

Examen 29/10/2024 D

1. Una potente empresa tiene en el mercado 3 redes sociales, TíTlac, Instakilo y TíyYo. En el Consejo de Administración de cada una, se ha planteado un debate ético sobre la conveniencia de adoptar medidas beligerantes en la lucha contra el cambio climático frente a la posición negacionista. En TíTlac, la votación ha sido: 9 Beligerantes y 10 Negacionistas. En Instakilo, 14 Beligerante y 15 Negacionistas y en TíyYo empate a 13. Por tanto la empresa, escuchadas las posiciones de las tres redes (2 negacionistas y 1 empate) optará por la posición negacionista. Si embargo, en una maniobra de última hora, se decide repetir la votación después de transferir por sorteo un miembro del consejo de administración de TíTlac a Instakilo y posteriormente una persona de Instakilo a TíyYo (también por sorteo). ¿Cuál es la probabilidad de que TíyYo tenga mayoría beligerante? ¿Cuál es la probabilidad de que si descubrimos al azar un miembro de TíyYo, éste sea de posición beligerante? Por otro lado, si sabemos que finalmente ha salido mayoría beligerante en TíyYo, ¿cuál es la probabilidad respectiva en cada caso de que TíTlac sea beligerante, negacionista o neutral?

2. El peso que se asigna a la entrada de tipo A de una determinada red neuronal sigue una distribución normal  $N(45, 2, 1, 3)$ . En las entradas de tipo B, una  $N(61, 4, 1, 1)$ . Calcular:

a) Probabilidad de que la suma de los pesos de 15 entradas de tipo A no supere el umbral de 697.0.

b) ¿Cuál es el peso acumulado de las entradas de una red neuronal con 15 entradas de tipo A y 7 entradas de tipo B que está por debajo del percentil 45?

c) Si tenemos dos redes neuronales de las anteriores, Red1 y Red2, ¿Cuál es la probabilidad de que el peso acumulado de las entradas en la primera sea, al menos, una unidad mayor que el de la segunda? Y, ¿la probabilidad de que tengan el mismo peso acumulado?

3. Dada la f.d.p de la v.a. X

$$f(x) = \begin{cases} k(1 - x^2) & \text{si } |x| \leq 8 \\ 0 & \text{si } |x| > 8 \end{cases}$$

Determina el valor de k, calcula  $E[X]$  y  $Var[X]$ .

4. En la construcción de un *Data Center* se necesitan miles de minicomponentes. Para el proveedor A, se está diseñando un plan de control de calidad de forma que tomando un lote de n unidades, se rechace si se observa alguna minicomponente defectuosa. Determina n para que, si el lote tiene un 9 por mil (o más) de unidades defectuosas, la probabilidad de aceptarlo sea menor del ~~5~~ 3%.

El proveedor B, sirve lotes de n=100 minicomponentes y sabemos que son defectuosos un 5 por mil de los minicomponentes. Si en este caso seguimos la regla de rechazar el lote si salen más de 3 minicomponentes defectuosos, ¿hoy nos ha enviado 10 lotes? ¿Cuál es la probabilidad de rechazar hoy 2 lotes? ¿Mañana llegarán 140 lotes, ¿Cuál será la probabilidad de rechazar 5 lotes como máximo?

5. Sabiendo que el número de averías en un día en el servidor A sigue una distribución Poisson  $X \sim Poi(6)$  y análogamente en el servidor B,  $Y \sim Poi(3)$  y que son independientes, calcula: a) la probabilidad de que A tenga 6 averías en un día sabiendo que entre ambos han tenido 8. b) Que A no tenga averías en una semana c) Que B tenga menos de 11 averías en una semana.

2 Puntos por ejercicio.

- 1.- Se ha detectado que los ciberataques que sufre la UA llegan de cuatro continentes (A,B,C y D). Se sospecha que la proporción de ataques es 3:1:1:1 respectivamente. Se han recogido los siguientes intentos de ataque informático en un mes: A: 325; B:110; C: 101; D: 99. ¿Hay motivos para refutar la sospecha? ¿Por qué?
- 2.- Con el comando `random.normalvariate(5,3)` de la librería `random` de `Python`, hemos generado 20 valores de una normal con el siguiente resultado:

7.71, 4.23, 7.84, 1.61, 7.42, 6.3, 8.66, 1.81, 11.28, 3.98, 4.32, 5.13, 4.67, 3.68, 6.02, 2.72, 5.82, 5.15, 6.11, 4.61

¿Tenemos razones para pensar que el comando funciona mal? ¿Por qué?

Nota:  $\bar{X} = 5.4535$   $S = 2.3$

- 3.- Bajo la suposición de que los datos anteriores si sean de la distribución normal,

a) Da un intervalo de confianza al 95 % de la media de la distribución.

b) ¿Podemos pensar que la media es superior a 5?

c) ¿Podemos pensar que la varianza no es 5.3?

- 4.- En un experimento sobre dos muestras de betatesters se les hace jugar hasta que no puedan aguantar más a dos versiones distintas del juego para medirles el umbral de soporte. Los datos obtenidos (en tiempo de aguarante) están en la siguiente tabla:

	Versión 1	Versión 2
n	14	10
$\bar{x}$	16.2	14.9
$s^2$	12.7	26.4

a) ¿Muestran los datos que hay evidencia de que hay diferencias en el tiempo que se soporta el juego?

b) Si debes realizar uno o varios contrastes, utiliza el *p-valor* en alguno de ellos.

- 5.- A la hora de plantear la resolución de problemas a un LLM, se han detectado los siguientes porcentajes de errores en función de si se planteaba a GPT4, Gemini, Claude v1, Cohere, Copilot:

	GPT4	Gemini	Claude	Cohere	Copilot
$X_i$	24	33	24	50	32
	37	20	40	20	62
	22	28	63	30	40
	55	12	18	13	15
	23	17	62	42	26
	38	17	30	28	37
	46	57	38	17	52
	25	42	23	73	12
	25	25	37	25	16
	23	63	26	22	25
$X_i$	31.27	31.73	35.09	33.73	31.55
$S_i^2$	11.01	16.56	14.86	17.42	15.55

La media global es  $\bar{X} = 32.67$  y la varianza global es  $S = 14.744$ . Hemos comprobado también que las varianzas son homogéneas. ¿Podemos afirmar si hay diferencias entre los modelos planteados? ¿Por qué?

## 2 Puntos por ejercicio.