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 examen

P1. *a)* Hay 4 opciones: están los 5 transformadores operativos ($\binom{5}{5} = 1$ forma), o bien hay 4 operativos ($\binom{5}{4} = 5$ formas), o bien hay 3 operativos ($\binom{5}{3} = 10$ formas) o bien hay 2 operativos ($\binom{5}{2} = 10$ formas). Luego, en total son $1 + 5 + 10 + 10 = 26$ configuraciones tales que el Metro sigue funcionando.

b) Sean $X_1, \ldots, X_n$ variables $\exp(\lambda)$ independientes, y sea $Z = \min(X_1, \ldots, X_n)$. Dado $z > 0$, es claro que $Z > z$ si y sólo si $X_i > z$ para todo $i = 1, \ldots, n$. Luego, por independencia, tenemos:

$$\begin{aligned}\mathbb{P}(Z > z) = \mathbb{P}(X_1 > z, \ldots, X_n > z) &= \mathbb{P}(X_1 > z) \times \cdots \times \mathbb{P}(X_n > z) \\ &= e^{-\lambda z} \times \cdots \times e^{-\lambda z} \\ &= e^{-n\lambda z}.\end{aligned}$$

Es decir, $F_Z(z) = 1 - \mathbb{P}(Z > z) = 1 - e^{-n\lambda z}$, lo cual significa que $Z \sim \exp(n\lambda)$.

c) Por la propiedad de pérdida de memoria de la distribución exponencial, una vez que ocurre una falla, es como si cada transformador que queda operativo se reiniciara. Luego, si llamamos $X_1^{(i)}, \ldots, X_{5-(i-1)}^{(i)}$ a los tiempos de falla de los $5 - (i - 1)$ transformadores que quedan operativos luego de la $(i - 1)$-ésima falla (considerados *desde* el instante de dicha falla), se tiene que éstas son variables exponenciales independientes de parámetro 1. Como $S_i = \min(X_1^{(i)}, \ldots, X_{5-(i-1)}^{(i)})$, por la parte anterior es directo que $S_i \sim \exp((5 - (i - 1)) \times 1) = \exp(6 - i)$.

d) Es claro que el instante en que el Metro deja de funcionar corresponde a T_4, pues es el instante en que falla el cuarto transformador (dejando solamente 1 operativo). Usando la linealidad de la esperanza, y el hecho que la esperanza de una variable $\exp(\lambda)$ es $1/\lambda$, obtenemos:

$$\begin{aligned}\mathbb{E}(T_4) = \sum_{i=1}^{4} \mathbb{E}(S_i) = \sum_{i=1}^{4} \frac{1}{6-i} &= \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} \\ &= \frac{12 + 15 + 20 + 30}{60} \\ &= \frac{77}{60}.\end{aligned}$$

Para la varianza: sabemos que $\mathrm{cov}(S_i, S_j) = 0$ para $i \neq j$, por independencia. Luego, la varianza de la suma de los S_i es igual a la suma de las varianzas. Recordando que la esperanza de una $\exp(\lambda)$ es $1/\lambda^2$, tenemos:

$$\begin{aligned}\mathrm{var}(T_4) = \sum_{i=1}^{4} \mathrm{var}(S_i) = \sum_{i=1}^{4} \frac{1}{(6-i)^2} &= \frac{1}{25} + \frac{1}{16} + \frac{1}{9} + \frac{1}{4} \\ &= \frac{144 + 225 + 400 + 900}{3600} \\ &= \frac{1669}{3600}.\end{aligned}$$

