

MAT-269: Sesión 8

Estimación ML bajo distribuciones de contornos elípticos II

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM



Un vector aleatorio $\mathbf{x} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ ¹ con $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} > \mathbf{0}$ tiene función de densidad

$$f(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})], \quad (1)$$

si y sólo si $g : \mathbb{R} \rightarrow [0, \infty)$ es tal que

$$\int_0^\infty u^{p/2-1} g(u) \, du < \infty,$$

y decimos que $g(\cdot)$ es la **generadora de densidad**.

¹Cuando $\mathbf{x} \sim EC_p(\mathbf{0}, \mathbf{I}; g)$ es usual escribir $\mathbf{x} \sim S_p(g)$.



Definición 1

Sea \mathbf{u} vector aleatorio $p \times 1$ con **distribución uniforme** sobre el conjunto

$$\mathcal{S}_p = \{\mathbf{x} \in \mathbb{R}^p : \|\mathbf{x}\| = 1\}, \quad (2)$$

esto es \mathcal{S}_p denota la **superficie de la esfera unitaria** en \mathbb{R}^p . En cuyo caso anotamos $\mathbf{u} \sim U(\mathcal{S}_p)$.

Resultado 1

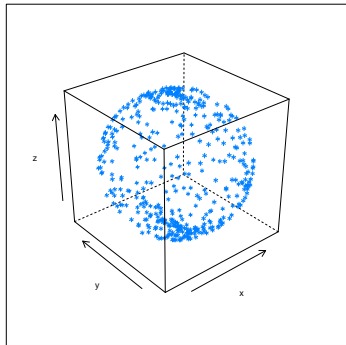
Si z_1, \dots, z_p son variables aleatorias IID con distribución $N(0, 1)$, entonces $\mathbf{u} = (u_1, \dots, u_p)^\top$, definido como:

$$\mathbf{u} = \frac{\mathbf{z}}{\|\mathbf{z}\|},$$

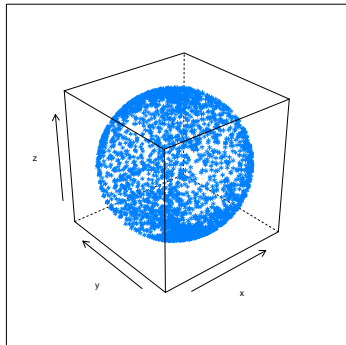
tiene **distribución uniforme sobre la esfera unitaria**.



Distribuciones de contornos elípticos



(a) $n = 750$



(b) $n = 3000$

Resultado 2

Suponga que $\mathbf{x} \sim S_p(g)$. Entonces \mathbf{x} tiene representación estocástica

$$\mathbf{x} \stackrel{d}{=} R\mathbf{u}, \quad (3)$$

donde $\mathbf{u} \sim U(\mathcal{S}_p)$ y $R \geq 0$ con distribución G , son independientes.

Resultado 3

Suponga que $\mathbf{x} \stackrel{d}{=} R\mathbf{u} \sim S_p(g)$ ($P(\mathbf{x} = \mathbf{0}) = 0$), entonces

$$\|\mathbf{x}\| \stackrel{d}{=} R, \quad \frac{\mathbf{x}}{\|\mathbf{x}\|} \stackrel{d}{=} \mathbf{u}.$$

Además $\|\mathbf{x}\|$ y $\mathbf{x}/\|\mathbf{x}\|$ son independientes.



Resultado 4

El vector de medias y la matriz de covarianza de $\mathbf{u} \sim U(\mathcal{S}_p)$ son:

$$E(\mathbf{u}) = \mathbf{0}, \quad \text{Cov}(\mathbf{u}) = \frac{1}{p} \mathbf{I}_p,$$

respectivamente.

Demostración:

Sea $\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{I})$, tenemos que $\mathbf{x} \stackrel{d}{=} \|\mathbf{x}\| \mathbf{u}$, con $\|\mathbf{x}\|$ independiente de \mathbf{u} . Sabemos que $\|\mathbf{x}\|^2 \sim \chi^2(p)$. Dado que

$$E(\mathbf{x}) = \mathbf{0}, \quad E(\|\mathbf{x}\|) > 0, \quad \text{y} \quad E(\|\mathbf{x}\|^2) = p, \quad \text{Cov}(\mathbf{x}) = \mathbf{I}_p,$$

es resultado sigue.



