1	2	3	4	5	Calificación

R y Probabilidad

Maestría en Ciencia de Datos - 2022

En color negro se encuentra la parte grupal del examen, y en color azul la parte individual.

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

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

(a) En una misma figura graficar las función $H(b)$ y $G(b)$ en función de b para valores de b entre 0 y 2 con paso 0.01.

(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.

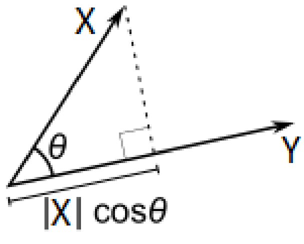
(d) Comparar la media anterior con b_G^*, y la mediana con b_H^*. ¿Son parecidos? Interprete.

2.- A partir de la secuencia $X_1, X_2, \ldots, X_n$ de variables aleatorias i.i.d. con distribución Normal(0,1) se construye el vector aleatorio $\vec{X} = (X_1, X_2, \ldots, X_n)$.

(a) ¿Cómo es la matriz de covarianza, Σ, de $\vec{X}$?

(b) El ángulo, θ, entre dos vectores $\vec{X}$ y $\vec{Y}$ pertenecientes al mismo espacio está definido por la siguiente ecuación,

$$X_1Y_1 + X_2Y_2 + \ldots X_nY_n = |X||Y|cos(\theta).$$



Donde $|Z| = \sqrt{Z_1^2 + Z_2^2 + \ldots Z_n^2}$ es la longitud del vector Z.

i. Considere el caso $n = 2$ (Normal bivariada). Utilizando la Ley de Grandes Números estimar la esperanza del ángulo, $\mathbb{E}(\theta)$, entre dos vectores $\vec{X_1}$ y $\vec{X_2}$ independientes.

ii. Idem i, pero ahora para $n = 3, 4, 5, 10, 20, 30, 100$. Graficar los resultados, es decir la estimación de la esperanza del ángulo en función de n. Los espacios en dimensiones grandes son raros, ¿no?

(c) ¿Qué otra cosa le parece interesante de estudiar cuando uno tiene uno o dos vectores en $\mathbb{R}^n$?

3.- Una población comienza con un miembro y evoluciona en tiempo discreto. En un tiempo $t = 1$ se divide en dos con probabilidad p o bien muere con probabilidad $1 - p$. Si se divide, entonces cada uno de sus hijos se comportan independientemente con las mismas dos alternativas en el tiempo $t = 2$, y así sucesivamente en tiempo discreto.

(a) Utilizando Ley de Grandes Números mostrar que

$$X_t\text{=el número de miembros en el tiempo } t\text{, verifica } \mathbb{E}\left(X_t\right) = (2p)^t.$$

Es decir, graficar una aproximación de $\mathbb{E}\left(X_t\right)$ como función de t para distintos valores de p, interprete.

(b) ¿Qué otra cosa le parece interesante de estudiar en este modelo?

4.- A partir de los datos de nacimientos, resolver los ejercicios 3,4,5,6 y 7 de la guía 4.

5.- A partir de los datos de nacimientos:

(a) Piense 7 preguntas no relacionadas con el problema de regresión que le gustaría responder con estos datos [1].

(b) Clasifique a las preguntas según si está estudiando una sola variable o la dependencia entre dos o más variables.

(c) Intente dar respuesta a las preguntas [2] del item (a) realizando algún gráfico.

[1]-"Esta bien perder con el enemigo, pero no con el miedo". Señor Miyagi

[2]"La respuesta sólo ha de ser importarte cuando la pregunta es correcta". Señor Miyagi.