Estadística bayesiana y aplicaciones en ciencias de datos

Tarea 1

(En equipos de máximo 3 personas)

1. Sea $X := \{X_i\}_{i=1}^n \sim \mathrm{N}_n(0, \Sigma)$ con Σ determinado por

$$E(X_i^2) = 1 \quad \text{y} \quad E(X_i X_j) = \rho, \quad \text{para } i \neq j \text{ y } \rho > 0$$

   a) Para que la matriz Σ sea positiva definida, de tal forma que la distribución este bien definida, se requiere

$$\rho \geq -\frac{1}{n-1}.$$

   Si lo anterior se satisface para toda $n \geq 2$, ¿Podrías decir que $\{X_i\}_{i=1}^\infty$ forman una sucesión de variables aleatorias intercambiables? Justifica tu respuesta.

   b) Demuestra que la distribución empírica de las $\{X_i\}_{i=1}^n$ converge a una distribución gaussiana, *i.e.*

$$F_\infty(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \delta_{(-\infty,x]}(X_i) = \Phi\left[\frac{(x-Y)}{\sqrt{1-\rho}}\right]$$

   donde Φ denota la función de distribución acumulada de $\mathrm{N}(0,1)$ y

$$Y := \lim_{n\to\infty} \bar{X}_n \sim \mathrm{N}(0, \rho).$$

   c) Desde el punto de vista subjetivo de la estadística bayesiana, su relación con el concepto de intercambiabilidad, y utilizando el teorema de representación de Bruno de Finetti. ¿Cuál modelo, *e.g.* $f(x \mid \theta)$, y cuál distribución inicial, *e.g.* $\pi(\theta)$, escogerías?

2. Demuestra que para toda sucesión intercambiable, $\{X_i\}_{i=1}^\infty$, se tiene que $\mathsf{Corr}(X_i, X_j) \geq 0$

3. Sea $X^{(n)} := \{X_i\}_{i=1}^n$ un vector aleatorio tal que $X^{(n)} \mid \theta \sim \prod_{i=1}^n \mathsf{Po}(x_i;\theta)$[1]. Supongamos que se escoge una distribución inicial $q \in \mathscr{G}_\nu := \{\mathsf{GIG}(\theta;\lambda,\delta,\gamma); \nu := (\lambda,\delta,\gamma) \in \mathbb{R} \times \mathbb{O}_\lambda\}$ donde

$$\mathsf{GIG}\,(x;\lambda,\delta,\gamma) = \frac{\left(\frac{\gamma}{\delta}\right)^{\lambda/2}}{2K_\lambda(\sqrt{\delta\gamma})}\, x^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\delta x^{-1} + \gamma x\right)\right\}, \quad x > 0, \tag{1}$$

   con

$$\mathbb{O}_\lambda = \begin{cases} \delta \geq 0, \gamma > 0, & \text{si } \lambda > 0 \\ \delta > 0, \gamma > 0, & \text{si } \lambda = 0 \\ \delta > 0, \gamma \geq 0, & \text{si } \lambda < 0, \end{cases}$$

   y $K_\lambda(\cdot)$ la función modificada de Bessel del tercer tipo con indice ν. (ver Cuadro 1)

   a) Encuentra la distribución posterior $f_{\Theta|X}(\theta \mid x^{(n)})$. ¿Es $\mathscr{G}_\nu$ una familia conjugada para el parámetro de intensidad de una distribución Poisson?

   b) Encuentra la distribución predictiva. ¿Reconoces alguna familia paramétrica conocida?

   c) Encuentra el estimador bayesiano, i.e. $\mathbb{E}[\Theta \mid X^{(n)} = x^{(n)}]$, bajo una función de pérdida cuadrática. *Hint:* Utiliza el resultado

$$E[Z^r] = \frac{K_{\lambda+r}(\delta\gamma)}{K_\lambda(\delta\gamma)} \left(\frac{\delta}{\gamma}\right)^r,$$

   cuando $Z \sim \mathsf{GIG}(\lambda,\delta,\gamma)$.

