

PAUTA EXAMEN

P1. a) 1) Llamemos A y B a los amigos que están peleados. La cantidad buscada es:

$$\binom{6}{4} + \binom{6}{4} + \binom{6}{5} = 2 \frac{6!}{4!2!} + \frac{6!}{5!1!} = 2 \frac{6 \cdot 5}{2} + 6 = 36.$$

El primer término se obtiene al seleccionar a A y escoger los otros 4 amigos entre el resto, excluyendo a B; el segundo término es lo análogo pero seleccionando a B. El último término corresponde a dejar fuera a A y B, y se escogen 5 amigos entre los 6 restantes.

2) Llamemos C y D a los amigos que solo van si los invitan a ambos. La cantidad buscada es:

$$\binom{6}{3} + \binom{6}{5} = \frac{6!}{3!3!} + \frac{6!}{5!1!} = \frac{6 \cdot 5 \cdot 4}{3 \cdot 2} + 6 = 26.$$

El primer término corresponde a seleccionar a C y D, y escoger a los otros 3 entre los 6 amigos restantes; el segundo término se obtiene al dejar fuera a C y D y escoger los 5 amigos entre los 6 restantes.

b) Llamemos A_i al evento en que el avión está en la zona i , y E_i al evento en que el avión es encontrado en la zona i , para $i = 1, 2, 3$.

1) Sea E el evento en que el avión es encontrado. Los E_i son disjuntos, por lo cual tenemos:

$$\mathbb{P}(E) = \mathbb{P}(E_1 \cup E_2 \cup E_3) = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(E_3).$$

Calculemos $\mathbb{P}(E_i)$ condicionando en A_i :

$$\mathbb{P}(E_i) = \mathbb{P}(E_i|A_i)\mathbb{P}(A_i) + \mathbb{P}(E_i|A_i^c)\mathbb{P}(A_i^c) = \alpha_i \frac{1}{3} + 0 \frac{2}{3} = \frac{\alpha_i}{3}.$$

Luego, la probabilidad de encontrar el avión es $\mathbb{P}(E) = (\alpha_1 + \alpha_2 + \alpha_3)/3$.

2) Queremos calcular $\mathbb{P}(A_i|E_1^c)$ para $i = 1, 2, 3$. Por regla de Bayes, tenemos para $i = 1$:

$$\mathbb{P}(A_1|E_1^c) = \mathbb{P}(E_1^c|A_1) \frac{\mathbb{P}(A_1)}{\mathbb{P}(E_1^c)} = (1 - \alpha_1) \frac{1/3}{1 - \alpha_1/3} = \frac{1 - \alpha_1}{3 - \alpha_1}.$$

Para $i \neq 1$, se cumple que la probabilidad de no encontrar al avión en la zona 1, dado que está en la zona i , es 1. Entonces, por regla de Bayes:

$$\mathbb{P}(A_i|E_1^c) = \mathbb{P}(E_1^c|A_i) \frac{\mathbb{P}(A_i)}{\mathbb{P}(E_1^c)} = 1 \frac{1/3}{1 - \alpha_1/3} = \frac{1}{3 - \alpha_1}.$$

P2. a) Para obtener C imponemos que la densidad integre 1. Usando integración por partes:

$$1 = C \int_0^\infty x e^{-x/\theta} dx = C \left(-x\theta e^{-x/\theta} \Big|_0^\infty + \int_0^\infty \theta e^{-x/\theta} dx \right) = C \left(0 + -\theta^2 e^{-x/\theta} \Big|_0^\infty \right) = C\theta^2,$$

y despejando obtenemos $C = 1/\theta^2$.

b) Calculemos:

$$\mathbb{E}(X) = \frac{1}{\theta^2} \int_0^\infty x^2 e^{-x/\theta} dx = \frac{1}{\theta^2} \left(-x^2 \theta e^{-x/\theta} + \int_0^\infty 2x \theta e^{-x/\theta} dx \right) = 2\theta \int_0^\infty \frac{1}{\theta^2} x e^{-x/\theta} dx,$$

donde la última integral vale 1 pues el integrando es la densidad de X . Se obtiene entonces $\mathbb{E}(X) = 2\theta$.

