



PAUTA CONTROL 3

- P1.** a) Sean X_1, \dots, X_n variables que denotan cuando cada lanzamiento cae dentro del círculo, es decir,

$$X_i = \begin{cases} 1, & \text{si el } i\text{-ésimo objeto cae dentro del círculo} \\ 0, & \text{si el } i\text{-ésimo objeto cae fuera del círculo.} \end{cases}$$

Con esta notación, se tiene claramente que $c_n = \sum_{i=1}^n X_i$. Notemos que cada X_i es una variable Bernoulli con parámetro p correspondiente a la probabilidad de caer dentro del círculo. Como el lugar en que cae el objeto es uniforme sobre la mesa, se tiene que

$$p = \frac{\text{área del círculo}}{\text{área de la mesa}} = \frac{\pi \cdot (a/2)^2}{a^2} = \frac{\pi}{4},$$

donde a es el largo del lado de la mesa (luego, $a/2$ es el radio del círculo). Por la ley de los grandes números fuerte y débil sabemos que el promedio de los X_i converge, cuando n crece indefinidamente, a la esperanza de X_1 casi seguramente y en probabilidad, i.e.:

$$\frac{c_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \rightarrow \mathbb{E}(X_1) = p = \frac{\pi}{4}.$$

Luego, cuando $n \rightarrow \infty$, se tiene que $4c_n/n \rightarrow \pi$ casi seguramente y en probabilidad.

- b) Denotemos X_i a la variable aleatoria que vale 1 si la i -ésima persona de la muestra prefiere el producto, 0 si no, para $i = 1, \dots, 100$. Cada X_i es una variable Bernoulli con parámetro p desconocido. Definamos las hipótesis

$$H_0 : p = 20\%$$

$$H_1 : p < 20\%.$$

Notemos que $\mathbb{E}(X_1) = p$ y $\text{var} X_1 = p(1-p)$. Por lo tanto, bajo H_0 , sabemos que el estadístico $Z = (\bar{X} - 0,2)/\sqrt{0,2 \cdot (1-0,2)/100}$ tiene distribución aproximadamente normal estándar. Dada la forma de las hipótesis, buscamos una región de rechazo del tipo $Z \in (-\infty, c]$, para una cierta constante c por determinar. Imponiendo que la probabilidad del error de tipo I sea $\alpha = 0,05$, tenemos que:

$$0,05 = \mathbb{P}(Z \in (-\infty, c] \mid H_0) \approx \mathbb{P}(\mathcal{N}(0,1) \leq c) = \mathbb{P}(\mathcal{N}(0,1) \geq -c).$$

Mirando la tabla de la normal, se obtiene que $-c = 1,64$ (también sirve 1,65). Entonces, la hipótesis del productor no se rechaza cuando $Z > c$, es decir:

$$c < Z = \frac{\bar{X} - 0,2}{\sqrt{0,2 \cdot (1-0,2)/100}} = 10 \cdot \frac{\bar{X} - 1/5}{\sqrt{1/5 \cdot 4/5}} = 25 \cdot (\bar{X} - 1/5) \Leftrightarrow \bar{X} > \frac{5 - 1,64}{25} = \frac{13,44}{100}.$$

Notemos que \bar{X} corresponde a $A/100$, donde A es la cantidad de personas que prefieren el producto. Por lo tanto, la mínima cantidad de estas personas tal que no se rechaza la hipótesis del productor es 14.

Si 10 personas prefieren el producto, de la última expresión se tiene que el valor observado de Z es $25 \cdot (10/100 - 1/5) = -25/10 = -2,5$. El p -valor es la probabilidad de obtener un resultado al menos tan extremo como el observado. Luego, mirando la tabla normal, tenemos:

$$p\text{-valor} = \mathbb{P}(Z \leq -2,5) \approx \mathbb{P}(\mathcal{N}(0,1) \leq -2,5) = \mathbb{P}(\mathcal{N}(0,1) \geq 2,5) = 0,62 \%.$$

- P2.** a) Sean X_1, \dots, X_n las mediciones a tomar, y $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ el promedio de ellas. Queremos asegurar que $|\bar{X} - d| < 0,5$ con probabilidad de al menos un 95 %, donde d coincide con la esperanza de las variables X_i . Usando la desigualdad de Chebyshev:

$$\mathbb{P}(|\bar{X} - d| < 0,5) = 1 - \mathbb{P}(|\bar{X} - d| \geq 0,5) \geq 1 - \frac{\text{var}(\bar{X})}{0,5^2} = 1 - \frac{4/n}{0,25} = 1 - \frac{16}{n},$$

donde hemos usado que $\text{var}(\bar{X}) = \sigma^2/n = \text{var}(X_1)/n$ debido a que las variables son i.i.d. Luego, para asegurar que la probabilidad de la izquierda sea al menos un 95 %, basta pedir que $1 - 16/n \geq 0,95$, es decir, $n \geq 16/0,05 = 320$.

