



PAUTA CONTROL 3

P1. 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) Trabajamos con el estadístico

$$W = \frac{(n-1)s^2}{\sigma^2},$$

el cual sabemos que se distribuye como una $\chi_{n-1}^2 = \chi_5^2$. Imponemos entonces:

$$90 \% = \mathbb{P}(c \leq W \leq d) = 1 - \mathbb{P}(W < c) - \mathbb{P}(W > d).$$

Adicionalmente, imponemos simetría de las probabilidades, es decir, $5 \% = \mathbb{P}(W < c) = 1 - \mathbb{P}(W \geq c)$ y $5 \% = \mathbb{P}(W > d)$. Mirando una tabla, obtenemos $c = 1,145$ y $d = 11,07$. El intervalo queda:

$$\begin{aligned} c &\leq W \leq d \\ c &\leq \frac{(n-1)s^2}{\sigma^2} \leq d \\ \frac{(n-1)s^2}{d} &\leq \sigma^2 \leq \frac{(n-1)s^2}{c} \\ \frac{\sum (X_i - \bar{X})^2}{d} &\leq \sigma^2 \leq \frac{\sum (X_i - \bar{X})^2}{c} \\ \frac{0,14275}{11,07} &\leq \sigma^2 \leq \frac{0,14275}{1,145}. \end{aligned}$$

- P2.** a) Calculemos primero el estimador de máxima verosimilitud. La función de verosimilitud corresponde a:

$$L(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i) = \prod_{i=1}^n r x_i^{r-1} \mathbb{1}_{[0,1]}(x_i) = r^n \left[\prod_{i=1}^n x_i \right]^{r-1},$$

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

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

Despejando r , se obtiene entonces $\hat{r}_{MV} = -n / \sum_{i=1}^n \ln(X_i)$. Veamos ahora el estimador del método de los momentos, para lo cual calculamos la esperanza de la variable en consideración:

$$\mathbb{E}(X_i) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x r x^{r-1} dx = r \frac{x^{r+1}}{r+1} \Big|_0^1 = \frac{r}{r+1} = \bar{X},$$

donde la última igualdad la hemos impuesto, para aplicar el método de los momentos. Despejando r , obtenemos

$$\hat{r}_{\text{mom}} = \frac{\bar{X}}{1 - \bar{X}}.$$

- b) 1) Sean X_1, \dots, X_n los datos, con $n = 9$, y sean μ y σ^2 los parámetros de la distribución normal de la cual proviene la variable en consideración. Planteamos las hipótesis.

$$H_0 : \mu = 50$$

$$H_1 : \mu < 50.$$

Trabajamos con el estadístico

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

el cual sabemos que tiene distribución t -student con $9 - 1$ grados de libertad. Calculemos el valor \bar{X} y s^2 para los datos proveídos:

$$\begin{aligned} \sum_{i=1}^9 X_i &= 9 \cdot 50 + \sum_{i=1}^9 (X_i - 50) = 9 \cdot 50 + [-1 + 1 - 4 - 1 + 1 - 2 + 1 - 4 + 0] \\ &= 9 \cdot 50 - 9, \end{aligned}$$

luego $\bar{X} = \frac{1}{n} \sum X_i = 49$. Además:

$$\begin{aligned} \sum_{i=1}^9 (X_i - \bar{X})^2 &= 0^2 + 2^2 + (-3)^2 + 0^2 + 2^2 + (-1)^2 + 2^2 + (-3)^2 + 1^2 \\ &= 0 + 4 + 9 + 0 + 4 + 1 + 4 + 9 + 1 \\ &= 32, \end{aligned}$$

por lo tanto $s^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2 = 32/8 = 4$. El p -valor corresponde a la probabilidad, bajo H_0 , de obtener algo tan o más extremo que el valor del estadístico observado en la muestra. Dada la forma de las hipótesis, esto significa que el estadístico T sea menor o igual que lo observado, es decir,

$$p\text{-valor} = \mathbb{P}(T \leq T_{\text{obs}} | \mu = 50) = \mathbb{P}(T \leq \frac{49 - 50}{\sqrt{4/9}}) = \mathbb{P}(T \leq -1,5) = \mathbb{P}(T \geq 1,5),$$

donde hemos utilizado la simetría de la distribución de t -student. Sabiendo que $T \sim t_8$, mirando la tabla obtenemos que la probabilidad anterior es aproximadamente 0,1 (utilizando el valor más cercano a 1,5, es decir, 1,397). Luego, el p -valor es de 10 %. Para $\alpha = 5 \% < 10 \%$, esto significa que no corresponde rechazar H_0 .

2) Planteamos las hipótesis

$$H_0 : \sigma = 4$$

$$H_1 : \sigma < 4.$$

Para resolver el test, utilizamos el estadístico

$$W = \frac{(n-1)s^2}{\sigma^2},$$

el cual sabemos que tiene distribución de chi-cuadrado con $9 - 1$ grados de libertad. El p -valor es la probabilidad, bajo H_0 , de que este estadístico sea al menos tan extremo que lo observado, lo cual significa obtener algo menor o igual que lo observado:

$$p\text{-valor} = \mathbb{P}(W \leq W_{\text{obs}} | \sigma = 4) = \mathbb{P}(W \leq \frac{8 \cdot 4}{4^2}) = \mathbb{P}(W \leq 2) = 1 - \mathbb{P}(W > 2).$$

Sabiendo que $W \sim \chi_8^2$, mirando la tabla obtenemos que $\mathbb{P}(W > 2)$ es aproximadamente 0,975 (utilizando el valor más cercano a 2, es decir, 2,18). Luego, el p -valor es de 2,5 %. Para $\alpha = 5 \% > 2,5 \%$, debemos rechazar H_0 y aceptar la hipótesis alternativa de que σ es estrictamente menor que 4.