Resultado 5

Si $\mathbf{x} \stackrel{d}{=} R\mathbf{u} \sim S_p(g)$ y $E(R^2) < \infty$. Entonces,

$$E(\mathbf{x}) = \mathbf{0}, \quad \text{Cov}(\mathbf{x}) = \frac{E(R^2)}{p} \mathbf{I}_p,$$

respectivamente.

Demostración:

En efecto, como R y \mathbf{u} son independientes, sigue que

$$E(\mathbf{x}) = E(R) E(\mathbf{u}) = \mathbf{0},$$

$$\text{Cov}(\mathbf{x}) = E(R^2) E(\mathbf{u}\mathbf{u}^\top) = E(R^2) \text{Cov}(\mathbf{u}) = \frac{E(R^2)}{p} \mathbf{I}_p,$$

siempre que $E(R) < \infty$ y $E(R^2) < \infty$.



Resultado 6

$\mathbf{x} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, $\boldsymbol{\Sigma} > \mathbf{0}$, si y sólo si

$$\mathbf{x} \stackrel{d}{=} \boldsymbol{\mu} + R\mathbf{B}\mathbf{u},$$

donde $R \geq 0$ es independiente de \mathbf{u} y \mathbf{B} es matriz tal que $\boldsymbol{\Sigma} = \mathbf{B}\mathbf{B}^\top$.

Resultado 7

Suponga que $\mathbf{x} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ y $E(R^2) < \infty$. Entonces

$$E(\mathbf{x}) = \boldsymbol{\mu}, \quad \text{Cov}(\mathbf{x}) = \frac{E(R^2)}{p} \boldsymbol{\Sigma}.$$



Distribuciones de contornos elípticos

Considere $\mathbf{x} \stackrel{d}{=} R\mathbf{u} \sim S_p(g)$, la densidad de R es determinada desde $g(\mathbf{x}^\top \mathbf{x})$ mediante transformar en coordenadas polares y luego obtener su distribución marginal. La densidad es dada por:

$$h(r) = \frac{2\pi^{p/2-1}}{\Gamma(p/2)} r^{p-1} g(r^2).$$

De ahí que $E(R^k) < \infty$ si y solo si,

$$\int_0^\infty r^{k+p-1} g(r^2) dr < \infty$$

La función característica de \mathbf{x} es

$$\phi(\mathbf{t}^\top \mathbf{t}) = E\{\exp(iR\mathbf{t}^\top \mathbf{u})\} = \int_0^\infty \psi(r^2 \mathbf{t}^\top \mathbf{t}) h(r) dr,$$

con $\psi(\mathbf{s}^\top \mathbf{s}) = E(e^{i\mathbf{s}^\top \mathbf{u}})$ la función característica de \mathbf{u} .



Distribuciones de contornos elípticos

Sabemos que²

$$E(\mathbf{x}) = \boldsymbol{\mu}, \quad E\{(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\top\} = \frac{E(R^2)}{p} \boldsymbol{\Sigma},$$

además

$$E\{(\mathbf{x} - \boldsymbol{\mu}) \otimes (\mathbf{x} - \boldsymbol{\mu})^\top \otimes (\mathbf{x} - \boldsymbol{\mu})\} = \mathbf{0}.$$

En efecto, todos los momentos impares de $\mathbf{x} - \boldsymbol{\mu}$ son cero. Los primeros momentos de R están relacionados con la función característica $\phi(\cdot)$,

$$E(R^2) = -2p\phi'(0), \quad E(R^4) = 4p(p+2)\phi''(0).$$

El **exceso de curtosis**, adopta la forma:

$$\begin{aligned} \frac{E\{(x_i - \mu_i)^4\} - 3[E\{(x_i - \mu_i)^2\}]^2}{[E\{(x_i - \mu_i)^2\}]^2} &= \frac{3E(R^4)/(p(p+2)) - 3(E(R^2)/p)^2}{(E(R^2)/p)^2} \\ &= 3\left(\frac{E(R^4)}{E^2(R^2)} \frac{p}{p+2} - 1\right) = 3\kappa. \end{aligned}$$

²Además, podemos escribir $\text{Cov}(\mathbf{x}) = -2\phi'(0)\boldsymbol{\Sigma}$.



