Ejercicios EYP1113L Control 2

Primera parte

13-06-2024

Introducción

Este documento contiene ejercicios para probar algunas características del cálculo numérico en R. En particular se encuentra el cálculo de probabilidades conjuntas, condicionales y marginales, cálculo de probabilidades con TLC y extemos. Por último, se incluyen ejercicios con datos para estimar parámetros por gráfico de probabilidades (QQ-plot), método de momentos y máxima verosimilitud.

Probabilidades conjuntas

1. Considere el siguiente modelo para las variables aleatorias X e Y, con probabiliades conjuntas

	$Y=0$	$Y=1$	$Y=2$	$Y=3$
$X=-1$	0.04	0.16	0.04	0.12
$X=0$	0.16	0.12	0	0
$X=1$	0.12	0.16	0	0.08

   Calcule las siguientes probabilidades

   (a) $P(X=-1)$

   (b) $P(X=0 \mid Y=2)$

   (c) $P(Y<2 \mid X=1)$

   (d) $P(X=1 \mid Y<2)$

2. Considere el siguiente modelo conjunto

$$X \sim \text{Poisson}(7) \quad , \quad Y|X=x \sim \text{Binomial}(x, 0.3)$$

   Calcule las siguientes probabilidades

   (a) $P(X=7, Y=4)$

   (b) $P(Y \leq 4|X=9)$

   (c) $P(X=10|Y=7)$

   (d) $P(X \leq 10|Y=7)$.

   (**Ayuda:** Del modelo se deduce que Y tiene distribución marginal Poisson(2.1))

3. Considere el siguiente modelo conjunto

$$X \sim \text{Binomial}(12, 0.7) \quad , \quad Y|X=x \sim \text{Binomial}(x, 0.3)$$

   Calcule las siguientes probabilidades

   (a) $P(Y \geq 5)$

(b) $P(3 \leq Y \leq 7|X = 6)$

(c) $P(X = 6|Y = 3)$

(**Ayuda:** Del modelo se deduce que Y tiene distribución marginal Binomial$(12, 0.21)$)

4. Considere el siguiente modelo conjunto para el par de variables aleatorias (X, Y),

$$X \sim \text{Poisson}(\lambda = 3), \quad Y|X = x \sim \text{Normal}(x, 1)$$

Calcule las siguientes probabilidades

(a) $P(Y > -1 \mid X = 0)$

(b) $P(X = 0, Y > -1)$

(c) $P(0 \leq X \leq 1,\ Y > -1)$

(d) $P(Y > -1 \mid 0 \leq X \leq 1)$

5. Considere el siguiente modelo

$$(X, Y) \sim \text{NormalBivariada}(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho)$$

con $\mu_X = \mu_Y = 2$, $\sigma_X = 1$, $\sigma_Y = 2$, y $\rho = 0.3$. Calcule las siguientes probabilidades.

(a) $P(1 < X < 2; 1 < Y < 3)$

(b) $P(X < 2.3)$

(c) $P(X > 1; Y < 2.3)$

(**Ayuda:** Usar comando `pmvnorm` del paquete `mvtnorm`)

Respuestas:

1. (a) 0.36, (b) 0, (c) 0.7777778, (d) 0.3684211
2. (a) 0.0144891, (b) 0.9011913, (c) 0.1460138, (d) 0.2793449
3. (a) 0.0865932, (b) 0.25569, (c) 0.0601106
4. (a) 0.8413447, (b) 0.0418881, (c) 0.1878513 (d) 0.9432736
5. (a) 0.1347078, (b) 0.6179114, (c) 0.4428822

Probabilidades TLC y valores extremos

1. Si $X_1, \ldots, X_{60} \overset{\text{iid}}{\sim}$ Poisson(4), entonces calcular aproximadamente $P(\overline{X} \leq 4.2)$. (Recordar corrección por continuidad)
2. Si $Y_1, \ldots, X_{45} \overset{\text{iid}}{\sim}$ Normal$(4, 1)$, calcular $P(3 < \overline{Y} < 4.2)$.
3. Si $Z_1, \ldots, Z_{40} \overset{\text{iid}}{\sim}$ Gamma$(2, 2)$, calcular aproximadamente $P(\overline{Z} > 1.1)$.
4. Si $W_1, \ldots, W_{120} \overset{\text{iid}}{\sim}$ Bernoulli(0.2), calcular aproximadamente $P(\overline{W} \geq 0.25)$. (Recordar corrección por continuidad)
5. Si $X_1, \ldots, X_5 \overset{\text{iid}}{\sim}$ Uniforme$(3, 6)$, calcular la probabilidad $P(X_{(5)} > 5.5)$, donde $X_{(5)} = \max\{X_1, \ldots, X_5\}$.
6. Si $T_1, \ldots, T_{10} \overset{\text{iid}}{\sim}$ Exponencial(0.5), calcular $P(T_{(10)} < 2)$, donde $T_{(10)} = \max\{T_1, \ldots, T_{10}\}$.
7. Si $S_1, \ldots, S_4 \overset{\text{iid}}{\sim}$ Normal$(0, 2)$, calcular la probabilidad $P(S_{(1)} > -0.3)$, donde $S_{(1)} = \min\{S_1, \ldots, S_4\}$.