c) Calculemos $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \frac{1}{\theta^2} \int_0^\infty x^3 e^{-x/\theta} dx = \frac{1}{\theta^2} \left(-x^3 \theta e^{-x/\theta} + \int_0^\infty 3x^2 \theta e^{-x/\theta} dx \right) = 3\theta \int_0^\infty \frac{1}{\theta^2} x^2 e^{-x/\theta} dx,$$

donde la última integral corresponde a la esperanza de X . Tenemos entonces que $\mathbb{E}(X^2) = (3\theta)(2\theta) = 6\theta^2$. Luego, la varianza es:

$$\text{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 6\theta^2 - (2\theta)^2 = 2\theta^2.$$

d) Si x_1, \dots, x_n son los valores que toma la muestra (necesariamente todos no negativos), la verosimilitud es

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n x_i \frac{1}{\theta^2} e^{-x_i/\theta} = \theta^{-2n} e^{-\sum x_i/\theta} \prod_{i=1}^n x_i.$$

Para encontrar el estimador de máxima verosimilitud, derivamos con respecto a θ e igualamos a cero. Es decir:

$$0 = -2n\theta^{-2n-1} e^{-\sum x_i/\theta} \prod_{i=1}^n x_i + \theta^{-2n} e^{-\sum x_i/\theta} \prod_{i=1}^n x_i \frac{1}{\theta^2} \sum x_i.$$

Simplificando la exponencial y el producto, y multiplicando por θ^{2n+2} , sigue que $0 = -2n\theta + \sum x_i$. Despejando θ , obtenemos que el estimador de máxima verosimilitud es $\hat{\theta} = \bar{X}/2$. Es directo que $\mathbb{E}(\hat{\theta}) = \mathbb{E}(\bar{X})/2 = \mathbb{E}(X)/2 = 2\theta/2 = \theta$, es decir, el estimador es insesgado.

e) Siguiendo la indicación, el estadístico $Z = (\bar{X} - 2\theta)/(\hat{\theta}\sqrt{2/n})$ es aproximadamente normal estándar. Buscamos un intervalo para Z centrado en 0, es decir:

$$1 - 5\% = \mathbb{P}(Z \in [-c, c]) \approx \mathbb{P}(\mathcal{N}(0, 1) \in [-c, c]) = 1 - 2\mathbb{P}(\mathcal{N}(0, 1) > c).$$

Es decir, buscamos c tal que $\mathbb{P}(\mathcal{N}(0, 1) > c) = 2,5\%$, obteniendo $c = 1,96$ (ver tabla). Teniendo en cuenta que $\bar{X} = 2\hat{\theta}$, el intervalo para θ queda:

$$-c \leq Z \leq c \Leftrightarrow -c \leq \frac{\bar{X} - 2\theta}{\hat{\theta}\sqrt{2/n}} \leq c \Leftrightarrow -c \leq \frac{2 \cdot 10 - 2\theta}{10\sqrt{2/50}} \leq c \Leftrightarrow -c \leq \frac{2 \cdot 10 - 2\theta}{2} \leq c,$$

es decir, $10 - c \leq \theta \leq 10 + c$. Luego, el intervalo buscado es $[8,04, 11,96]$. Falta demostrar la indicación: si llamamos $\mu = 2\theta$ y $\sigma^2 = 2\theta^2$ a la esperanza y varianza de X , respectivamente, se tiene que

$$Z = \frac{\bar{X} - 2\theta}{\hat{\theta}\sqrt{2/n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \frac{\sigma/\sqrt{n}}{\hat{\theta}\sqrt{2/n}}.$$

Por TLC, sabemos que la primera fracción en la última expresión converge en distribución a una normal estándar cuando $n \rightarrow \infty$. Por otro lado, por la LFGN sabemos que \bar{X} converge casi seguramente a $\mu = 2\theta$, es decir, $\hat{\theta} = \bar{X}/2$ converge casi seguramente a $\theta = \sigma/\sqrt{2}$. Esto prueba que la segunda fracción converge casi seguramente a 1. Por propiedad vista en cátedra, se concluye que Z converge en distribución al límite de la primera fracción, es decir, a una normal estándar.