---

[1] Notación alternativa para $X_i \mid \theta \overset{\text{iid}}{\sim} \mathsf{Po}(\theta)$ para $i = 1, \ldots, n$

4. Sea $X^{(n)}$ como en la pregunta anterior. Si escogemos $q \in \mathscr{G}_\nu := \{\mathrm{Iga}(\theta; a, b); \nu := (a,b) \in \mathbb{R}_+^2\}$, donde Iga denota la distribución de una variable aleatoria $Z = Y^{-1}$ con $Y \sim \mathsf{Ga}(a,b)$. $\mathsf{Ga}(a,b)$ denota una distribución gamma con media a/b.

   a) ¿Es $\mathscr{G}_\nu$ una familia conjugada para el parámetro de una distribución Poisson?

   b) ¿Reconoces alguna familia paramétrica conocida para la distribución posterior?

5. Sea $X \sim \mathscr{E}(\theta)$, donde $\mathscr{E}(\theta)$ denota una familia exponencial natural en una dimensión, es decir, con función de densidad dada por
$$f(x \mid \theta) = c(x)\, e^{\theta x - k(\theta)},$$
donde $k(\theta) := \log \int e^{\theta x}\, c(x) \eta(\mathrm{d}x)$ denota el *cumulante característico* y η es la medida de conteo o medida de Lebesgue. Encuentra las familias correspondientes cuando (*identifica quién es* $c(x)$)

   a) $k(\theta) = \theta^2/2$

   b) $k(\theta) = -\ln(-\theta), \quad \theta < 0$

   c) $k(\theta) = -\ln(1 - e^\theta), \quad \theta < 0$

   d) $k(\theta) = e^\theta$

   e) $k(\theta) = \ln(1 + e^\theta)$.

6. Sea $X \sim \mathsf{Bin}(n, \theta)$ con n conocido y una distribución inicial $q(\theta) = \mathsf{Be}(\theta; \sqrt{n}/2, \sqrt{n}/2)$.

   a) ¿Cuál es en estimador de Bayes, d^q, bajo una perdida cuadrática?

   b) ¿Cómo es el riesgo de Bayes, $r(q)$, para d^q?

   c) ¿Cómo se compara este riesgo de Bayes con el correspondiente al estimador $d_0(x) = x/n$, cuando $n = 10, 50$ y 100?

7. Sea $X \sim \mathsf{N}(\theta, 1)$ y $q(\theta) = \mathsf{N}(\theta; 0, n)$.

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

   b) Sea $n = 1$. Demuestra que bajo la función de pérdida $l(\theta, d) = e^{3\theta^2/4}(\theta - d)^2$ el estimador bayesiano es $d^q(x) = 2x$

   c) ¿Cuál es el riesgo bayesiano para este último estimador? ¿ Cómo se compara este estimador con el del inciso a)?

8. Considérese la siguiente función de pérdida
$$l(\theta, d) = e^{c(\theta - d)} - c(\theta - d) - 1.$$

   a) Demuestra que $l(\theta, d) > 0$ y dibuja su comportamiento, como función de $(\theta - d)$, cuando $c = 0{,}1, 0{,}5, 1, 2$.

   b) Proporciona una expresión para el estimador de Bayes bajo esta función de pérdida.

   c) Si $X_1, \ldots, X_n \sim \mathsf{N}(\theta, 1)$ y $q(\theta) = 1$, encuentra el estimador de Bayes correspondiente.

9. ¿Cuál es el estimador de Bayes correspondiente a una función de perdida $l(\theta, d) = |\theta - d|$?

10. Considérese las siguientes funciones de "pérdida intrínsecas":

$$l_e(\theta, d) = E_\theta\left[\log\left(\frac{f(x \mid \theta)}{f(x \mid d)}\right)\right] = \int_{\mathbb{X}} \log\left(\frac{f(x \mid \theta)}{f(x \mid d)}\right) f(x \mid \theta)\,\mathrm{d}x \quad \text{(Entropía)}$$

$$l_H(\theta, d) = E_\theta\left[\left(\sqrt{\frac{f(x \mid \theta)}{f(x \mid d)}} - 1\right)^2\right] = \int_{\mathbb{X}} \left(\sqrt{\frac{f(x \mid \theta)}{f(x \mid d)}} - 1\right)^2 f(x \mid \theta)\,\mathrm{d}x \quad \text{(Distancía de Hellinger)}$$

