Ingeniería Matemática
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE
MA3403-2 Probabilidades y Estadística
Primavera 2014

Roberto Cortez
Ignacio López
Dafne Pincheira

Pauta control 3

P1. *a)* 1) El estimador de los momentos se obtiene al igualar el primer momento muestral, es decir $\overline{X}$, con el primer momento real de la variable, es decir θ/λ. Despejando λ, se obtiene entonces

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

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

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

donde hemos usado que todas las indicatrices valen 1 pues los x_i se obtuvieron desde una variable Gamma(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:

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

2) 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 Gamma$(n\theta,\lambda)$. 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 Gamma$(n\theta-1,\lambda)$. 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:

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

b) Modelamos las mediciones como variables aleatorias normales independientes con parámetros μ y σ^2 desconocidos. Planteamos las hipótesis

$$H_0 : \mu = 130,$$
$$H_1 : \mu < 130.$$

Para $n = 40$, trabajamos con el estadístico

$$T = \frac{\overline{X} - 130}{s/\sqrt{n}},$$

el cual, bajo H_0, tiene distribución t-student con $n - 1 = 39$ grados de libertad. Dada la forma de H_1, rechazaremos H_0 cuando $T < c$, donde la constante c se determina imponiendo que la probabilidad del error de tipo I sea el $\alpha = 5\,\%$ deseado:

$$5\,\% = \mathbb{P}(T < c \mid H_0) = \mathbb{P}(t_{39} < c) = \mathbb{P}(t_{39} > -c),$$

donde en el último paso hemos utilizado la simietría de la distribución t-student. Desde una tabla (usando 40 grados de libertad), obtenemos que $c = -1{,}684$. Notar que el valor observado del estadístico T es

$$T_{\text{obs}} = \frac{128{,}6 - 130}{2{,}1/\sqrt{40}} = -\sqrt{40}\frac{1{,}4}{2{,}1} \approx -6{,}3\frac{2}{3} = -4{,}2.$$

Como $T_{\text{obs}} < c = -1{,}684$, corresponde rechazar H_0 y fallar en favor de la hipótesis de que $\mu < 130$.

P2. *a*) Sabiendo que la densidad de X_1 es f, calculemos $\mathbb{E}(X_1)$:

$$\mathbb{E}(X_1) = \int_0^1 x \cdot 2x dx = \left.\frac{2x^3}{3}\right|_{x=0}^{x=1} = 2/3.$$

Para la varianza, calculamos primero $\mathbb{E}(X_1^2)$:

$$\mathbb{E}(X_1^2) = \int_0^1 x^2 \cdot 2x dx = \left.\frac{2x^4}{4}\right|_{x=0}^{x=1} = 1/2.$$

Luego:

$$\text{var}(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 = 1/2 - 4/9 = 1/18.$$

b) El instante en que falla la última componente es $S_{15} = \sum_{i=1}^{15} X_i$. Como todos los X_i tienen densidad f, es claro que que

$$\mathbb{E}(S_{15}) = \sum_{i=1}^{15} \mathbb{E}(X_i) = 15\mathbb{E}(X_1) = 10.$$

Además, por independencia de los X_i, tenemos

$$\text{var}(S_{15}) = \sum_{i=1}^{15} \text{var}(X_i) = 15\text{var}(X_1) = 15/18.$$

Utilizando la desigualdad de Chebyshev, acotamos la probabilidad deseada:

$$\mathbb{P}(S_{15} \in (8, 12)) = 1 - \mathbb{P}(|S_{15} - 10| \geq 2) \geq 1 - \frac{15/18}{2^2} = 57/72.$$

c) Llamemos $\mu = \mathbb{E}(X_1) = 2/3$ y $\sigma^2 = \mathrm{var}(X_1) = 1/18$. Impongamos que la duración promedio sea mayor que 14horas = 14/24días con la probabilidad deseada de 90 %, o equivalentemente, que el complemento tenga probabilidad 10 %:

$$10\,\% = \mathbb{P}(\overline{X} < 14/24) = \mathbb{P}\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{14/24-\mu}{\sigma/\sqrt{n}}\right) \approx \mathbb{P}\left(\mathcal{N}(0,1) < \frac{14/24-\mu}{\sigma/\sqrt{n}}\right)$$
$$= \mathbb{P}\left(\mathcal{N}(0,1) > -\frac{14/24-\mu}{\sigma/\sqrt{n}}\right),$$