Ahora usemos el TLC. Sabiendo que $(\bar{X} - d)/\sqrt{\sigma^2/n}$ tiene distribución aproximadamente normal estándar, tenemos:

$$\mathbb{P}(|\bar{X} - d| < 0,5) = \mathbb{P}\left(\frac{|\bar{X} - d|}{2/\sqrt{n}} < \frac{0,5}{2/\sqrt{n}}\right) \approx \mathbb{P}\left(|Z| < \frac{0,5}{2/\sqrt{n}}\right) = 1 - 2\mathbb{P}(Z \geq 0,25\sqrt{n}),$$

donde $Z \sim \mathcal{N}(0,1)$, y hemos usado la simetría de la densidad de Z . Luego, para que la probabilidad de la izquierda sea de al menos un 95 %, basta pedir que $1 - 2\mathbb{P}(Z \geq 0,25\sqrt{n}) \geq 95 \%$, es decir, $\mathbb{P}(Z \geq 0,25\sqrt{n}) \leq 2,5 \%$. Mirando una tabla de una normal estándar, debe tenerse entonces que $0,25\sqrt{n} \geq 1,96$, es decir, $\sqrt{n} \geq 4 \cdot 1,96 = 7,84$. Como $7,84^2 = 61,4656$, el menor natural que cumple la desigualdad anterior es $n = 62$.

- b) Sean X_1, \dots, X_n los datos de la muestra, con $n = 7$. Consideremos el estadístico

$$V = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2,$$

el cual sabemos que tiene distribución de chi-cuadrado con $n - 1$ grados de libertad, es decir, $V \sim \chi_6^2$. Para encontrar el intervalo, imponemos que $90 \% = \mathbb{P}(c < V < d)$, donde c y d son constantes tales que cumplen simetría de probabilidad de la distribución de V , es decir, $\mathbb{P}(V \leq c) = \mathbb{P}(V \geq d)$. Con todo esto:

$$90 \% = \mathbb{P}(c < V < d) = 1 - \mathbb{P}(V \leq c) - \mathbb{P}(V \geq d) \Rightarrow \mathbb{P}(V \leq c) = \mathbb{P}(V \geq d) = 5 \%.$$

Mirando la tabla de la distribución chi-cuadrado, obtenemos que $d = 12,59$. Además, $\mathbb{P}(V > c) = 1 - \mathbb{P}(V \leq c) = 95 \%$, y de la tabla obtenemos que $c = 1,635$. Por lo tanto, el intervalo de confianza para σ^2 al nivel 90 % se obtiene imponiendo que $c < V < d$ y despejando σ^2 . Es decir:

$$c < V < d \Leftrightarrow c < \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 < d \Leftrightarrow \frac{U}{d} < \sigma^2 < \frac{U}{c},$$

donde $U = \sum_{i=1}^n (X_i - \bar{X})^2$. Calculemos el valor de U obtenido en la muestra:

$$\bar{X} = \frac{1}{7}(5,0 + 2,4 - 1,0 + 2,6 + 4,0 - 2,0 + 3,0) = \frac{1}{7} \cdot 14,0 = 2,0,$$

luego:

$$U = \sum_{i=1}^n (X_i - \bar{X})^2 = 3,0^2 + 0,4^2 + 3,0^2 + 0,6^2 + 2,0^2 + 4,0^2 + 1,0^2$$

$$= 9,0 + 0,16 + 9,0 + 0,36 + 4,0 + 16,0 + 1,0 = 39,52,$$

y el intervalo de confianza para σ^2 queda

$$\left[\frac{39,52}{12,59}, \frac{39,52}{1,635} \right].$$

- P3.** a) El estimador de los momentos se obtiene al igualar el primer momento muestral, es decir \bar{X} , con el primer momento real de la variable, es decir θ/λ . Despejando λ , se obtiene entonces

$$\hat{\lambda}_{\text{mom}} = \frac{\theta}{\bar{X}}.$$

Para el otro estimador, calculemos la verosimilitud: si $f(\cdot; \lambda)$ es la densidad de la gama(θ, λ), tenemos que

$$L(x_1, \dots, x_n; \lambda) = \prod_{i=1}^n f(x_i; \lambda) = \prod_{i=1}^n \frac{\lambda e^{-\lambda x_i} (\lambda x_i)^{\theta-1}}{\Gamma(\theta)} \mathbb{1}_{[0, \infty)}(x_i) = \frac{\lambda^{n\theta} e^{-\lambda \sum_{i=1}^n x_i} [\prod_{i=1}^n x_i]^{\theta-1}}{\Gamma(\theta)^n},$$

donde hemos usado que todas las indicatrices valen 1 pues los x_i se obtuvieron desde una variable gama(θ, λ), la cual siempre toma valores ≥ 0 . Para maximizar lo anterior, derivamos con respecto a λ e igualamos a cero. Simplificando términos, se obtiene:

$$0 = n\theta \lambda^{n\theta-1} e^{-\lambda \sum_{i=1}^n x_i} - \lambda^{n\theta} e^{-\lambda \sum_{i=1}^n x_i} \sum_{i=1}^n x_i \Rightarrow \hat{\lambda}_{\text{MV}} = \frac{n\theta}{\sum_{i=1}^n x_i} = \frac{\theta}{\bar{X}}.$$