Distribuciones de contornos elípticos

- ▶ **Normal:** $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con

$$g(u) = c_1 \exp(-u/2).$$

- ▶ **t-Student:** $\mathbf{x} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, donde

$$g(u) = c_2(1 + u/\nu)^{-(\nu+p)/2}, \quad \nu > 0.$$

- ▶ **Normal contaminada:** $\mathbf{x} \sim CN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$, con $\epsilon \in [0, 1)$ y $\gamma > 0$,

$$g(u) = c_1 \left\{ (1 - \epsilon) \exp(-u/2) + \epsilon \gamma^{-p/2} \exp(-u/(2\gamma)) \right\}.$$

- ▶ **Cauchy:** $\mathbf{x} \sim \text{Cauchy}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con

$$g(u) = c_3(1 + u)^{-(p+1)/2}.$$

- ▶ **Logística:** $\mathbf{x} \sim L_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con

$$g(u) = c_4 \exp(-u) / \{1 + \exp(-u)\}^2.$$

- ▶ **Exponencial Potencia:** $\mathbf{x} \sim PE_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$, donde

$$g(u) = c_5 \exp(-u^\lambda/2), \quad \lambda > 0.$$



Estimación ML: caso independiente

Sea $\mathbf{x}_1, \dots, \mathbf{x}_n$ vectores aleatorios IID cada uno con densidad $EC_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. La densidad conjunta (de la muestra) es:

$$L(\boldsymbol{\theta}) = |\boldsymbol{\Sigma}|^{-n/2} \prod_{i=1}^n g[(\mathbf{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})].$$

La media muestral $\bar{\mathbf{x}}$ y la matriz de covarianza

$$\mathbf{S} = \frac{1}{n-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top,$$

son estimadores insesgados de $\boldsymbol{\mu}$ y

$$\boldsymbol{\Psi} = \frac{E(R^2)}{p} \boldsymbol{\Sigma},$$

respectivamente, donde $R^2 = (\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})$.



Estimación ML: caso independiente

Por la ley de los grandes números, tenemos

$$\bar{\mathbf{x}} \xrightarrow{P} \boldsymbol{\mu}, \quad \mathbf{S} \xrightarrow{P} \boldsymbol{\Sigma},$$

Además, (asintóticamente) $\text{Cov}(\bar{\mathbf{x}}) = \frac{1}{n}\boldsymbol{\Sigma}$, mientras que

$$\begin{aligned} n \text{Cov}(\text{vec } \mathbf{S}) &= \text{E}\{(\text{vec } \mathbf{S} - \text{vec } \boldsymbol{\Sigma})(\text{vec } \mathbf{S} - \text{vec } \boldsymbol{\Sigma})^\top\} \\ &\rightarrow (\kappa + 1)(\mathbf{I}_{p^2} + \mathbf{K}_p)(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) + \kappa(\text{vec } \boldsymbol{\Sigma})(\text{vec } \boldsymbol{\Sigma})^\top. \end{aligned}$$

Es decir,

$$\sqrt{n} \begin{pmatrix} \bar{\mathbf{x}} - \boldsymbol{\mu} \\ \text{vec } \mathbf{S} - \text{vec } \boldsymbol{\Sigma} \end{pmatrix} \rightarrow \text{N} \left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega} \end{pmatrix} \right),$$

donde

$$\boldsymbol{\Omega} = 2(\kappa + 1)\mathbf{N}_p(\boldsymbol{\Sigma} \otimes \boldsymbol{\Sigma}) + \kappa(\text{vec } \boldsymbol{\Sigma})(\text{vec } \boldsymbol{\Sigma})^\top,$$

con $\mathbf{N}_p = \frac{1}{2}(\mathbf{I}_{p^2} + \mathbf{K}_p)$.



Note que

$$\begin{aligned} E\{[(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})]^2\} &= E\{(\mathbf{z}^\top \mathbf{z})^2\} = E\{(R^2 \mathbf{u}^\top \mathbf{u})^2\} = E(R^4) E\{(\mathbf{u}^\top \mathbf{u})^2\} \\ &= \frac{p^2 E(R^4)}{E^2(R^2)} = p(p+2)(\kappa+1), \end{aligned}$$

pues $\mathbf{z} = \mathbf{B}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \sim S_p(g)$ y $\mathbf{z} \stackrel{d}{=} R\mathbf{u}$.

Como $\bar{\mathbf{x}} \xrightarrow{P} \boldsymbol{\mu}$ y $\mathbf{S} \xrightarrow{P} \boldsymbol{\Sigma}$,

$$\frac{1}{n} \sum_{i=1}^n [(\mathbf{x}_i - \bar{\mathbf{x}})^\top \mathbf{S}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})]^2 \xrightarrow{P} p(p+2)(\kappa+1).$$

De este modo, un **estimador consistente** para κ es (Mardia, 1970):³

$$\hat{\kappa} = \frac{1}{p(p+2)} \frac{1}{n} \sum_{i=1}^n [(\mathbf{x}_i - \bar{\mathbf{x}})^\top \mathbf{S}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})]^2 - 1.$$

³Biometrika 57, 519-530.



