Estadística bayesiana

Tarea 2

Fecha de entrega: 13 de octubre

1. Sea $X \mid \theta \sim \text{Bin}(n, \theta)$ con n conocido y donde se asume que a priori $\theta \sim \text{Be}(\sqrt{n}/2, \sqrt{n}/2)$.

   (a) Encuentre el estimador de Bayes d^* bajo una pérdida cuadrática y obtenga el riesgo de Bayes.

   (b) Considere ahora el estimador $d_0(x) = x/n$. Encuentre el riesgo de Bayes y compárelo con el de d^*. ¿Qué sucede cuando se toma $n = 10, 50$ y 100?

2. Sea $X \mid \theta \sim \mathcal{N}(\theta, 1)$ donde se asume que a priori $\theta \sim \mathcal{N}(0, n)$.

   (a) Encuentra el riesgo bayesiano bajo una pérdida cuadrática. ¿Cómo se comporta dicho riesgo cuando n crece?

   (b) Sea $n = 1$. Demuestre que bajo la función de pérdida,

   $$l(\theta, d) = \exp\left(\frac{3\theta^2}{4}\right)(\theta - d)^2,$$

   el estimador bayesiano es $\tilde{d}(x) = 2x$.

   (c) ¿Cuál es el riesgo bayesiano asociado a $\tilde{d}(x)$ y cómo se compara con el del inciso (a)?

3. Sea $X_1, \ldots, X_n$ una muestra aleatoria con distribución Pareto(α, β), cuya función de densidad está dada para $x > \beta$ por

   $$f(x \mid \alpha, \beta) = \frac{\beta\alpha^\beta}{x^{\beta+1}}.$$

   Supóngase que se observa $Y = \min\{X_1, \ldots, X_n\}$ y que α es conocida pero β no.

   (a) Muestre que la distribución gamma es una familia conjugada para el parámetro β.

   (b) Encuentre el estimador bayesiano bajo una función de pérdida cuadrática.

4. Sea $X_1, \ldots, X_n$ una colección de variables que se asumen intercambiables tales que $X_i \mid \theta \sim \text{Ber}(\theta)$.

(a) Utilizando la distribución impropia $f(\theta) \propto \theta^{-1}(1-\theta)^{-1}$ encuentre la distribución posterior de $\theta \mid \mathbf{x}$.

(b) Obtén la aproximación de Laplace para esta distribución posterior.

(c) Muestra que la distribución inicial $f(\theta)$ es equivalente a una distribución uniforme para

$$\beta \quad = \quad \log\left(\frac{\theta}{1-\theta}\right).$$

(d) Para dicha transformación, encuentra la distribución posterior y la aproximación de Laplace correspondiente.

(e) ¿Para qué parametrización tiene más sentido utilizar la aproximación de Laplace?

5. Considera la integral

$$I \quad = \quad \int_0^{10} \exp(-2|x-5|)dx.$$

(a) Suponga que $X \sim \mathrm{Unif}(0,10)$. Mostrar que la integral se puede ver como una esperanza con respecto a dicha distribución. De esta forma, utilizando el lenguaje de programación de tu preferencia, deriva una aproximación a I utilizando integración de Monte Carlo.

(b) Explica como se puede estimar I utilizando el método de muestreo por importancia con distribución instrumental $g(x) = \mathcal{N}(x \mid 5,1)$. Detalla el algoritmo e impleméntalo.

(c) ¿Cuál de los dos métodos anteriores prefieres?

6. Sea $X \sim \mathcal{N}(0,1)$.

(a) Explica cómo utilizar la integración de Monte Carlo para encontrar $\mathbb{P}(X > a)$ para $a \in \mathbb{R}$.

(b) ¿Cuáles son las dificultades para valores grandes de a?

(c) Supóngase que se desea estimar dicha probabilidad pero utilizando el método de muestreo por importancia utilizando como distribución instrumental $g = \mathcal{N}(\mu,1)$. Considerando los casos $a = 3$ y $\mu = 4$ y $a = 4.5$ y $\mu = 4.5$, comente las ventajas de este enfoque comparado con simular directamente de una distribución $\mathcal{N}(0,1)$.

(d) Explica cómo encontrar $\mathbb{P}(X > 4.5)$ utilizando el muestro por importancia con distribución instrumental dada por una exponencial de parámetro $\lambda = 1$ y truncada en 4.5 cuya densidad está dada por

$$g(x) \quad = \quad \exp(-(x-4.5))\mathbb{I}(x)_{(4.5,\infty)}$$

7. Demuestra que para el muestreo por importancia, se tiene que

$$g^*(x) = \frac{|h(x)|f(x)}{\int |h(t)|f(t)dt},$$

minimiza la varianza del estimador

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} h(x_i)w(x_i).$$

8. Suponga que bajo un enfoque bayesiano se tiene que $f(\theta \mid x) = cg(\theta)$ donde $g(\theta) \propto f(x \mid \theta)f(\theta)$ y c es la constante de normalización. Para una función $h(\theta)$ se desea encontrar $\mathbb{E}(h(\theta) \mid x)$. Suponga además que se tiene una muestra $\theta_1, \ldots, \theta_N$ provenientes de una distribución $q(\theta)$ cuyo soporte es el mismo que el de $f(\theta \mid x)$.

   (a) Si c es conocida ¿cómo se puede estimar $\mathbb{E}(h(\theta) \mid x)$ a partir de la muestra generada de $q(\theta)$?

   (b) Si c es desconocida, ¿cómo se podría estimar c a partir de la muestra generada de $q(\theta)$?

   (c) Finalmente, detalle cómo estimar $\mathbb{E}(h(\theta) \mid x)$ cuando c es desconocida.

9. Sea $X \mid \theta \sim \mathcal{N}(\theta, 1)$ donde se asume que a priori que $\theta \sim \mathcal{C}(0, 1)$.

   (a) Utilizando la integración de Monte Carlo encuentra $\mathbb{E}(\theta \mid x)$ y $\mathsf{Var}(\theta \mid x)$. Realiza tus aproximaciones para $m = 10, 100, 1000, 10000$ y 100000 simulaciones.

   (b) ¿Cómo se comparan los estimadores del inciso anterior con los estimadores correspondientes a una distribución inicial $\mathcal{N}(0, 1)$?

10. Considera la distribución de Kumaraswamy definida en el intervalo $(0, 1)$ y cuya densidad está dada por

$$f(x \mid a, b) = abx^{a-1}(1 - x^a)^{b-1}.$$

   (a) Obtén la cdf de dicha distribución e implementa un algoritmo para simular una muestra de tamaño 10,000.

   (b) Utilizando el algoritmo de aceptación y rechazo con distribución uniforme como distribución instrumental, simula una muestra de tamaño 10,000. ¿Cuántas variables tuviste que generar hasta obtener las 10,000 deseadas?

   (c) Utilizando el muestreo por importancia con distribución uniforme como la distribución instrumental, genera una muestra de tamaño 10,000.

   (d) Obtener de forma analítica la esperanza de la distribución y compara para diferentes tamaños de muestra el estimador de Monte Carlo de los tres métodos anteriores.

Actividades de DataCamp

Realizar los siguientes cursos de manera individual

1. Introduction to Writing Functions in R
2. Intermediate R
3. Writing Efficient R Code