- b) Por la ley fuerte de los grandes números, sabemos que \bar{X} converge casi seguramente al promedio de la variable subyacente, cuando $n \rightarrow \infty$. Es decir, $\bar{X} \rightarrow \theta/\lambda$. Por lo tanto, cuando $n \rightarrow \infty$, se tiene la siguiente convergencia casi segura:

$$\hat{\lambda} = \frac{\theta}{\bar{X}} \rightarrow \frac{\theta}{\theta/\lambda} = \lambda.$$

- c) Calculemos la esperanza de $\hat{\lambda}$:

$$\mathbb{E}(\hat{\lambda}) = \mathbb{E}\left(\frac{\theta}{\frac{1}{n} \sum_{i=1}^n X_i}\right) = n\theta \mathbb{E}\left(\frac{1}{Y}\right),$$

donde $Y = \sum_{i=1}^n X_i$ tiene distribución gama($n\theta, \lambda$). Aplicando la fórmula para la esperanza de una función de una variable aleatoria, tenemos que:

$$\mathbb{E}\left(\frac{1}{Y}\right) = \int_0^\infty \frac{1}{y} \frac{\lambda e^{-\lambda y} (\lambda y)^{n\theta-1}}{\Gamma(n\theta)} dy = \frac{\lambda}{n\theta-1} \int_0^\infty \frac{\lambda e^{-\lambda y} (\lambda y)^{n\theta-2}}{\Gamma(n\theta-1)} dy,$$

donde hemos usado la propiedad fundamental de la función Γ . Notemos que la integral anterior vale 1, pues el integrando es exactamente la densidad de una gama($n\theta-1, \lambda$). Por lo tanto, la esperanza de $\hat{\lambda}$ es

$$\mathbb{E}(\hat{\lambda}) = n\theta \mathbb{E}\left(\frac{1}{Y}\right) = \lambda \frac{n\theta}{n\theta-1}.$$

Modificando $\hat{\lambda}$, obtenemos entonces que el siguiente estimador de λ es insesgado:

$$\tilde{\lambda} = \frac{n\theta - 1}{n\theta} \hat{\lambda} = \frac{n\theta - 1}{n\theta} \frac{\theta}{\bar{X}} = \frac{\theta - 1/n}{\bar{X}}.$$

Calculemos la varianza de $\hat{\lambda}$, para lo cual calculamos $\mathbb{E}(\hat{\lambda}) = n^2\theta^2\mathbb{E}(1/Y^2)$. Usando nuevamente la fórmula de la esperanza de una función de una variable, tenemos:

$$\mathbb{E}\left(\frac{1}{Y^2}\right) = \int_0^\infty \frac{1}{y^2} \frac{\lambda e^{-\lambda y} (\lambda y)^{n\theta-1}}{\Gamma(n\theta)} dy = \frac{\lambda^2}{(n\theta - 1)(n\theta - 2)} \int_0^\infty \frac{\lambda e^{-\lambda y} (\lambda y)^{n\theta-3}}{\Gamma(n\theta - 2)} dy,$$

donde nuevamente la integral vale 1 pues el integrando es la densidad de una $\text{gama}(n\theta - 2, \lambda)$. Por lo tanto:

$$\text{var}(\hat{\lambda}) = \mathbb{E}(\hat{\lambda}^2) - \mathbb{E}(\hat{\lambda})^2 = \lambda^2 \frac{n^2\theta^2}{(n\theta - 1)(n\theta - 2)} - \lambda^2 \frac{n^2\theta^2}{(n\theta - 1)^2} = \frac{n^2\theta^2\lambda^2}{(n\theta - 1)^2(n\theta - 2)}.$$

Luego, la varianza de $\tilde{\lambda}$ es

$$\text{var}(\tilde{\lambda}) = \text{var}\left(\frac{n\theta - 1}{n\theta} \hat{\lambda}\right) = \frac{(n\theta - 1)^2}{n^2\theta^2} \text{var}(\hat{\lambda}) = \frac{(n\theta - 1)^2}{n^2\theta^2} \frac{n^2\theta^2\lambda^2}{(n\theta - 1)^2(n\theta - 2)} = \frac{\lambda^2}{n\theta - 2},$$

lo cual claramente tiende a 0 cuando $n \rightarrow \infty$. Como además $\tilde{\lambda}$ es insesgado, por propiedad vista en cátedra se concluye que $\tilde{\lambda}$ es consistente.