



PAUTA EXAMEN

- P1.** a) 1) La cantidad buscada es $30!$. Esta expresión corresponde a la cantidad de formas de ordenar a las personas en las sillas disponibles.
- 2) La cantidad buscada es $\binom{30}{6,6,6,6,6}$. Esta cantidad se obtiene al separar a las 30 personas en 5 grupos de tamaño 6, correspondientes a las capacidades de las mesas disponibles.
- 3) La cantidad buscada es

$$5 \cdot \binom{27}{3,6,6,6,6} + 5 \cdot 4 \cdot \binom{27}{4,5,6,6,6} + \binom{5}{3} \cdot \binom{27}{5,5,5,6,6}.$$

El primer término de la suma anterior corresponde a la cantidad de formas de escoger 1 mesa en la que habrá 3 sillas vacías, y multiplicar por la cantidad de formas de repartir a las 27 personas en grupos de tamaños 3, 6, 6, 6 y 6, correspondientes a las capacidades de las mesas una vez fijados los puestos vacíos. El segundo término es lo análogo, pero escogiendo una mesa con 2 puestos vacíos y otra con 1 puesto vacío. El último término corresponde a que en cada mesa haya a lo más un puesto vacío.

- b) Sean los eventos J_A y J_B correspondientes a que el jugador que lanza es A ó B, respectivamente. Sea X la variable que denota la distancia del dardo al centro del blanco en el primer lanzamiento.

- 1) Nos piden calcular la probabilidad de que ocurra J_A dado que $X < r/4$. Por regla de Bayes, tenemos:

$$\mathbb{P}(J_A|X < r/4) = \frac{\mathbb{P}(X < r/4|J_A)\mathbb{P}(J_A)}{\mathbb{P}(X < r/4)}.$$

Dado J_A , sabemos que X es una uniforme en $[0, r/2]$, por lo cual la primera probabilidad en el numerador es $1/2$. Como la elección del jugador es al azar, la segunda probabilidad del numerador también es $1/2$. La probabilidad del denominador se calcula utilizando la regla de probabilidades totales:

$$\mathbb{P}(X < r/4) = \mathbb{P}(X < r/4|J_A)\mathbb{P}(J_A) + \mathbb{P}(X < r/4|J_B)\mathbb{P}(J_B) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{3}{8},$$

donde hemos utilizado que X es uniforme en $[0, r]$ cuando el jugador que lanza es B. Obtenemos entonces:

$$\mathbb{P}(J_A|X < r/4) = \frac{1/2 \cdot 1/2}{3/8} = \frac{2}{3}.$$

- 2) Sea Y la variable que denota la distancia del dardo al centro del blanco correspondiente al segundo lanzamiento. Nos piden calcular la probabilidad de que $Y < r/4$ dado que $X < r/4$. Por definición de probabilidad condicional, tenemos que:

$$\mathbb{P}(Y < r/4|X < r/4) = \frac{\mathbb{P}(Y < r/4, X < r/4)}{\mathbb{P}(X < r/4)},$$

donde la probabilidad del denominador ya fue calculado. Para el numerador, utilizamos nuevamente la regla de probabilidades totales:

$$\begin{aligned}\mathbb{P}(Y < r/4, X < r/4) &= \mathbb{P}(Y < r/4, X < r/4|J_A)\mathbb{P}(J_A) + \mathbb{P}(Y < r/4, X < r/4|J_B)\mathbb{P}(J_B) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{2} = \frac{5}{32},\end{aligned}$$

donde hemos utilizado que los lanzamientos son independientes, una vez que se fija el jugador que lanza. Se obtiene entonces que:

$$\mathbb{P}(Y < r/4|X < r/4) = \frac{5/32}{3/8} = \frac{5}{12}.$$

P2. a) La integral de la densidad debe valer 1, por lo cual:

$$1 = \int_{-\infty}^{\infty} f_X(x)dx = \int_m^{\infty} Cx^{-(\alpha+1)}dx = C \left. \frac{x^{-\alpha}}{-\alpha} \right|_m^{\infty} = C \frac{m^{-\alpha}}{\alpha}.$$

Además, conocemos el valor de la esperanza de X , con lo que tenemos:

$$\frac{\alpha}{\alpha-1} = \mathbb{E}(X) = \int_{-\infty}^{\infty} xf_X(x)dx = \int_m^{\infty} xCx^{-(\alpha+1)}dx = C \left. \frac{x^{-\alpha+1}}{-\alpha+1} \right|_m^{\infty} = C \frac{m^{-\alpha+1}}{\alpha-1}.$$

Obtenemos entonces dos ecuaciones para C y m como incógnitas. Haciendo el cociente se obtiene $m = 1$ y despejando este valor en cualquiera de las ecuaciones se llega a $C = \alpha$.

b) Calculemos $\text{var}(X)$ y veamos cuándo se obtiene ∞ .

