

---

# Programación para el análisis de datos

## Maestría en Ciencia de Datos - 2023

1- Volvemos al vuelo de la mosca y a su modelado...

- (a) Para el modelo A planteado escriba en un archivo llamado “vuelo\_mosca.csv” el índice de la simulación, las coordenadas (x,y) de las posiciones donde la mosca hace un cambio de vuelo, el tiempo en el que ocurrió ese cambio, y algún nombre relacionado con las distribuciones que está usando. Por ejemplo, escribir (simul,tiempo,coordenadax,coordenaday, distri\_angulos,distri\_tiempos) para cada uno de los eventos que ocurren dentro de la “mesa”.
- (b) Entregar en forma prolija el código del modelo teniendo en cuenta que el cuerpo principal del código no ocupe más de media carilla. Las funciones que hayan creado no cuentan como parte del cuerpo principal pero también serán entregadas.
- (c) Correr el modelo para dos distribuciones distintas de ángulos/tiempos/velocidades.
- (d) Cargar los datos “vuelo\_mosca.csv”.
  - i. Estudiar la distribución de la v.a. número de cambios de dirección (de vuelo) antes de salir de la mesa. ¿ Se parece a alguna distribución conocida? ¿Cuál sería el o los parámetros de esa distribución? Haga algún gráfico para comparar la distribución de los datos con esa distribución conocida.
  - ii. Mostrar de alguna manera el efecto de la distribución sobre la v.a. número de cambios de dirección.
  - iii. **Parte individual:** Escribir 4 preguntas relacionadas con los datos creados (o sea sobre el vuelo) y responderlas utilizando alguna de las técnicas de estadística descriptiva.

2.- Sea  $X$  una variable aleatoria Exponencial( $\lambda = 2$ ), y definimos las funciones

$$H(b) = \mathbb{E}((X - b)^2), \quad \text{y} \quad G(b) = \mathbb{E}(|X - b|).$$

- (a) En una misma figura graficar las funciones  $H(b)$ , y  $G(b)$  en función de  $b$  para valores de  $b$  entre 0 y 2 con paso 0.001.
- (b) Determinar aproximadamente los valores  $b_H^*$ , y  $b_G^*$  que minimizan la función  $H(b)$ , y  $G(b)$  respectivamente.
- (c) Tomar una muestra aleatoria de 10000 v.a. independientes exponenciales con  $\lambda = 2$  y calcular
  - i. la media(o promedio).
  - ii. la mediana.
- (d) Comparar la media anterior con  $b_H^*$ , y la mediana con  $b_G^*$ . ¿Son parecidos? Interprete.
- (e) Sea la función  $M_k(b) = \mathbb{E}((X - b)^k)$ . Calcular el mínimo de esta función para  $k = 2, 4, 6, 8$ , y 10. Graficar en el eje de las ordenadas estos mínimos y en el eje de las abscisas el valor  $k$ . Interpretar el gráfico, y agregarle a este mismo gráfico la función de densidad de la exponencial.