P3. a) En cátedra se vio que \bar{X} es insesgado, por lo tanto $\mathbb{E}(\bar{X}) = \mathbb{E}(X_1) = 1/\lambda$, donde hemos utilizado que la esperanza de una variable exponencial de parámetro λ es $1/\lambda$. Rehagamos el cálculo:

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda} = \frac{1}{\lambda}.$$

Para lo otro, en la demostración de la LDGN se vio que $\text{var}(\bar{X}) = \frac{\text{var}(X_1)}{n}$, y sabiendo que la varianza de una variable exponencial de parámetro λ es $1/\lambda^2$, se obtiene lo deseado. Explícitamente:

$$\text{var}(\bar{X}) = \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \left[\sum_{i=1}^n \text{var}(X_i) + \sum_{i \neq j} \text{cov}(X_i, X_j) \right] = \frac{1}{n^2} \sum_{i=1}^n \frac{1}{\lambda^2} = \frac{1}{n\lambda^2},$$

donde hemos utilizado que las constantes salen al cuadrado de la varianza, y que las covarianzas son todas 0 porque las variables son independientes.

b) Por LFGN, sabemos que \bar{X} converge c.s. cuando $n \rightarrow \infty$ a $\mathbb{E}(X_1) = 1/\lambda$, lo que equivale a que $1/\bar{X}$ converge c.s. a λ .

c) Veamos lo primero:

$$\begin{aligned}
& |\hat{\lambda} - \lambda| \leq \alpha\lambda \\
\Leftrightarrow & -\alpha\lambda \leq \frac{1}{\bar{X}} - \lambda \leq \alpha\lambda \\
\Leftrightarrow & \lambda(1 - \alpha) \leq \frac{1}{\bar{X}} \leq \lambda(1 + \alpha) \\
\Leftrightarrow & \frac{1}{\lambda(1 + \alpha)} \leq \bar{X} \leq \frac{1}{\lambda(1 - \alpha)} \\
\Leftrightarrow & -\frac{\alpha}{\lambda(1 + \alpha)} \leq \bar{X} - \frac{1}{\lambda} \leq \frac{\alpha}{\lambda(1 - \alpha)}.
\end{aligned}$$

También notemos que:

$$\begin{aligned}
& |\bar{X} - \frac{1}{\lambda}| \leq \frac{\alpha}{\lambda(1 + \alpha)} \\
\Leftrightarrow & -\frac{\alpha}{\lambda(1 + \alpha)} \leq \bar{X} - \frac{1}{\lambda} \leq \frac{\alpha}{\lambda(1 + \alpha)} \\
\Rightarrow & -\frac{\alpha}{\lambda(1 + \alpha)} \leq \bar{X} - \frac{1}{\lambda} \leq \frac{\alpha}{\lambda(1 - \alpha)},
\end{aligned}$$

y usando la equivalencia anterior, se concluye que el evento $\{|\bar{X} - \frac{1}{\lambda}| \leq \frac{\alpha}{\lambda(1 + \alpha)}\}$ está contenido en el evento $\{|\hat{\lambda} - \lambda| \leq \alpha\lambda\}$. Pasando al complemento, se concluye que $\{|\hat{\lambda} - \lambda| > \alpha\lambda\} \subseteq \{|\bar{X} - \frac{1}{\lambda}| > \frac{\alpha}{\lambda(1 + \alpha)}\}$, y aplicando $\mathbb{P}(\cdot)$, obtenemos

$$\mathbb{P}(|\hat{\lambda} - \lambda| > \alpha\lambda) \leq \mathbb{P}\left(|\bar{X} - 1/\lambda| > \frac{\alpha}{\lambda(1 + \alpha)}\right).$$

d) Utilizando lo anterior y aplicando la desigualdad de Chebyshev a la variable \bar{X} , tenemos:

$$\mathbb{P}(|\hat{\lambda} - \lambda| > \alpha\lambda) \leq \mathbb{P}\left(|\bar{X} - 1/\lambda| > \frac{\alpha}{\lambda(1 + \alpha)}\right) \leq \frac{\text{var}(\bar{X})}{\left(\frac{\alpha}{\lambda(1 + \alpha)}\right)^2} = \frac{\frac{1}{n\lambda^2}}{\frac{\alpha^2}{\lambda^2(1 + \alpha)^2}} = \frac{(1 + \alpha)^2}{n\alpha^2}.$$

e) Para $\alpha = 25\% = 1/4$, imponemos que la probabilidad de que el error relativo sea mayor que α no sea mayor que $\beta = 10\% = 1/10$, para lo cual utilizamos la cota anterior:

$$\mathbb{P}(|\hat{\lambda} - \lambda| > \alpha\lambda) \leq \frac{(1 + \alpha)^2}{n\alpha^2} \leq \beta \quad \Leftrightarrow \quad n \geq \frac{(1 + \alpha)^2}{\beta\alpha^2} = \frac{(5/4)^2}{(1/10) \cdot (1/4)^2} = 10 \cdot 5^2 = 250.$$

Luego, tomando 250 mediciones nos aseguramos que $\mathbb{P}(|\hat{\lambda} - \lambda| > \alpha\lambda) \leq 10\%$, o equivalentemente, pasando al complemento, $\mathbb{P}(|\hat{\lambda} - \lambda| \leq \alpha\lambda) \geq 90\%$, como queríamos.