Estimación ML: caso independiente

Para obtener los **estimadores máximo verosímiles** de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ asumiremos que $g(\cdot)$ es conocido. La función de log-verosimilitud adopta la forma:

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \log |\boldsymbol{\Sigma}| + \sum_{i=1}^n \log g[(\mathbf{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})].$$

Diferenciando $\ell(\boldsymbol{\theta})$ con relación a $\boldsymbol{\mu}$, obtenemos⁴

$$\begin{aligned} d_{\boldsymbol{\mu}} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n d_{\boldsymbol{\mu}} \log g(u_i) = \sum_{i=1}^n \frac{g'(u_i)}{g(u_i)} d_{\boldsymbol{\mu}} u_i \\ &= - \sum_{i=1}^n W_g(u_i) [(d\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) + (\mathbf{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} d\boldsymbol{\mu}] \\ &= (d\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} \sum_{i=1}^n \omega_i(\boldsymbol{\theta}) (\mathbf{x}_i - \boldsymbol{\mu}), \end{aligned}$$

donde $\omega_i(\boldsymbol{\theta}) = -2W_g(u_i)$ con $W_g(u_i) = g'(u_i)/g(u_i)$, para $i = 1, \dots, n$.

⁴ Considere $u_i = (\mathbf{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})$ para $i = 1, \dots, n$.



Análogamente, diferenciando $\ell(\boldsymbol{\theta})$ con relación a $\boldsymbol{\Sigma}$, tenemos

$$\begin{aligned}d_{\boldsymbol{\Sigma}} \ell(\boldsymbol{\theta}) &= -\frac{n}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} d \boldsymbol{\Sigma} + \sum_{i=1}^n W_g(u_i) (\mathbf{x}_i - \boldsymbol{\mu})^{\top} (d \boldsymbol{\Sigma}^{-1}) (\mathbf{x}_i - \boldsymbol{\mu}) \\&= -\frac{n}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} d \boldsymbol{\Sigma} - \sum_{i=1}^n W_g(u_i) (\mathbf{x}_i - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} (d \boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \\&= -\frac{n}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} d \boldsymbol{\Sigma} - \frac{1}{2} \sum_{i=1}^n \omega_i(\boldsymbol{\theta}) \operatorname{tr} \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) (\mathbf{x}_i - \boldsymbol{\mu})^{\top} \boldsymbol{\Sigma}^{-1} d \boldsymbol{\Sigma} \\&= \frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} \{ \mathbf{Q}_{\omega}(\boldsymbol{\mu}) - n \boldsymbol{\Sigma} \} \boldsymbol{\Sigma}^{-1} d \boldsymbol{\Sigma}\end{aligned}$$

donde

$$\mathbf{Q}_{\omega}(\boldsymbol{\mu}) = \sum_{i=1}^n \omega_i(\boldsymbol{\theta}) (\mathbf{x}_i - \boldsymbol{\mu}) (\mathbf{x}_i - \boldsymbol{\mu})^{\top}.$$



Estimación ML: caso independiente

La condición de primer orden, lleva al siguiente sistema de ecuaciones:

$$\sum_{i=1}^n \omega_i(\hat{\boldsymbol{\theta}})(\mathbf{x}_i - \hat{\boldsymbol{\mu}}) = \mathbf{0}$$
$$\mathbf{Q}_\omega(\hat{\boldsymbol{\mu}}) - n\hat{\boldsymbol{\Sigma}} = \mathbf{0},$$

que **no** tiene solución en forma explícita y por tanto métodos iterativos son requeridos.

Por ejemplo, usando una estimación inicial $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, actualizamos las estimaciones para $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, como:

$$\boldsymbol{\mu}^{(k+1)} = \frac{1}{\sum_{i=1}^n \omega_i(\boldsymbol{\theta}^{(k)})} \sum_{i=1}^n \omega_i(\boldsymbol{\theta}^{(k)}) \mathbf{x}_i,$$
$$\boldsymbol{\Sigma}^{(k+1)} = \frac{1}{n} \sum_{i=1}^n \omega_i(\boldsymbol{\theta}^{(k)}) (\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})(\mathbf{x}_i - \boldsymbol{\mu}^{(k+1)})^\top,$$

a la convergencia del algoritmo⁵, hacemos $(\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}})$.

