Medidas de probabilidad aleatorias

Tarea 1

(En equipos de a lo máximo 2 gentes)

1. Sea $\{X_i\}_{i=1}^n \sim \mathrm{N}_n(0, \Sigma)$ 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, ¿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) Utilizando el teorema de representación de Bruno de Finetti. ¿Cuál es el modelo, *e.g.* $f(x \mid \theta)$, y la distribución inicial, *e.g.* $\pi(\theta)$, más adecuadas que justifican este modelo bayesiano?

2. Demuestra que para toda sucesión intercambiable, $\{X_i\}_{i=1}^\infty$, se tiene que:

   a) $\mathsf{Cov}(X_i, X_j) \geq 0$

   b) Las fidis correspondientes a $\{X_i\}_{i=1}^\infty$ satisfacen las propiedades de estacionariedad fuerte y reversibilidad.

3. Sea $\{X_t; t \geq 0\}$ un proceso de Lévy con valores en $\mathbb{R}_+$ no-decreciente y con terna $(0, 0, \nu)$, donde $\nu(\mathrm{d}x) = x^{\alpha-1}e^{-\beta x}\mathrm{d}x$, con $-1 < \alpha < \infty$ y $\beta > 0$

   a) ¿Para que valores de α, $\{X_t; t \geq 0\}$ se puede ver como un proceso Poisson compuesto? En caso afirmativo, ¿cuál es la distribución de los "brincos" correspondientes?.

   b) Si $\beta = 0$ y $\alpha \in (-1, 0)$, ¿Cuál es el exponente característico, ψ, correspondiente?

   c) ¿Para cuales valores de α, $\{X_t; t \geq 0\}$ es un proceso de Lévy con "actividad infinita"?

4. Sea $\xi := \{\xi_t; t \geq 0\}$ un subordinador con terna $(0, 0, \nu)$. Define una medida finita α sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ y el proceso normalizado $P(B) := \xi_{\alpha(B)}/\xi_a$, para todo $B \in \mathcal{B}(\mathbb{R})$ y donde $a := \alpha(\mathbb{R})$.

   a) Si $\nu(\mathbb{R}_+) = \infty$ y definimos $P_0(\cdot) := \alpha(\cdot)/a$ y $\mu_n(u) := \int_{\mathbb{R}_+} s^n e^{-us}\nu(ds)$, $n \in \mathbb{N}$, $u > 0$, demuestra que

$$\mathsf{Var}(P(B)) = P_0(B)\{1 - P_0(B)\}\, I_a \tag{1}$$

   donde $I_a := a\int_{\mathbb{R}_+} u e^{-a\psi(u)}\mu_2(u)\mathrm{d}u$ con $\psi(u) = \int_{\mathbb{R}_+}(1 - e^{-ux})\nu(dx)$. *Hint:*

$$\mathsf{E}[P(B)^2] = \int_{\mathbb{R}_+} u\, \mathsf{E}\left[e^{-u\xi_a}\, \xi_{\alpha(B)}^2\right] \mathrm{d}u$$

b) Evalúa (1) cuando $\nu(dx) = \frac{e^{-x}}{x}\,dx$

c) En el caso anterior, ¿Qué ocurre con la varianza cuando $a \to \infty$?

d) Si para una función g, $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$-medible, definimos la variable aleatoria ϕ_P mediante la funcional

$$\phi_P := \int_{\mathbb{R}} g(x) P(\mathrm{d}x).$$

Determina el valor de $\mathsf{E}[\phi_P]$.

e) Con la notación de arriba, si definimos

$$\lambda_P := \int_{\mathbb{R}} g^2(x) P(\mathrm{d}x) - \phi_P^2.$$

Demostrar que

$$E[\lambda_P]/\mathsf{Var}(\phi_P) = a. \tag{2}$$

Hint: muestra que $\mathsf{Var}(\phi_P) = \lambda_{P_0}/(a+1)$.

f) Simplifica (2) cuando $g(x) = x$ y $P_0 = \mathrm{N}(0, 1)$.

5. Si $P \sim$ N-Prior(α, ν), $B_1, B_2 \in \mathcal{B}(\mathbb{R})$ y $C := B_1 \cap B_2$, demuestra que

a)

$$\mathsf{Cov}(P(B_1), P(B_2)) = \frac{a\,\alpha(C) - \alpha(B_1)\alpha(B_2)}{a^2}\, I_a,$$

con I_a definido como en la pregunta 4.

b) y si B_1 y B_2 son disjuntos, entonces

$$\mathsf{Corr}(P(B_1), P(B_2)) = -\sqrt{\frac{\alpha(B_1)\alpha(B_2)}{[a - \alpha(B_1)][a - \alpha(B_2)]}}$$

6. Bajo el mismo esquema del último inciso de la pregunta anterior y utilizando la representación de Blackwell & MacQueen (1973)

$$\mathsf{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{i=1}^{n} \frac{a\,P_0(A_i) + \sum_{j=1}^{i-1} \delta_{X_j}(A_i)}{a + i - 1}, \quad n \geq 1 \tag{3}$$

a) ¿Cómo simularías un conjunto de observaciones de la distribución conjunta (3)? Describe un algoritmo.

b) Realiza un programa, en el lenguaje de tu preferencia, que implemente dicho algoritmo en el caso en donde $P_0 = \mathrm{N}(\mu, 1)$. Para $n = 100$ y $a = 1$, dibuja la función de distribución acumulada de los datos simulados (e.g. la distribución empírica observada) Repite dicho procedimiento 50 veces. ¿Qué observas?

c) ¿Que ocurre cuando a varía? En particular, cuando $a = 1, 5, 10, 50$ y cuando $a = 1, 1/2, 1/10, 1/100$.

El siguiente ejercicio es opcional.

A: (Punto extra) MPA vía pesos geométricos. Considera la medida de probabilidad aleatoria definida por

$$P(B) = \sum_{i=1}^{\infty} \lambda(1-\lambda)^{i-1} \delta_{X_i}(B), \tag{4}$$

donde $X_i \overset{\text{iid}}{\sim} P_0$, para $i = 1, \ldots, \infty$, e independientes de $\lambda \sim \mathsf{Beta}(a, b)$. ¿Puedes encontrar $\mathsf{Cov}(P(B_1), P(B_2))$, donde $B_1, B_2 \in \mathcal{B}(\mathbb{R})$?

Una respuesta satisfactoria (a mi juicio!) adicionará 1 punto a la calificación final del curso.

SUERTE!