   a) Muestre que l_e y l_H son no-negativas e iguales a 0 cuando $d = \theta$.
   b) Bajo que condiciones $d = \theta$ es solución única de $l_e(\theta, d) = 0$ ($l_H(\theta, d) = 0$ respectivamente)
   c) Proporciona expresiones para ambas funciones de pérdida cuando $X \mid \theta \sim \mathsf{N}(0, \theta)$
   d) Si $X \mid \theta \sim \mathsf{Ga}(\alpha, \theta)$ y $q(\theta) = \mathsf{Ga}(\theta; \nu, x_0)$ encuentra el estimador de Bayes para θ bajo la función de pérdida de Hellinger.

11. Sea $X \sim \mathrm{N}(x; \theta, 1)$ y supón que se escoge una distribución inicial con densidad

$$q(\theta) = \frac{1}{2}\mathrm{N}(\theta; \mu, 1) + \frac{1}{2}\mathrm{N}(\theta; -\mu, 1)$$

   a) Encuentra la distribución posterior.
   b) Encuentra el estimador bayesiano bajo una función de pérdida cuadrática.
   c) Encuentra la distribución predictiva.

12. Supóngase que se observa una variable aleatoria con función de densidad dada por $f(x \mid \theta) = \theta^{-1}\, I(0 < x < \theta)$ y además se asume una distribución inicial $q(\theta) = \theta\, e^{-\theta}\, I(\theta > 0)$. Encuentra el estimador bayesiano bajo una función de pérdida cuadrática.

13. Sea $\{X_i\}_{i=1}^n$ una muestra aleatoria con distribución Pareto(α, β), cuya función de densidad está dada por

$$f(x) = \frac{\beta\gamma^\beta}{x^{\beta+1}}\, I(x \geq \gamma), \quad \beta, \gamma > 0$$

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

   a) Muestre que $\pi \in \mathscr{G}_\nu := \{\mathsf{Ga}(\beta; a, b); \nu := (a, b) \in \mathbb{R}_+^2\}$ es una familia conjugada para el parámetro β
   b) Encuentra el estimador bayesiano bajo una función de pérdida cuadrática.

Funciones de Bessel modificadas	
Propiedades	Expansión Asintótica cuando $x \downarrow 0$
1.- $K_{-\nu}(x) = K_\nu(x)$	1.- $K_\nu(x) \sim \frac{1}{2}\Gamma(\nu)\left(\frac{x}{2}\right)^{-\nu}, \quad \nu > 0$
2.- $K_{1/2}(x) = \sqrt{\frac{\pi}{2x}}e^{-x}$	2.- $K_\nu(x) \sim \frac{1}{2}\Gamma(-\nu)\left(\frac{x}{2}\right)^{\nu}, \quad \nu < 0$
3.- $K_{\nu+\varepsilon}(x) > K_\nu(x), \quad \nu, \varepsilon, x > 0$	3.- $K_0(x) \sim -\log(x)$
4.- $K_{\nu+1}(x) = \frac{2\nu}{x}K_\nu(x) + K_{\nu-1}(x)$	
Representación integral	Expansión Asintótica cuando $x \uparrow \infty$
$K_\nu(x) = \frac{1}{2}\int_0^\infty y^{\nu-1}\exp\left(-\frac{x}{2}(y + y^{-1})\right) dy$	$K_\nu(x) \sim \sqrt{\frac{\pi}{2x}}e^{-x}$
Derivadas $\partial/\partial x$	
1.- $K_0'(x) = -K_1(x)$	4.- $(\log K_\nu(x))' = \frac{\nu}{x} - R_\nu(x)$
2.- $K_\nu'(x) = -\frac{1}{2}(K_{\nu+1}(x) + K_{\nu-1}(x))$	
3.- $K_\nu'(x) = \frac{\nu}{x}K_\nu(x) - K_{\nu+1}(x)$	donde $R_\nu(x) := \frac{K_{\nu+1}(x)}{K_\nu(x)}$, $x > 0$

SUERTE!