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_1^{\infty} x^2 \alpha x^{-(\alpha+1)}dx = \alpha \int_1^{\infty} x^{-(\alpha-1)}dx,$$

donde sabemos que la integral anterior es finita si y sólo si $(\alpha-1) > 1$, es decir, cuando $\alpha > 2$. Para un tal α , tenemos entonces:

$$\mathbb{E}(X^2) = \alpha \left. \frac{x^{-\alpha+2}}{-\alpha+2} \right|_1^{\infty} = \frac{\alpha}{\alpha-2}.$$

Por lo tanto, la varianza de X para $\alpha \leq 2$ es ∞ , mientras que para $\alpha > 2$ es

$$\begin{aligned}\text{var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{\alpha}{\alpha-2} - \frac{\alpha^2}{(\alpha-1)^2} = \frac{\alpha^3 - 2\alpha^2 + \alpha - (\alpha^3 - 2\alpha^2)}{(\alpha-1)^2(\alpha-2)} \\ &= \frac{\alpha}{(\alpha-1)^2(\alpha-2)}.\end{aligned}$$

c) Sea $Y = \ln(X)$. Calculemos la distribución acumulada de Y :

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(\ln(X) \leq y) = \mathbb{P}(X \leq e^y).$$

Si $y < 0$, entonces $e^y < 1$ y la probabilidad anterior vale 0, pues sabemos que X toma valores en el intervalo $[1, \infty)$. Para $y \geq 0$, tenemos:

$$F_Y(y) = \int_1^{e^y} \alpha x^{-(\alpha+1)}dx = -x^{-\alpha} \Big|_1^{e^y} = -(e^y)^{-\alpha} + 1 = 1 - e^{-\alpha y}.$$

Derivando, y teniendo en cuenta que $F_Y(y) = 0$ para $y < 0$, obtenemos que $f_Y(y) = \alpha e^{-\alpha y} \mathbb{1}_{[0, \infty)}(y)$, es decir, Y tiene distribución exponencial de parámetro α .

d) La función de verosimilitud corresponde a:

$$L(x_1, \dots, x_n) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \alpha x_i^{-(\alpha+1)} \mathbb{1}_{[1, \infty)}(x_i) = \alpha^n \left[\prod_{i=1}^n x_i \right]^{-(\alpha+1)},$$

donde hemos utilizado que todas las indicatrices valen 1, pues los x_i corresponden a la realización de X , la cual sabemos que toma valores en $[1, \infty)$. Para maximizar lo anterior con respecto a α , tomamos $\ln(\cdot)$, derivamos e igualamos a 0:

$$0 = \frac{d}{d\alpha} \ln(L(x_1, \dots, x_n)) = \frac{d}{d\alpha} \left(n \ln(\alpha) - (\alpha + 1) \sum_{i=1}^n \ln(x_i) \right) = \frac{n}{\alpha} - \sum_{i=1}^n \ln(x_i).$$

Despejando α , se obtiene entonces $\hat{\alpha} = n / \sum_{i=1}^n \ln(X_i)$.

e) Sean $Y_i = \ln(X_i)$, para $i = 1, \dots, n$. Por lo hecho anteriormente, sabemos que Y_i es una variable exponencial de parámetro α . En términos de los Y_i , se tiene que $\hat{\alpha} = n / \sum_{i=1}^n Y_i = 1/\bar{Y}$. Por la LFGN, sabemos que $\bar{Y} \rightarrow \mathbb{E}(Y_1) = 1/\alpha$ casi seguramente cuando $n \rightarrow \infty$. Luego $\hat{\alpha} \rightarrow 1/(1/\alpha) = \alpha$ casi seguramente cuando $n \rightarrow \infty$, como deseábamos.