⁵Esto es, cuando la secuencia $\{\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}\}$ se 'estabiliza'.



Funciones de pesos $\omega(\theta)$ para algunas distribuciones elípticas

- ▶ **Normal:** $\mathbf{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, tenemos:

$$\omega_i(\boldsymbol{\theta}) = 1, \quad i = 1, \dots, n.$$

- ▶ **t-Student:** $\mathbf{x} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, $\nu > 0$,

$$\omega_i(\boldsymbol{\theta}) = \frac{\nu + p}{\nu + u_i}, \quad i = 1, \dots, n,$$

la distribución Cauchy $_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, es obtenida para $\nu = 1$.

- ▶ **Normal contaminada:** $\mathbf{x} \sim CN_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$, con $\epsilon \in [0, 1)$ y $\gamma > 0$,

$$\omega_i(\boldsymbol{\theta}) = \frac{(1 - \epsilon) \exp(-u/2) + \epsilon \gamma^{-(p/2+1)} \exp(-u/(2\gamma))}{(1 - \epsilon) \exp(-u/2) + \epsilon \gamma^{-p/2} \exp(-u/(2\gamma))}, \quad i = 1, \dots, n.$$

- ▶ **Logística:** $\mathbf{x} \sim L_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, en cuyo caso,

$$\omega_i(\boldsymbol{\theta}) = 2 \tanh(u_i/2), \quad i = 1, \dots, n.$$

- ▶ **Exponencial Potencia:** $\mathbf{x} \sim PE_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$, $\lambda > 0$,

$$\omega_i(\boldsymbol{\theta}) = \lambda u_i^{\lambda-1}, \quad i = 1, \dots, n.$$



Ejemplo: Datos de creatinina (Shih y Weisberg, 1986)⁶

La **eliminación de creatinina** a través de la orina es una medida importante de la **función renal**.

Los datos corresponden aun estudio para determinar la eliminación de creatinina en 34 pacientes registrando las siguientes variables:

- ▶ **peso corporal (WT)** en kg.,
- ▶ **concentración de creatinina sérica (SC)** en mg/decilitro,
- ▶ **Edad (Age)** en años, y
- ▶ la **depuración de creatinina** endógena (CR).

Un modelo típico sugerido en farmacocinética es:

$$E\{\log(\text{CR})\} = \beta_0 + \beta_1 \log(\text{WT}) + \log_2 \log(\text{SC}) + \beta_3 \log(140 - \text{Age}).$$

A continuación asumiremos que $(\log(\text{CR}), \log(\text{WT}), \log(\text{SC}), \log(140 - \text{Age}))^T$ sigue una **distribución normal multivariada**.

⁶Technometrics 28, 231-239.



Ejemplo: Datos de creatinina

```
> library(heavy)
> data(creatinine)
> creatinine
```

	WT	SC	Age	CR
1	71.0	0.71253	38	132.0
2	69.0	1.48161	78	53.0
3	85.0	2.20545	69	50.0
4	100.0	1.42505	70	82.0
5	59.0	0.67860	45	110.0
6	73.0	0.75777	65	100.0
7	63.0	1.11969	76	68.0
8	81.0	0.91611	61	92.0
...				
26	67.0	1.19886	21	80.0
27	68.0	7.60001	81	4.3
28	72.2	2.10001	43	43.2
29	NA	1.35719	78	75.0
30	NA	1.05183	38	41.0
31	107.0	NA	62	120.0
32	75.0	NA	70	52.0
33	62.0	NA	63	73.0
34	52.0	NA	68	57.0