P2. *a)* Es directo ver que $e^{-(x/r)^{-\beta}}$ es la primitiva de $\frac{\beta}{r}\left(\frac{x}{r}\right)^{-(\beta+1)} e^{-(x/r)^{-\beta}}$. Sabiendo que $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$, se concluye que $F_X(x) = e^{-(x/r)^{-\beta}} \mathbb{1}_{[0,\infty)}(x)$. Vemos lo otro: sea $Y = X^{-\beta}$, y usando el cálculo previo, tenemos para $y > 0$:

$$\begin{aligned}\mathbb{P}(Y > y) = \mathbb{P}(X^{-\beta} > y) = \mathbb{P}(X < y^{-1/\beta}) &= F_X(y^{-1/\beta}) \\ &= e^{-(y^{-1/\beta}/r)^{-\beta}} \\ &= e^{-r^\beta y}.\end{aligned}$$

Luego, $F_Y(y) = 1 - \mathbb{P}(Y > y) = 1 - e^{-r^\beta y}$, lo cual significa que $Y \sim \exp(r^\beta)$.

b) La función de verosimilitud es:

$$\begin{aligned}L = L(x_1, \dots, x_n; r) = \prod_{i=1}^n f_X(x_i; r) &= \prod_{i=1}^n \frac{\beta}{r}\left(\frac{x_i}{r}\right)^{-(\beta+1)} e^{-(x_i/r)^{-\beta}} \\ &= \beta^n r^{n\beta} \prod_{i=1}^n x_i^{-(\beta+1)} e^{-(x_i/r)^{-\beta}}.\end{aligned}$$

Luego:

$$\log L = n \log \beta + n\beta \log r - \sum_{i=1}^n [(\beta + 1) \log x_i + r^\beta x_i^{-\beta}].$$

Para maximizar lo anterior con respecto a r derivamos e igualamos a 0:

$$0 = \frac{\partial \log L}{\partial r} = \frac{n\beta}{r} - \beta r^{\beta-1} \sum_{i=1}^n x_i^{-\beta},$$

es decir, $0 = n - r^\beta \sum_{i=1}^n x_i^{-\beta}$. Despejando r y reemplazando los x_i por la muestra, obtenemos que el estimador de máxima verosimilitud es $\hat{r} = (\frac{1}{n} \sum_{i=1}^n X_i^{-\beta})^{-1/\beta}$.

c) Definimos $Y_i = X_i^{-\beta}$. Por la parte (a) sabemos que $Y_1, \dots, Y_n$ es una m.a.s. con distribución común $\exp(r^\beta)$. Por la LFGN, se deduce que $1/\hat{r}^\beta = \overline{Y}$ converge casi seguramente a $\mu := \mathbb{E}(Y_1) = \mathbb{E}(\exp(r^\beta)) = 1/r^\beta$ cuando $n \to \infty$. Tomando $(\cdot)^{-1/\beta}$, esto equivale a que $\hat{r} \to r$.

d) Escribamos Z en términos de la muestra $Y_1, \dots, Y_n$: sabemos que $\hat{r}^{-\beta} = \overline{Y}$, y como cada Y_i es una variable $\exp(r^\beta)$, también tenemos que $\mu = \mathbb{E}(Y_1) = r^{-\beta}$ y que $\sigma^2 = \mathrm{var}(Y_1) = r^{-2\beta}$. Luego,

$$Z = \frac{\hat{r}^{-\beta} - r^{-\beta}}{r^{-\beta}/\sqrt{n}} = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}},$$

es decir, Z es el estadístico que aparece en el TCL asociado a la muestra $Y_1, \dots, Y_n$. Para n grande, el TCL asegura que Z posee distribución aproximada $\mathcal{N}(0, 1)$.

e) Imponemos el nivel deseado al estadístico Z, el cual posee distribución aproximada $\mathcal{N}(0, 1)$, usando un intervalo simétrico respecto al origen:

