Conceptos básicos de inferencia estadística | Semestre 2018-1

Tarea 5

Fecha de entrega: 5 de octubre

1. Sea X una variable aleatoria continua con varianza finita.

   a) Mostrar que $f(a) = E[(X-a)^2]$ es mínima para $a = E(X)$.

   b) Mostrar que $g(a) = E[|X-a|]$ es mínima para $a = med(X)$. Donde $med(X)$ es la mediana de X y se define como la constante c tal que $P(X < c) = P(X > c) = 0.5$ (para X continua).

2. Simular n observaciones de una distribución $Ga(\alpha, \lambda)$, con $\alpha > 0$ y $\lambda > 0$, para $n = 10^1, 10^2, \ldots, 10^8$; con cada muestra calcular $m'_1 = \bar{x}_n$. ¿Se parece m'_1 a $\mu'_1 = E(X)$?

3. Encontrar la distribución de $X_{(1)}$ (mínimo) y $X_{(n)}$ (máximo) de una muestra aleatoria de tamaño n de una distribución:

   a) $U(0, \beta)$, $\beta > 0$.

   b) $Exp(\lambda)$, $\lambda > 0$.

   ¿Se parecen las distribuciones anteriores a alguna de las vistas en clase? (para la uniforme considerar $\beta = 1$). ¿Cómo simularía observaciones de las distribuciones de $X_{(1)}$ y $X_{(n)}$?

4. Simular $N = 5{,}000$ muestras de tamaño n de una distribución $f(\cdot \,|\, \boldsymbol{\theta})$. Con cada muestra calcular $m'_1 = \bar{x}_n$. Las observaciones $\bar{x}_n^{(1)}, \ldots, \bar{x}_n^{(N)}$ constituyen una realización de una muestra aleatoria de tamaño N de la distribución de $\bar{X}_n$. Utilizar las $\bar{x}_n^{(i)}$ para determinar si distribución de $\bar{X}_n$ es aproximadamente normal: con estadísticos descriptivos (simetría, curtosis) y con métodos gráficos (histograma, densidad estimada). Considerar $n = 10, 30, 50, 100$ y 1000 y $f(\cdot \,|\, \theta) = Exp(\cdot \,|\, \lambda)$ y $Ga(\cdot \,|\, \alpha, \lambda)$, con $\alpha > 0$ y $\lambda > 0$.

5. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $N(\theta, \theta^2)$, con $\theta > 0$. Estimar θ con el método de momentos:

   a) Con los primeros momentos no centrales).

   b) Con los segundos momentos no centrales.

   ¿Son iguales los estimadores encontrados?

6. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $Ga(\alpha, \lambda)$, con $\alpha, \lambda > 0$. Estimar α y λ con el método de momentos utilizando la media y la varianza.

7. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $U(\alpha, \beta)$, con $\alpha < \beta$. Estimar α y β con el método de momentos utilizando la media y la varianza.

8. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $Be(\alpha, \beta)$, con $\alpha > 0$ y $\beta > 0$. Estimar α y β con el método de momentos utilizando la media y la varianza.

9. Sobre la distribución F.

   a) Argumentar que si X tiene una distribución F, entonces $Y = X^{-1}$ también tiene una distribución F. Si los grados de libertad de la distribución de X son m (numerador) y n (denominador), ¿cuáles son los grados de libertad de la distribución de Y?

b) Mostrar que si $X \sim F(m,n)$, entonces

$$Y = \frac{mX/n}{1+mX/n} \sim Be\left(\frac{m}{2}, \frac{n}{2}\right).$$

Comprobar el resultado anterior con simulación.

c) Utilizar el resultado anterior (teórico) para calcular la media y la varianza de una distribución $F(m,n)$.

10. Sobre la distribución t

a) Argumentar que si X tiene una distribución t, entonces $Y = X^2$ tiene una distribución F. Si la distribución de X tiene n grados, ¿cuáles son los grados de libertad de la distribución de Y?

b) Mostrar que

$$\lim_{n\to} t(x \,|\, n) = N(x \,|\, 0,1), \qquad \text{para } x \in \mathbb{R} \text{ fijo },$$

donde $t(x \,|\, n)$ es la fdp de la distribución t con n grados de libertad y $N(x \,|\, 0,1)$ es la fdp de la distribución normal estándar.

c) Verificar gráficamente el resultado anterior. Graficar $t(x \,|\, n)$ para distintos valores de n y comparar con la gráfica de $N(x \,|\, 0,1)$.

d) Mostrar que si $X \sim t(n)$, entonces

$$Y = \frac{1}{1+X^2/n} \sim Be\left(\frac{n}{2}, \frac{1}{2}\right).$$

Comprobar el resultado anterior con simulación.

11. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución con media μ y varianza σ^2. Se define el estadístico T como

$$T = \sum_{j=1}^{n} \alpha_j X_j,$$

donde $\alpha_1, \ldots, X_n$ son constantes conocidas.

a) Mostrar que $E(T) = \mu$ si y sólo si $\sum_{j=1}^{n} \alpha_j = 1$.

b) Mostrar que si $\sum_{j=1}^{n} \alpha_j = 1$, entonces $V(T)$ se minimiza si $\alpha_j = 1/n$ para $j = 1, \ldots, n$.

⋆ Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $f(\cdot \,|\, \boldsymbol{\theta})$, con valor esperado μ y varianza σ^2. Mostrar que

$$V(S_n^2) = \frac{1}{n}\left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right),$$

donde

$$S_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2 \qquad \text{y} \qquad \mu_r = E[(X-\mu)^r], \quad \text{para } r = 1,2,3,\ldots.$$