Examen General de Estadística

Semestre 2025-2 (Julio-Agosto 2025)

Instrucciones: Deberá responder todas las preguntas justificando sus respuestas.
Se requieren 4/6 de calificación para aprobar el examen.
Tiempo máximo de examen: 5 horas (9:00-14:00 hrs.).

Inferencia Estadística

1. Sean X_1 y X_2 variables aleatorias discretas con funciones de masa de probabilidad $p_1(\cdot)$ y $p_2(\cdot)$, respectivamente, definidas sobre los conjuntos disjuntos:

$$\text{soporte de } X_1 = \{1, 2, \ldots, m\}, \quad \text{soporte de } X_2 = \{m+1, \ldots, n\}.$$

Sea X una variable aleatoria definida como:

$$X = \begin{cases} X_1, & \text{con probabilidad } \theta, \\ X_2, & \text{con probabilidad } 1-\theta, \end{cases} \qquad \text{donde } 0 < \theta < 1.$$

Calcula la entropía de X, denotada $H(X)$, en términos de $H(X_1)$, $H(X_2)$ y θ.
Hint:

(1) Observa que se trata de una mezcla de distribuciones con soportes disjuntos.

(2) Considera introducir una variable aleatoria $Z \sim \text{Bernoulli}(\theta)$ que indique de qué componente proviene X.

2. Sean $X_1, \ldots, X_n$ una muestra aleatoria de $X \sim F_X(x)$, con $\mathbb{E}(X) = \mu$ y $\text{Var}(X) = \sigma^2$. Sea $X_1^*, \ldots, X_n^*$ una muestra aleatoria con reemplazo de tamaño n, generada a partir de la función de distribución empírica $\hat{F}_n(x)$ asociada a las observaciones. Definimos:

$$\bar{X}_n^* := \frac{1}{n} \sum_{i=1}^{n} X_i^*.$$

Calcula:

2.1. $\mathbb{E}(\bar{X}_n^* \mid X_1, \ldots, X_n)$.

2.2. $\text{Var}(\bar{X}_n^* \mid X_1, \ldots, X_n)$.

2.3. $\mathbb{E}(\bar{X}_n^*)$.

2.4. $\text{Var}(\bar{X}_n^*)$.

Hint. Recuerda que:

- $\mathbb{E}(Y) = \mathbb{E}\left[\mathbb{E}(Y \mid X)\right]$ y $\text{Var}(Y) = \mathbb{E}\left[\text{Var}(Y \mid X)\right] + \text{Var}\left[\mathbb{E}(Y \mid X)\right]$.
- $\mathbb{E}(\bar{X}_n) = \mu$, $\text{Var}(\bar{X}_n) = \sigma^2/n$ y $\mathbb{E}\left(S_n^2\right) = \sigma^2$, en donde $S_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2$.

3. Sea $X_{1:n}$ una muestra aleatoria proveniente de una distribución con función de densidad:

$$f(x \mid \theta) = \frac{\theta}{x^2} \mathbb{I}_{\{x \geq \theta\}}, \quad \theta > 0.$$

Responde lo siguiente:

3.1. Encuentra una estadística suficiente mínima $S_n = S(X_{1:n})$ para θ.

3.2. Determina la distribución exacta de S_n.

3.3. Considera a S_n como un estimador de θ:

i) ¿Es S_n insesgado para θ? En caso contrario, construye a partir de S_n un estimador insesgado T_n.

ii) Calcula la varianza de T_n.

iii) Obtén el error cuadrático medio (ECM) de T_n.

iv) Demuestra que T_n es un estimador consistente de θ.

v) Encuentra la distribución exacta de T_n.

Inferencia Bayesiana

4. Supongamos que $s_m = \{x_1, \ldots, x_m\}$ es una muestra aleatoria de una distribución Poisson $P(x|\theta)$. Supongamos además que

$$p(\theta) = \omega_1 Ga(\theta|a_1, b_1) + \omega_2 Ga(\theta|a_2, b_2) + \omega_3 Ga(\theta|a_3, b_3),$$

donde $\omega_1 + \omega_2 + \omega_3 = 1$, $\omega_1, \omega_2, \omega_3 \geq 0$ y $a_1, a_2, a_3, b_1, b_2, b_3 > 0$ y todos son conocidos.

(a) Demuestre que la distribución inicial es conjugada de la Poisson, encontrando los nuevos pesos y los parámetros finales.

(b) Supongamos ahora que y es condicionalmente independiente dado θ de s_m y encuentra la densidad predictiva

$$p(y|s_m) = \int_0^\infty p(y|\theta) p(\theta|s_m)\, d\theta.$$

(c) Finalmente, hagamos $(\omega_1, \omega_2, \omega_3) = (0.4, 0.2, 0.4)$, $s_m = \{0, 2, 2, 2, 0\}$ y que $(a_1, b_1) = (2, 1)$, $(a_2, b_2) = (1, 2)$ y $(a_3, b_3) = (2, 2)$. Usando los resultados anteriores, calcula $p(y = 1|s_m)$.

5. Se sabe que el tiempo de espera (en segundos) entre los registros de dos emisiones consecutivas de un trazador radioactivo es una variable aleatoria T con densidad de probabilidad

$$p(t|\theta) \propto e^{-\theta t} I_{(0.1,\infty)}(t) \qquad (\theta > 0),$$

donde el *truncamiento* en 0.1 se debe a que el contador permanece bloqueado una décima de segundo cuando registra una señal. Supongamos que la información inicial sobre θ puede describirse mediante una distribución $\mathrm{Ga}(\theta|1, \beta)$.

(a) Encuentre la forma del estimador bayesiano de $\mu = \theta^{-1}$ bajo una función de pérdida cuadrática.

(b) ¿Qué pasa con este estimador conforme $\beta \to 0$? Discute este resultado.

6. Para llegar a su clase de estadística bayesiana, una estudiante puede viajar en metro o en bicicleta. La estudiante quiere establecer con cuál de estos dos medios de transporte tiene mayor posibilidad de llegar a tiempo a su clase, basándose en su experiencia previa: en $n = 7$ ocasiones que tomó el metro, siempre llegó a tiempo; de $m = 9$ ocasiones que viajó en bicicleta, llegó tarde la primera vez. Denotemos por θ y ψ las probabilidades de llegar a tiempo a su clase cuando viaja en metro y en bicicleta, respectivamente. Supón que estas pruebas son independientes y asigna distribuciones iniciales $\mathrm{U}(0, 1)$ a θ y a ψ .

(a) Obtén la distribución final de (θ, ψ).

(b) Proporciona una estimación bayesiana para $\lambda = \theta - \psi$.

(c) Encuentra un intervalo de máxima densidad para θ al 95 %.

(d) Verifica la hipótesis $H_0 : \theta > \psi$.