Estadística I
Grado en Matemáticas, UAM, 2015-2016

Hoja 3. Muestreo aleatorio y estadísticos

Sobre el estadístico "media muestral"

1. Sea X una variable aleatoria continua y positiva tal que $\mathbf{E}(X) = \mu > 0$ y $\mathbf{V}(X) = 1$. Estima cuál debe ser el (mínimo) tamaño n de una muestra aleatoria para que, con probabilidad del 95 %, el error relativo entre la media muestral y μ sea inferior al 1 %.

2. Consideremos una muestra aleatoria de tamaño $n = 50$ de una variable X que sigue una distribución $\mathcal{N}(4, 1)$ (una normal con media 4 y varianza 1).

a) Utiliza la desigualdad de Chebyshev para obtener una cota superior para $\mathbf{P}(|\overline{X} - 4| > 0.3)$.

b) Calcula numéricamente la probabilidad $\mathbf{P}(|\overline{X} - 4| > 0.3)$ utilizando que $\overline{X}$ es una normal. Compara el resultado con el del apartado a). Medita al respecto.

3. Sea X una $\textsc{geo}(p)$, es decir,

$$\mathbf{P}(X = k) = p\,(1 - p)^{k-1}\,, \qquad \text{para cada } k = 1, 2, \ldots$$

a) Comprueba que la suma $n\overline{X}$ es una binomial negativa $\textsc{binneg}(n, p)$, cuya función de masa viene dada por

$$\mathbf{P}(X = k) = \binom{k-1}{n-1} p^n\,(1 - p)^{k-n}\,, \qquad \text{para cada } k = n, n+1, n+2, \ldots$$

b) ¿Qué función de masa tiene la variable $\overline{X}$?

Otros estadísticos

4. Definimos la asimetría y la curtosis de una variable Z con media μ y desviación típica σ como

$$\text{asim}(Z) = \frac{\mathbf{E}((Z-\mu)^3)}{\sigma^3} \qquad \text{y} \qquad \text{curt}(Z) = \frac{\mathbf{E}((Z-\mu)^4)}{\sigma^4}.$$

Sea $X_1, \ldots, X_n$ una muestra aleatoria de X. Comprueba que

$$\text{asim}(\overline{X}) = \frac{\text{asim}(X)}{\sqrt{n}} \qquad \text{y que} \qquad \text{curt}(\overline{X}) = \frac{\text{curt}(X)}{n} + 3\,\frac{n-1}{n}.$$

5. Llamemos H al valor del segundo término más grande de entre $X_1, X_2, \ldots, X_n$. Prueba que

$$F_H(t) = n(1 - F_X(t))F_X(t)^{n-1} + F_X(t)^n.$$

¿Y si H fuera el valor del segundo término más pequeño? ¿Y si fuera el tercero?

6. Sea $X \sim \textsc{exp}(\lambda)$, con $\lambda > 0$. Consideramos el estadístico $m_n = \text{mín}(X_1, \ldots, X_n)$. Comprueba que $m_n \sim \textsc{exp}(n\lambda)$ y calcula $n\mathbf{E}(m_n)$.

7. Sea $X \sim \textsc{unif}([0, a])$, con $a > 0$. Consideramos el estadístico $M_n = \text{máx}(X_1, \ldots, X_n)$.

a) Determina las funciones de distribución y densidad de M_n, y calcula $\mathbf{E}(M_n)$ y $\mathbf{V}(M_n)$.

b) Para un $\varepsilon > 0$ dado, calcula

$$\mathbf{P}(a - M_n \geq \varepsilon\, a).$$

8. Sean $X_1, \ldots, X_n$ variables normales estándar independientes. Definimos las variables

$$Z_j = \sum_{i=1}^{j} X_i \qquad \text{para cada } j = 1, \ldots, n.$$

El vector $\mathbb{Z} = (Z_1, \ldots, Z_n)$ sigue una normal multidimensional. ¿Por qué? Determina su vector $\mathbf{m}$ de medias y su matriz $\mathbf{V}$ de varianzas/covarianzas.

9. Sean $Y, Z_1, Z_2, \ldots, Z_n$ variables normales estándar independientes. Sea $\rho \in (0,1)$. Definimos el vector aleatorio $\mathbb{X} = (X_1, \ldots, X_n)$ mediante

$$X_j = \sqrt{\rho}\, Y + \sqrt{1-\rho}\, Z_j \qquad \text{para cada } j = 1, \ldots, n.$$

Dando por sentado que $\mathbb{X} \sim \mathcal{N}(\mathbf{m}, V)$, halla el vector de medias $\mathbf{m} = \mathbf{0}$ y la matriz V.

Sobre la χ^2, la t de Student y la F de Fisher

10. Si $Z \sim \chi^2_n$, comprueba que, para $n > 4$,

$$\mathbf{E}(1/Z^2) = \frac{1}{(n-2)(n-4)}\,.$$

11. Calcula el coeficiente de asimetría de $Z \sim \chi^2_n$ en términos de n.

(Sugerencia: si X es normal estándar, $\mathbf{E}(X^2) = 1$, $\mathbf{E}(X^4) = 3$ y $\mathbf{E}(X^6) = 15$).

12. a) Usando que $\Gamma(x+1) = x\Gamma(x)$ para $x > 0$ y que $\Gamma(1/2) = \sqrt{\pi}$, comprueba que

$$\Gamma(n + \tfrac{1}{2}) = \frac{(2n)!}{4^n\, n!}\,\sqrt{\pi} \quad \text{para cada entero } n \geq 0.$$

b) Para cada $n \geq 1$, la función de densidad de una variable t de Student con n grados de libertad viene dada por

$$f_{t_n}(t) = D_n \Big(\frac{1}{1 + t^2/n}\Big)^{(n+1)/2}, \quad \text{donde } D_n = \frac{1}{\sqrt{n\pi}}\,\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\,.$$

Comprueba, usando la fórmula de Stirling, que

$$\lim_{n\to\infty} f_{t_n}(x) = \phi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}$$

13. Comprueba que si $Z \sim F_{n,m}$ entonces

$$\mathbf{V}(Z) = \frac{2(m+n-2)m^2}{n(m-4)(m-2)^2}\,.$$

Muestreo normal

14. a) Sea $(X_1, \ldots, X_{100})$ una muestra aleatoria de X normal con parámetros $\mu = 3$ y $\sigma^2 = 4$. Calcula $\mathbf{P}\big(|\overline{X} - \mu| > 1/2\,,\, S^2 < 4.2\big)$.

b) Sea $(X_1, \ldots, X_n)$ una muestra aleatoria de una variable $X \sim \mathcal{N}(\mu, \sigma^2)$. Para $t > 0$ fijo, calcula la probabilidad de que $n\,(\overline{X} - \mu)^2 + (n-1)\,S^2$ sea $\geq t$.