donde en el penúltimo paso hemos utilizado el TCL para aproximar por una normal estándar, y en el último paso usamos la simetría de la normal. De una tabla normal obtenemos que el lado derecho en la última desigualdad debe ser aproximadamente igual a $c = 1{,}28$. Es decir:

$$c \approx -\frac{14/24-\mu}{\sigma/\sqrt{n}} \;\Leftrightarrow\; \sqrt{n} \approx \frac{c\sigma}{\mu - 14/24} = \frac{1{,}28 \times \sqrt{1/18}}{2/24}$$
$$\Leftrightarrow\; n \approx (1{,}28)^2 \times 144/18 \approx 13{,}1.$$

Luego, $n = 13$ componentes bastan para cumplir lo deseado.

d) Notar que $\overline{X} = S_n/n$. Por la ley fuerte de los grandes números, sabemos que $\overline{X}$ converge casi seguramente a $\mathbb{E}(X_1) = 2/3$ cuando $n \to \infty$. Luego:

$$r = \lim_{n\to\infty} \frac{n}{S_n} = \lim_{n\to\infty} \frac{1}{\overline{X}} = \frac{1}{2/3} = 3/2.$$

P3. *a*) Trabajamos con el estadístico

$$T = \frac{\overline{X}-\mu}{s/\sqrt{n}},$$

el cual sabemos que posee distribución t-student con $n-1 = 15$ grados de libertad. Imponemos el nivel deseado a este estadístico, usando un intervalo simétrico respecto a 0:

$$95\,\% = \mathbb{P}(T \in [-c, c]) = 1 - \mathbb{P}(|T| > c) = 1 - 2\mathbb{P}(T > c),$$

donde hemos usado la simetría de la distribución t-student. Despejando la probabilidad anterior, obtenemos $\mathbb{P}(t_{15} > c) = 2{,}5\,\%$, y de una tabla se obtiene $c = 2{,}131$. Con esto, despejando μ obtenemos:

$$95\,\% = \mathbb{P}\left(-c \le \frac{\overline{X}-\mu}{s/\sqrt{n}} \le c\right) = \mathbb{P}\left(\overline{X} - \frac{s}{\sqrt{n}}c \le \mu \le \overline{X} + \frac{s}{\sqrt{n}}c\right).$$

Aproximando $sc = \sqrt{8{,}1} \times 2{,}1313 \approx 3 \times 2$, obtenemos finalmente que el intervalo de confianza buscado para μ es

$$\left[20{,}0 - \frac{3\times 2}{4}, 20{,}0 + \frac{3\times 2}{4}\right] = [18{,}5,\; 21{,}5].$$

b) Usamos el estadístico

$$U = (n-1)\frac{s^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2},$$

el cual sabemos que posee distribución chi-cuadrado con $n-1 = 15$ grados de libertad. Imponemos el nivel deseado a este estadístico:

$$95\,\% = \mathbb{P}(c \le U \le d) = 1 - \mathbb{P}(U < c) - \mathbb{P}(U > d).$$

Imponemos simetría de probabilidad para las constantes c y d, es decir, imponemos que $\mathbb{P}(U < c) = \mathbb{P}(U > d)$. Usando esto en la igualdad previa y despejando, se obtiene que $\mathbb{P}(U > d) = 2{,}5\,\%$, y por lo tanto $\mathbb{P}(U > c) = 1 - \mathbb{P}(U < c) = 97{,}5\,\%$. De una tabla de la distribución chi-cuadrado, se deduce que $c = 6{,}26$ y $d = 27{,}5$. Despejando σ^2, se obtiene:

$$95\,\% = \mathbb{P}\left(c \leq (n-1)\frac{s^2}{\sigma^2} \leq d\right) = \mathbb{P}\left(\frac{(n-1)s^2}{d} \leq \sigma^2 \leq \frac{(n-1)s^2}{c}\right).$$