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

donde hemos usado la simetría de la distribución $\mathcal{N}(0, 1)$. Despejando, se obtiene $\mathbb{P}(\mathcal{N}(0, 1) > c) \approx \alpha/2$, y el valor de c se despeja de una tabla o de un software estadístico (en cátedra utilizamos la notación $c \approx z_{\alpha/2}$, pero acá lo dejaremos como c simplemente). Con esto, obtenemos el intervalo buscado para r:

$$\begin{aligned}1 - \alpha \approx \mathbb{P}\left(-c \le \frac{\hat{r}^{-\beta} - r^{-\beta}}{r^{-\beta}/\sqrt{n}} \le c\right) &= \mathbb{P}\left(-\frac{c}{\sqrt{n}} \le \frac{r^\beta}{\hat{r}^\beta} - 1 \le \frac{c}{\sqrt{n}}\right) \\ &= \mathbb{P}\left(\hat{r}\left(1 - \frac{c}{\sqrt{n}}\right)^{1/\beta} \le r \le \hat{r}\left(1 + \frac{c}{\sqrt{n}}\right)^{1/\beta}\right).\end{aligned}$$

P3. *a*) La cantidad de datos es $n = 50+73+87 = 210$. Las probabilidades estimadas en base a la muestra son $\hat{p}_j = n_j/n$, donde $n_1 = 50$, $n_2 = 73$ y $n_3 = 87$ son la cantidad de clientes que prefieren el envase A, B y C, respectivamente. Las probabilidades de la hipótesis nula son equiprobables entre las $k = 3$ posibilidades, es decir, $p_j^0 = 1/k = 1/3$, para $j = 1, 2, 3$. Con esto, calculemos el valor observado en la muestra del estadístico Δ:

$$\Delta_{\text{obs}} = n\sum_{j=1}^{k} \frac{(\hat{p}_j - p_j^0)^2}{p_j^0} = n\sum_{j=1}^{k} \frac{\left(\frac{n_j}{n} - \frac{n}{nk}\right)^2}{\frac{1}{k}} = \frac{k}{n}\sum_{j=1}^{k}(n_j - n/k)^2$$
$$= \frac{1}{70}[(50-70)^2 + (73-70)^2 + (87-70)^2],$$

es decir, $\Delta_{\text{obs}} = [400+9+289]/70 = 698/70 \approx 10$. El p-valor corresponde a la probabilidad, bajo H_0, de obtener algo tan o más extremo que lo observado en la muestra. Sabemos que bajo H_0, el estadístico Δ tiene distribución aproximada de una χ^2_{k-1} (con $k-1 = 2$), con lo cual obtenemos:

$$p\text{-valor} = \mathbb{P}(\Delta \geq \Delta_{\text{obs}} \mid H_0) \approx \mathbb{P}(\chi^2_2 \geq 10) \approx 0{,}005,$$

donde en el último paso hemos utilizado la tabla de la distribución chi-cuadrado (también es válido obtener 0,01). Notemos que $\alpha = 5\,\% > 0{,}5\,\% \approx p$-valor, por lo tanto corresponde rechazar la hipótesis nula. Es decir, los clientes sí poseen preferencia por alguno de los envases.

b) Claramente

$$\mathbb{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n, \qquad \beta = \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} \in \mathbb{R}^{k+1}.$$

La matriz $\mathbb{X}$ está compuesta por las potencias de los x_i, incluyendo una columna de 1's correspondiente al coeficiente β_0. Específicamente:

$$\mathbb{X} = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{bmatrix} \in \mathbb{R}^{n\times(k+1)}.$$

c) Sabemos que $\hat{b}$ corresponde a la covarianza muestral sobre la varianza muestral de x, es decir

$$\hat{b} = \frac{\frac{1}{n}\sum x_i y_i - \overline{x}\,\overline{y}}{\frac{1}{n}\sum x_i^2 - \overline{x}^2} = \frac{2{,}4 - 3\times 2}{10{,}2 - 3^2} = \frac{-3{,}6}{1{,}2} = -3.$$

Además, sabemos que $\hat{a} = \overline{y} - \hat{b}\overline{x} = 2 + 3\times 3 = 11$. Luego, el valor aproximado del y asociado a $x = 2$ es $y = a + bx \approx \hat{a} + \hat{b}x = 11 - 3\times 2 = 5$.