P3. a) Sean X_1, \dots, X_n los puntajes de las personas, con $n = 16$. Consideremos el estadístico

$$T = \frac{\bar{X} - \mu}{\sqrt{s^2/n}},$$

el cual sabemos que tiene distribución de t -student con $n - 1 = 15$ grados de libertad, donde s^2 es el estimador insesgado de la varianza. Buscamos un intervalo simétrico con respecto a T al nivel 95 %, es decir, imponemos que

$$95 \% = \mathbb{P}(-c \leq T \leq c) = 1 - 2\mathbb{P}(T > c) \Rightarrow \mathbb{P}(T > c) = 2,5 \%,$$

donde hemos usado la simetría de la distribución t -student. Usando la tabla de la t -student, obtenemos $c = 2,131$. El intervalo buscado se obtiene despejando μ :

$$-c \leq T \leq c \Leftrightarrow -c \leq \frac{\bar{X} - \mu}{\sqrt{s^2/n}} \leq c \Leftrightarrow \bar{X} - c \frac{\sqrt{s^2}}{\sqrt{n}} \leq \mu \leq \bar{X} + c \frac{\sqrt{s^2}}{\sqrt{n}}.$$

Reemplazando los valores de n , de \bar{X} y $\sqrt{s^2}$ obtenidos en la muestra, y de c ya calculado, se tiene entonces que el intervalo de confianza para μ al nivel buscado es:

$$\left[540 - 2,131 \cdot \frac{50}{\sqrt{16}}, 540 + 2,131 \cdot \frac{50}{\sqrt{16}} \right] = \left[540 - \frac{50 \cdot 2,131}{4}, 540 + \frac{50 \cdot 2,131}{4} \right].$$

b) Consideremos las variables X_1, \dots, X_n , donde X_i vale 1 si el i -ésimo conductor de la muestra olvidó su licencia, y 0 si no. Notemos que X_i es una variable Bernoulli con parámetro p desconocido, donde p es la probabilidad de que un conductor olvide su licencia. Consideremos las hipótesis $H_0 : p = 2 \%$ y $H_1 : p < 2 \%$.

1) Notemos que la esperanza y varianza de cada X_i es p y $p(1 - p)$, respectivamente. Consideremos el estadístico

$$Z = \frac{\bar{X} - 2 \%}{\sqrt{2 \%(1 - 2 \%) / n}},$$

el cual, bajo H_0 , sabemos que tiene distribución aproximadamente normal estándar, por el TLC. El valor obtenido de este estadístico, debido a que los 100 conductores de la muestra traían su licencia, corresponde a

$$z = \frac{0 - 2\%}{\sqrt{2\%(1 - 2\%)/n}} = \frac{-\frac{1}{50}}{\sqrt{(\frac{1}{50} \cdot \frac{49}{50})/100}} = \frac{-10 \cdot \frac{1}{50}}{\frac{7}{50}} = -\frac{10}{7} \approx -1,43.$$

El p -valor corresponde a la probabilidad, bajo H_0 , de obtener un valor al menos tan extremo como el observado en la muestra, donde, debido a la forma de la hipótesis H_1 , esto significa obtener $Z < z$. Es decir:

$$p\text{-valor} = \mathbb{P}(Z < -1,43|H_0) \approx \mathbb{P}(\mathcal{N}(0, 1) < -1,43) = \mathbb{P}(\mathcal{N}(0, 1) > 1,43) = 7,64\%,$$

donde la última igualdad se obtiene mirando la tabla de la normal.

- 2) La región de rechazo es de la forma $Z < -c$ para una cierta constante c que se escoge de modo de obtener el nivel deseado:

$$2,28\% = \mathbb{P}(Z < -c) \approx \mathbb{P}(\mathcal{N}(0, 1) < -c) = \mathbb{P}(\mathcal{N}(0, 1) > c),$$

lo cual, mirando la tabla, implica que $c = 2,0$. Luego, la afirmación no se rechaza cuando $Z \geq c$. Si suponemos que todos los conductores adicionales también traen su licencia, esto significa que \bar{X} se mantiene en 0. Luego:

$$-2,0 \leq Z = \frac{0 - 2\%}{\sqrt{2\%(1 - 2\%)/n}} = -\frac{\sqrt{n}}{7} \Rightarrow n \leq 7^2 \cdot 2^2 = 196.$$

Luego, aún si los siguientes 96 conductores traen su licencia, no hay suficiente evidencia para rechazar la afirmación al nivel deseado.