Estadística (M) - Segundo cuatrimestre 2020

Práctica 5 - Optimalidad

Se recomienda empezar por los ejercicios marcados sin $\star$.

A) *Estadísticos suficientes*

1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $\mathcal{P}(\lambda)$. Demostrar que $T = \sum_{i=1}^n X_i$ es un estadístico suficiente para λ:

   (a) Aplicando la definición.

   (b) Aplicando el Teorema de Factorización.

2. Sean $X_1, \ldots, X_n$ variables aleatorias independientes. Probar que:

   (a) Si $X_i \sim Bi(n_i, p)$, $1 \leq i \leq n$, con n_i conocidos, entonces $\sum_{i=1}^n X_i$ es suficiente para p.

   (b) Si $X_i \sim \Gamma(\alpha, \lambda)$ entonces $(\prod_{i=1}^n X_i, \sum_{i=1}^n X_i)$ es suficiente para (α, λ).
   Si α es conocido, entonces $\sum_{i=1}^n X_i$ es suficiente para λ.
   Si λ es conocido, entonces $\prod_{i=1}^n X_i$ es suficiente para α.

   (c) Si $X_i \sim \beta(r, s)$ entonces $(\prod_{i=1}^n X_i, \prod_{i=1}^n (1 - X_i))$ es suficiente para (r, s).
   Si r es conocido, entonces $\prod_{i=1}^n (1 - X_i)$ es suficiente para s.
   Si s es conocido, entonces $\prod_{i=1}^n X_i$ es suficiente para r.

   (d) Si X_i tiene función de probabilidad puntual

   $$f(x; \theta, p) = (1-p) p^{x-\theta} \text{ para } x = \theta, \theta+1, \theta+2, \ldots$$

   con $(\theta, p) \in \mathbb{R} \times (0,1)$, entonces $\left(X_{(1)}, \sum_{i=1}^n X_i\right)$ es suficiente para (θ, p).
   Si p es conocido, entonces $X_{(1)}$ es suficiente para θ.
   Si θ es conocido, entonces $\sum_{i=1}^n X_i$ es suficiente para p.

3. Sean $X_1, \ldots, X_n$ v.a.i.i.d. $\mathcal{U}(\theta_1, \theta_2)$. Mostrar que $\left(X_{(1)}, X_{(n)}\right)$ es suficiente para (θ_1, θ_2).

$\star$ 4. Sea $X_1, \ldots, X_n$ una m.a. de una distribución con densidad

$$f(x; \mu, \sigma) = \frac{1}{\sigma} \exp\left\{-\frac{(x-\mu)}{\sigma}\right\} I_{[\mu, \infty)}(x)$$

con $\mu \in \mathbb{R}$ y $\sigma > 0$.

   (a) Exhibir un estadístico suficiente bidimensional para (μ, σ).

   (b) Hallar un estadístico suficiente para μ cuando σ es fijo y conocido.

   (c) Hallar un estadístico suficiente para σ cuando μ es fijo y conocido.

$\star$ 5. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución P_θ discreta o absolutamente continua. Supóngase que $T(\mathbf{X})$ es un estadístico suficiente para θ. Demostrar que el EMV de θ depende de $\mathbf{X}$ sólo a través de $T(\mathbf{X})$.

B) <u>*Teorema de Rao-Blackwell*</u>

1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $Bi(1, \theta)$. Luego, $\delta(\mathbf{X}) = X_1$ es un estimador insesgado de θ y $T(\mathbf{X}) = \sum_{i=1}^{n} X_i$ es un estadístico suficiente.

   (a) Usando el Teorema de Rao–Blackwell, hallar un estimador insesgado $\delta^*(T)$ que sea mejor que δ.

   (b) Probar que $V_\theta(\delta^*) < V_\theta(\delta) \quad \forall\, \theta$ si $n > 1$.

2. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $\mathcal{P}(\lambda)$ y $T = \sum_{i=1}^{n} X_i$ el estadístico suficiente.

   (a) Sea $\mu = e^{-\lambda}$. Mostrar que $\hat{\mu} = I\{X_1 = 0\}$ es un estimador insesgado de μ.

   (b) Aplicar el Teorema de Rao–Blackwell para obtener un estimador insesgado $\delta^*(T)$ mejor que $\hat{\mu}$ y comparar los ECM de ambos estimadores.

3. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $Bi(1, \theta)$. Dados m y k tales que $k \leq m < n$, se quiere estimar $q(\theta) = \binom{m}{k}\theta^k(1-\theta)^{m-k}$. Sea $S = \sum_{i=1}^{m} X_i$ y $h(S) = I\{S = k\}$.

   (a) Probar que $h(S)$ es un estimador insesgado de $q(\theta)$.

   (b) Construir un estimador insesgado de $q(\theta)$ mejor que $h(S)$.

*4. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $N(\mu, 1)$ y $\xi \in \mathbb{R}$ un número fijo y conocido. Se desea estimar $p = P(X_1 \leq \xi)$.

