Estadística

Victoria Torroja Rubio

19/1/2026

Índice general

Despacho: 303E
Correo: beatrizg@ucm.es
Evaluación:

- 80 % examen final
- 20 % evaluación continua (en este caso serán trabajos de R)

Bibliografía:

- 'Interferencia estadística' de M.A. Gómez Villegas
- 'Statistical interference' de Casella Berger

Capítulo 0

Estadística descriptiva

Práctica:

- Hacer en R el ejemplo 1.3.2 (lo del colesterol) del libro (página 11). Lo que hay que hacer aparece en el ejercicio 1.9.3.
- Programar el coeficiente de simetría y de curtosis en R.

Ejemplo. Página 44, Ejercicio 1. Consideramos la población dada por la distribución de Bernuilli $B\left(\frac{1}{2}\right)$. Se consideran todas las m.a.s de tamaño 3, es decir,

$$(X_1, X_2, X_3).$$

Recordamos que

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2, \ \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Tendremos que

$$\overline{X} = \begin{cases} 0 \to \frac{1}{8} \\ \frac{1}{3} \to \frac{3}{8} \\ \frac{2}{3} \to \frac{3}{8} \\ 1 \to \frac{1}{8} \end{cases}.$$

De esta forma tenemos que

$$P_{\overline{X}}(0) = \frac{1}{8}, \quad P_{\overline{X}}\left(\frac{1}{3}\right) = \frac{3}{8}, \text{ etc.}$$

Lo que estamos haciendo es considerar los diferentes casos para X_1, X_2 y X_3 y calcular la media en cada caso. A partir de ahí, estamos tratando a la media como una variable aleatoria discreta.

Ejemplo. Página 44, Ejercicio 2. Una forma de media θ y varianza 1. Se toma una m.a.s de tamaño n. Así, tenemos que $X \sim F_\theta(x) = P_\theta(X \leq x)$. Tendremos que $E_\theta[X] = \theta$ y $V_\theta(X) = 1$. Tenemos que calcular θ para que

$$P_\theta\left(\left|\overline{X} - \theta\right| \leq 0{,}5\right) \geq 0{,}95.$$

Tenemos que aplicar la desigualdad de Chebychev, primer calculamos

$$E_\theta\left[\overline{X}\right]=\frac{1}{n}\sum_{i=1}^{n}E_\theta\left[X_i\right]=\theta.$$

Por otro lado, $D_\theta\left(X\right)=\sqrt{V_\theta\left(X\right)}=1$, así

$$D_\theta\left(\overline{X}\right)=\sqrt{V_\theta\left(\overline{X}\right)}=\sqrt{V_\theta\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)}=?\frac{1}{n}\sum_{i=1}^{n}V_\theta\left(X_i\right)=\frac{1}{n}\to 0.$$

Aplicando la desigualdad de Chebychev

$$P_\theta\left(\left|\overline{X}-\theta\right|\geq\frac{K}{n}\right)\leq\frac{1}{K}.$$

Ejemplo. **Página 45, Ejercicio 6.** Sea una m.a.s de tamaño n calculamos la distribución de la media muestral $\overline{X}$, cuando la población:

(b) $X\sim\gamma\left(p,a\right)$. Denotamos $\theta=\left(p,a\right)\in\mathbb{R}^+\times\mathbb{R}^+$ y decimos que $\theta\in\Theta\subset\mathbb{R}^2$. Recordamos que

$$f\left(x|_\theta\right)=f_\theta\left(x\right)=\frac{a^p x^{p-1}e^{-ax}}{\Gamma\left(p\right)},\ x>0.$$

Hacemos el ejercicio con la ecuación característica:

$$\varphi_{\overline{X}}\left(t\right)=E_\theta\left[e^{it\overline{X}}\right]=E_\theta\left[e^{i\frac{t}{n}\sum_{i=1}^{n}X_i}\right]=\prod_{i=1}^{n}\varphi_{X_i}\left(\frac{t}{n}\right)=\varphi_X\left(\frac{t}{n}\right)^n=\left(1-i\frac{t}{na}\right)^{-pn}.$$

La función característica se corresponde con $\gamma\left(np,na\right)$.