Aproximando $(n-1)s^2/d = 15 \times 8{,}1/27{,}5 \approx 4{,}5$ y $(n-1)s^2/d = 15 \times 8{,}1/6{,}26 \approx 19{,}5$, el intervalo buscado es $[4{,}5,\ 19{,}5]$.

c) Llamemos $X_1, \ldots, X_n$ a la muestra original, e $Y_1, \ldots, Y_m$ a la muestra de la compañera. Anotemos $Z_1, \ldots, Z_{n+m}$ a la muestra consolidada, con $(Z_1, \ldots, Z_n) = (X_1, \ldots, X_n)$ y $(Z_{n+1}, \ldots, Z_{n+m}) = (Y_1, \ldots, Y_m)$. Podemos obtener $\overline{Z}$ a partir de $\overline{X}$ e $\overline{Y}$ como sigue:

$$\overline{Z} = \frac{1}{n+m}\sum_{i=1}^{n+m} Z_i = \frac{1}{n+m}\left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j\right) = \frac{1}{n+m}(n\overline{X} + m\overline{Y}),$$

y entonces el promedio obtenido en la muestra consolidada es $(16 \times 20 + 20 \times 20)/36 = 20{,}0$. Similarmente: si llamamos s_X^2, s_Y^2 y s_Z^2 a los estimadores insesgados de la varianza correspondientes a la muestra original, de la compañera, y consolidada, respectivamente, podemos obtener s_Z^2 a partir de s_X^2 y s_Y^2 usando el hecho que las tres muestras poseen el mismo promedio (igual a 20,0):

$$\begin{aligned} s_Z^2 = \frac{1}{n+m-1}\sum_{i=1}^{n+m}(Z_i - \overline{Z})^2 &= \frac{1}{n+m-1}\left(\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_i - \overline{Y})^2\right) \\ &= \frac{1}{n+m-1}\left((n-1)s_X^2 + (m-1)s_Y^2\right). \end{aligned}$$

Es decir, el valor de s_Z^2 es $(15 \times 8{,}1 + 19 \times 10)/35 = 8{,}9$. Para obtener el intervalo de confianza deseado, se repite el procedimiento de la primera parte, esta vez utilizando una t-student con $n + m - 1 = 35$ grados de libertad, por lo cual la constante c que se obtiene de la tabla es $c = 2{,}042$ (usando 30 grados de libertad; también es válido $c = 2{,}021$ si se usan 40 grados de libertad). Aproximando $s_Z c/\sqrt{n+m} = \sqrt{8{,}1} \times 2{,}042/\sqrt{36} \approx 1$, el intervalo obtenido es entonces

$$\left[\overline{Z} - \frac{s_Z}{\sqrt{n+m}}c, \overline{Z} + \frac{s_Z}{\sqrt{n+m}}c\right] = [20{,}0 - 1,\ 20{,}0 + 1] = [19{,}0,\ 21{,}0].$$

d) Se utiliza el estadístico

$$W = \frac{\overline{Z} - \mu}{\sigma/\sqrt{n+m}},$$

el cual posee distribución normal estándar. Se impone el nivel deseado a W usando un intervalo simétrico:

$$95\,\% = \mathbb{P}(W \in [-c, c]) = 1 - \mathbb{P}(|W| > c) = 1 - 2\mathbb{P}(W > c),$$

donde hemos usado la simetría de la normal. Esto implica que $\mathbb{P}(W > c) = 2{,}5\,\%$, y de la tabla normal se obtiene $c = 1{,}96$. Despejando μ:

$$95\,\% = \mathbb{P}\left(-c \leq \frac{\overline{Z} - \mu}{\sigma/\sqrt{n+m}} \leq c\right) = \mathbb{P}\left(\overline{Z} - \frac{\sigma}{\sqrt{n+m}}c \leq \mu \leq \overline{Z} + \frac{\sigma}{\sqrt{n+m}}c\right).$$

Aproximando $\sigma c/\sqrt{n+m} = \sqrt{9{,}0} \times 1{,}96/\sqrt{36} \approx 1$, el intervalo de confianza buscado es $[19{,}0,\ 21{,}0]$ (lo mismo de la parte anterior, debido a la aproximación utilizada; si el cálculo fuese exacto se obtendría una pequeña diferencia).