos”}$. Realiza las siguientes tareas:*

- i) Considera que X es una variable aleatoria uniforme en el intervalo $[0, 100]$. ¿Cuánto vale $E(X)$ y $Var(X)$?*
- ii) Simula los tiempos de espera de una muestra aleatoria de personas con tamaño muestral $n = 30$ y calcula la media muestral, \bar{X} y la varianza muestral S^2 .*
- iii) ¿Cuanto valen $E(\bar{X})$ y $Var(\bar{X})$ para $n = 30$? Y si aumentáramos a $n = 100$?*
- iv) Simula 500 muestras aleatorias de tamaño $n = 30$ de la variable X . Recomendación: Introduce las muestras en una matriz de 30 filas y 500 columnas.*
- v) Calcula las 500 medias muestrales, es decir $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_{500}$. Recomendación: Utiliza el comando `colMeans`.*
- vi) Calcula la media y varianza de las 500 medias muestrales y comparalas con los valores obtenidos en el tercer apartado.*
- vii) Repite los tres pasos anteriores para 500 muestras aleatorias de tamaño $n = 100$.*
- viii) Representa las 500 medias de tamaño $n = 30$ y las 500 medias de tamaño $n = 100$ mediante dos histogramas en una gráfica de dos paneles.*
- ix) Superpón a cada histograma la función densidad normal con $\mu = E(X)$ y $\sigma^2 = Var(X)/n$.*
- x) Comenta el resultado.*