   (a) Mostrar que $I\{X_1 \leq \xi\}$ es un estimador insesgado de p.

   (b) A partir de $T = \bar{X}$, construir otro estimador insesgado $\delta^*(T)$ mejor que el anterior.
   Sugerencia: probar y usar que $\bar{X}$ y $X_1 - \bar{X}$ son independientes.

C) <u>*Familias exponenciales*</u>

1. Mostrar que las siguientes familias de distribuciones son exponenciales. En cada caso, hacer una elección explícita de $a(\theta)$, $h(\mathbf{x})$, $c_j(\theta)$ y $t_j(\mathbf{x})$ y exhibir un estadístico suficiente para θ.

   (a) $\mathcal{P}(\theta)$; $\theta > 0$.

   (b) $N(\mu, \sigma^2)$; $\theta = (\mu, \sigma^2)$, con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$.

   (c) $\Gamma(\alpha, \lambda)$; $\theta = (r, \lambda)$, con $\alpha, \lambda > 0$.

   (d) $\mathcal{M}(n, \theta_1, \ldots, \theta_k)$; $\theta = (\theta_1, \ldots, \theta_k)$, con $0 < \theta_j < 1$, $\sum_{j=1}^{k} \theta_j = 1$ y n conocido.

   (e) $BN(r, \theta)$; $0 < \theta < 1$ y r conocido.

   (f) $\beta(r, s)$; $\theta = (r, s)$, con $r, s > 0$.

2. Sea $\{f(x; \theta) : \theta \in \Theta\}$ una familia exponencial. Probar que el soporte de $f(x; \theta)$ no depende de θ. Mostrar que la familia $\{\mathcal{U}(0, \theta) : \theta > 0\}$ no es exponencial.

D) *Completitud y estimadores IMVU*

1. Analizar si los estadísticos suficientes hallados en los ejercicios 1 y 2 de la parte *A)* son completos.

2. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $N(\mu, 1)$.

   (a) Mostrar que $T = (X_1, \sum_{i=2}^n X_i)$ es suficiente para μ pero no completo.

   (b) Hallar un estadístico completo y un estimador IMVU de μ.

3. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $\mathcal{U}(0, \theta)$. Probar que $X_{(n)}$ es completo para θ.

4. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $Bi(1, \theta)$, con $\theta \in \Theta \subset (0, 1)$.

   (a) Mostrar que si Θ tiene más de n puntos, el estadístico $T = \sum_{i=1}^n X_i$ es completo.

   (b) Deducir que $\bar{X}$ es IMVU de θ.

   (c) Mostrar que $q(\theta) = \theta/(1-\theta)$ no es estimable mediante un estimador insesgado.

5. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $N(\mu, \sigma^2)$.

   (a) Mostrar que si σ^2 es conocido, $\bar{X}$ es IMVU para μ.

   (b) Mostrar que si μ es conocido, $\frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$ es IMVU para σ^2.

   (c) Mostrar que si ambos parámetros son desconocidos, entonces $\bar{X}$ es IMVU para μ y s^2 es IMVU para σ^2.

6. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $\mathcal{E}(\theta)$. Hallar un estimador IMVU de θ.

7. En el ejercicio 2 de la parte *B)*, hallar estimadores IMVU para λ, λ^2 y μ.

8. En el ejercicio 3 de la parte *B)*, hallar un estimador IMVU para $q(\theta)$.

* 9. En el ejercicio 4 de la parte *B)*, hallar un estimador IMVU para p.

* 10. Sean δ_1 y δ_2 dos estimadores IMVU de $q_1(\theta)$ y $q_2(\theta)$ respectivamente, y $c_1, c_2 \in \mathbb{R}$. Demostrar que $c_1\delta_1 + c_2\delta_2$ es IMVU de $c_1 q_1(\theta) + c_2 q_2(\theta)$.

E) *Desigualdad de Rao-Cramer*

1. Sea $X_1, \ldots, X_n$ una muestra aleatoria de una distribución $Bi(1, \theta)$.

   (a) Calcular el número de información de Fisher $I_1(\theta)$.

   (b) Mostrar que $\bar{X}$ alcanza la cota de Rao–Cramer y deducir que es IMVU para θ.

2. Sea $X_1, \ldots, X_n$ una m.a. de una distribución $N(\mu, \sigma_0^2)$, con σ_0^2 conocido.

   (a) Mostrar que $\bar{X}$ es un estimador IMVU para μ, usando la desigualdad de Rao–Cramer.

   (b) Mostrar que $\bar{X}^2 - \sigma_0^2/n$ es un estimador IMVU de μ^2, aunque no alcanza la cota de Rao–Cramer.

* 3. Sea X una v.a. con distribución $\mathcal{P}(\lambda)$ y sean μ y $\delta^*(T)$ como en el ejercicio 2 de la parte *B)*. Mostrar que aunque $\delta^*(T)$ es un estimador IMVU de μ, la cota de Rao–Cramer es estrictamente menor que $V_\lambda(\delta^*(T))$.