8. Si $W_1, \ldots, W_{14} \overset{\text{iid}}{\sim} \text{Beta}(3, 4)$, calcular $P(W_{(1)} \leq 0.15)$, donde $W_{(1)} = \min\{W_1, \ldots, W_{14}\}$

Respuestas:

(1) 0.7901297, (2) 0.9101438, (3) 0.1855467, (4) 0.1047035, (5) 0.5981224, (6) 0.0101859, (7) 0.0980767, (8) 0.4928458,

Estimación de parámetros

`ENS.xlsx`

La base de datos consiste en un extracto de la Encuesta Nacional de la Salud (ENS) tomada el 2009. Consiste en encuestar y medir ciertos indicadores corporales de cada persona, junto con su región de origen, edad, y otros factores. Datos se encuentra en hoja "ENS2009", y descripción de las variables se encuentra en la segunda hoja "Variables".

Responda las siguientes preguntas:

1. Para la variable Colesterol (`COLES`), ajustar por método de momentos una distribución Log-Normal(λ, ζ).
2. Para la variable Colesterol (`COLES`), ajustar por método de gráfico de probabilidad (QQplot) una distribución Log-Normal(λ, ζ).
3. Ajustar una distribución Logística(μ, σ) a la presión arterial diastólica (`PAD`) por método de máxima verosimilitud.
4. Para la misma presión arterial diastólica (`PAD`) ajustar distribución Logística(μ, σ) por método de gráfico de probabilidad (QQplot).
5. Para la variable Colesterol (`COLES`), ajustar por máxima verosimilitud una distribución Gamma(k, ν).
6. Estimar los parámetros de una distribución Gamma(k, ν) para ajustar a la presión arterial sistólica (`PAS`) por medio de método de momentos.
7. Ajustar una distribución Weibull(β, η) por máxima verosimilitud al nivel de colesterol de alta densidad (`HDL`).
8. Para la misma variable de colesterol `HDL`, ajustar una distribución Weibull(β, η) por método de gráfico de probabilidad (QQplot).

Respuestas:

(1) $\hat{\lambda} = 5.2478849$, $\hat{\zeta} = 0.2176618$, (2) $\hat{\lambda} = 5.2473349$, $\hat{\zeta} = 0.2217887$, (3) $\hat{\mu} = 75.8890199$, $\hat{\sigma} = 6.4495854$, (4) $\hat{\mu} = 76.3900596$, $\hat{\sigma} = 6.3019782$, (5) $\hat{k} = 20.8023593$, $\hat{\nu} = 0.1068317$, (6) $\hat{k} = 31.8377709$, $\hat{\nu} = 0.2479596$, (7) $\hat{\beta} = 3.6430157$, $\hat{\eta} = 52.2573649$, (8) $\hat{\beta} = 5.1366997$, $\hat{\eta} = 51.4488406$

`Abalon.xlsx`

Base de datos consiste en 400 locos (Abalón Rojo) observados en 4 centros de estudio de Chile. Se registran varias características físicas

- `largo`: El largo del loco (en cm)
- `diametro`: El diámetro del loco (en cm)
- `alto`: El alto del loco (en cm)
- `pesot`: Peso total, cuerpo más concha (en g)
- `pesocu`: Peso cuerpo (en g)
- `pesoco`: Peso concha (en g)

- `anillos`: Número de anillos en la concha
- `centro`: Nombre del centro de estudio

Responda las siguientes preguntas:

1. Ajustar un modelo Weibull$(\beta, 16)$ a la variable `largo`, estimando su parámetros β por máxima verosimilitud.
2. Ajustar una distribución Gamma$(k, 2)$ a la variable `diametro`, estimando su parámetro k por máxima verosimilitud.
3. A la variable número de `anillos`, ajustar una distribución Poisson(λ) por medio del método de momentos.

Respuestas:

(1) $\hat{\beta} = 5.1846195$, $\eta = 16$, (2) $\hat{k} = 20.4652057$, $\nu = 2$, (3) $\hat{\lambda} = 9.95$.