- P3.** a) La hipótesis nula es que no hay preferencia de pistas. Es decir, $p_i = p_i^0 = 1/4$, $\forall i = 1, 2, 3, 4$, donde p_i es la probabilidad de que un conductor escoja la pista i . La hipótesis alternativa es que algún p_i es distinto de $1/4$. Los estimadores de los p_i son las frecuencias observadas, es decir, $\hat{p}_1 = 294/1000$, $\hat{p}_2 = 276/1000$, $\hat{p}_3 = 238/1000$ y $\hat{p}_4 = 192/1000$. Trabajamos con el estadístico

$$\Delta = n \sum_{i=1}^4 \frac{(\hat{p}_i - p_i^0)^2}{p_i^0},$$

el cual, bajo H_0 , sabemos que se distribuye aproximadamente como una chi-cuadrado con $4 - 1 = 3$ grados de libertad. Buscamos c tal que $0,05 = \mathbb{P}(\Delta > c | H_0) \approx \mathbb{P}(\chi_3^2 > c)$. Mirando una tabla, obtenemos $c = 7,81$. Por otro lado, el valor que toma el estadístico Δ es:

$$\begin{aligned} \Delta &= 1000 \left[\frac{(\frac{294}{1000} - \frac{250}{1000})^2}{\frac{250}{1000}} + \frac{(\frac{276}{1000} - \frac{250}{1000})^2}{\frac{250}{1000}} + \frac{(\frac{238}{1000} - \frac{250}{1000})^2}{\frac{250}{1000}} + \frac{(\frac{192}{1000} - \frac{250}{1000})^2}{\frac{250}{1000}} \right] \\ &= \frac{1}{250} [44^2 + 26^2 + 12^2 + 58^2] = \frac{1}{250} [1936 + 676 + 144 + 3364] = \frac{6120}{250} = 24,48. \end{aligned}$$

Como 24,48 es mayor que el c encontrado, debemos rechazar la hipótesis nula. Es decir, los conductores sí tienen preferencia por alguna de las pistas.

- b) 1) Si llamamos $Z = XY$, entonces X_1Y_1, \dots, X_nY_n es una muestra aleatoria simple de Z . Por LFGN, el promedio de esta muestra converge a la esperanza de Z , es decir, $\frac{1}{n} \sum_{i=1}^n X_iY_i$ converge casi seguramente a $\mathbb{E}(Z) = \mathbb{E}(XY)$. El mismo argumento sirve para probar que $\frac{1}{n} \sum_{i=1}^n X_i^2$ converge casi seguramente a $\mathbb{E}(X^2)$.
- 2) Sabemos que \hat{a}_n es la covarianza muestral sobre la varianza muestral de X , es decir:

$$\hat{a}_n = \frac{\frac{1}{n} \sum_{i=1}^n X_iY_i - \bar{X}\bar{Y}}{\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2}.$$

Por la parte anterior, se tiene que $\frac{1}{n} \sum_{i=1}^n X_iY_i$ converge a $\mathbb{E}(XY)$, y que $\frac{1}{n} \sum_{i=1}^n X_i^2$ converge a $\mathbb{E}(X^2)$. Además, sabemos por LFGN que \bar{X} y \bar{Y} convergen a $\mathbb{E}(X)$ y $\mathbb{E}(Y)$ (todas las convergencias son casi seguramente, cuando $n \rightarrow \infty$). Por lo tanto, \hat{a}_n converge casi seguramente a

$$\frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\mathbb{E}(X^2) - \mathbb{E}(X)^2} = \frac{\text{cov}(X, Y)}{\text{var}(X)}.$$

Por último, sabemos que $\hat{b}_n = \bar{Y} - \hat{a}_n\bar{X}$, donde cada término converge. Por lo tanto, \hat{b}_n converge casi seguramente a

$$\mathbb{E}(Y) - \mathbb{E}(X) \frac{\text{cov}(X, Y)}{\text{var}(X)}.$$