(c) Basta con sustituir $\lambda=1$ en el apartado anterior.

Distribuciones asociadas a la normal

Consideremos que $\left(X_1,\ldots,X_n\right)$ es una m.a.s con cada variable independientes y siguen una distribución $X\sim N\left(0,1\right)$. Consideremos la variable aleatoria

$$Y:=\sum_{i=1}^{n}X_i^2=X_1^2+\cdots+X_n^2.$$

Tenemos que

$$\varphi_Y\left(t\right)=\prod_{i=1}^{n}\varphi_{X_i^2}\left(t\right).$$

En general, si $Z=X^2$,

$$F_Z\left(z\right)=P\left(Z\leq z\right)=P\left(X^2\leq z\right)=P\left(-\sqrt{z}\leq X\leq\sqrt{z}\right)=F_X\left(\sqrt{z}\right)-F_X\left(-\sqrt{z}\right).$$

La función de densidad será

$$f_Z(t) = \frac{1}{2\sqrt{z}} f_X\left(\sqrt{z}\right) + \frac{1}{2\sqrt{z}} f_X\left(-\sqrt{t}\right) = \frac{1}{\sqrt{2\pi}} t^{-\frac{1}{2}} e^{-\frac{1}{2}t},\ t > 0.$$

Podemos ver que se trata de la función de densidad de una distribución Gamma, así, $Z \sim \gamma\left(\frac{1}{2}, \frac{1}{2}\right)$. Definimos $\chi_1^2 \sim \gamma\left(\frac{1}{2}, \frac{1}{2}\right)$, el uno viene de que tenemos un grado de libertad (se genera a partir de una normal $N(0,1)$). Por la propiedad de reproductividad de la distribución Gamma, tenemos que $Y \sim \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$. A esta distribución la llamamos $\chi_n^2 \sim \gamma\left(\frac{n}{2}, \frac{1}{2}\right)$ (decimos que tiene n grados de libertad porque viene de n distribuciones normales $N(0,1)$). Recordamos que la función característica de una distribución γ es

$$\varphi(t) = \left(1 - \frac{it}{a}\right)^{-p}.$$

Ejemplo. **Ejercicio 19.** Sea $(X_1, \ldots, X_n)$ una m.a.s distribuida idénticamente a X tal que

$$P(X = x_i) = p_i,\ i \in \mathbb{N}.$$

Sea $\left(X_{(1)}, \ldots, X_{(n)}\right)$ la muestra ordenada de menor a mayor. Demostremos que la distribución para $X_{(k)}$ es discreta y

$$P\left(X_{(1)} \leq x_i\right) = \sum_{j=k}^{n} \binom{n}{j} (F(x_i))^j (1 - F(x_i))^{n-j}.$$

El resto viene dado en el enunciado. Tenemos que $X_{(1)} = \text{mín}\{X_1, \ldots, X_n\}$ y $X_{(n)} = \text{máx}\{X_1, \ldots, X_n\}$. Tenemos que

$$F_{X_{(n)}}(t) = P\left(X_{(n)} \leq t\right) = P\left(\bigcap_{1 \leq i \leq n} \{X_i \leq t\}\right) = \prod_{i=1}^{n} P(X_i \leq t) = F_X(t)^n.$$

Por otro lado,

$$F_{X_{(1)}}(t) = 1 - P\left(X_{(1)} > t\right) = 1 - P\left(\bigcap_{1 \leq i \leq n} \{X_i > t\}\right) = 1 - \prod_{i=1}^{n} P(X_i > t) = 1 - (1 - F_X(t))^n.$$

Ahora, buscamos $F_{X_{(k)}}(t)$ para $k \in \{1, \ldots, n\}$. Tenemos que

$$F_{X_{(k)}}(t) = P\left(X_{(k)} \leq t\right) = \sum_{j=k}^{n} P(\text{'exactamente } j \text{ por debajo de } t\text{'}) = \sum_{j=k}^{n} \binom{n}{j} F_X(t)^j (1 - F_X(t))^{n-j}.$$

Finalmente, recordamos que $P_{X_{(k)}}(x_i) = F_{X_{(k)}}(x_i) - F_{X_{(k)}}(x_{i-1})$.