Ejemplo: Datos de creatinina

```
# ajuste con el modelo normal
> fm0 <- heavyFit(~ log(WT) + log(SC) + log(140 - Age) + log(CR),
+               data = creatinine, family = normal())

> fm0 # Salida:
Call:
heavyFit(x = ~log(WT) + log(SC) + log(140 - Age) + log(CR),
        data = creatinine, family = normal())
Converged in 1 iterations

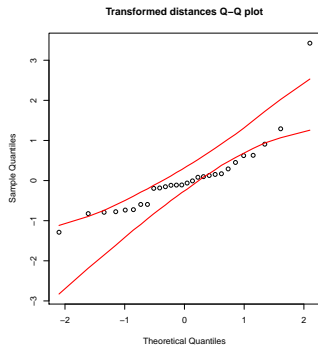
Center:
      log(WT)          log(SC) log(140 - Age)          log(CR)
4.2763         0.2038         4.4370         4.3069

Scatter matrix estimate:
      log(WT)          log(SC)          log(140 - Age)          log(CR)
log(WT)          0.023876
log(SC)          -0.000072          0.248036
log(140 - Age)  -0.002317          -0.051800          0.0455740
log(CR)          0.022111          -0.307312          0.0873372          0.457392

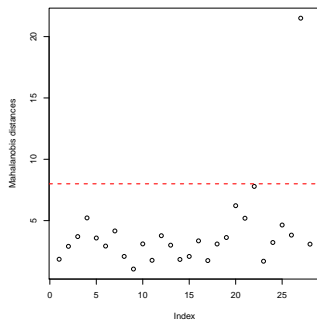
Number of Observations: 28
```



Ejemplo: Datos de creatinina



(a)



(b)

Ejemplo: Datos de creatinina

```
# ajuste con el modelo Student-t
> fm1 <- heavyFit(~ log(WT) + log(SC) + log(140 - Age) + log(CR),
+               data = creatinine, family = Student(df = 4))

> fm1
Call:
heavyFit(x = ~log(WT) + log(SC) + log(140 - Age) + log(CR),
        data = creatinine, family = Student(df = 6.67344))
Converged in 38 iterations

Center:
      log(WT)      log(SC) log(140 - Age)      log(CR)
      4.2872      0.1262      4.4446      4.4247

Scatter matrix estimate:
      log(WT)      log(SC)      log(140 - Age)      log(CR)
log(WT)      0.021237
log(SC)      0.004514      0.138614
log(140 - Age) -0.004014      -0.030218      0.039962
log(CR)      0.012819      -0.133791      0.051319      0.180992

Number of Observations: 28
```



Ejemplo: Datos de creatinina

Debemos destacar que, en general:

$$\text{Cov}(\mathbf{x}) = \frac{E(R^2)}{p} \Sigma = \Psi.$$

Desde el procedimiento de estimación se obtuvo,⁷ $\hat{\nu} = 6.673439$. Para la distribución $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, sigue que:

$$\hat{\Psi}_{\text{Student}} = \frac{\hat{\nu}}{\hat{\nu} - 2} \hat{\Sigma} = \begin{pmatrix} 0.03032 & 0.00645 & -0.00573 & 0.01831 \\ 0.00645 & 0.19793 & -0.04315 & -0.19105 \\ -0.00573 & -0.04315 & 0.05706 & 0.07338 \\ 0.01831 & -0.19105 & 0.07328 & 0.25845 \end{pmatrix},$$

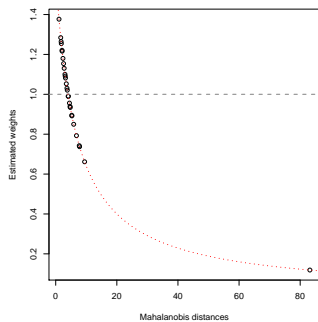
mientras que para la distribución $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, tenemos $\text{Cov}(\mathbf{x}) = \boldsymbol{\Sigma} (= \boldsymbol{\Psi})$. En efecto,

$$\hat{\Psi}_{\text{Normal}} = \begin{pmatrix} 0.02386 & -0.00007 & -0.00232 & 0.02211 \\ -0.00007 & 0.24804 & -0.05180 & -0.30731 \\ -0.00232 & -0.05180 & 0.04557 & 0.08734 \\ 0.02211 & -0.30731 & 0.08734 & 0.45739 \end{pmatrix}.$$

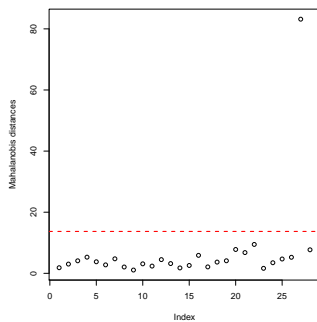
⁷Para detalles sobre como estimar ν , no se pierda la próxima clase! ☺



Ejemplo: Datos de creatinina



(a)



(b)

Referencias bibliográficas



Mardia, K.V. (1970).

Measures of multivariate skewness and kurtosis with applications.
Biometrika 57, 519-530.



Maronna, R.A. (1976).

Robust M -estimators of multivariate location and scatter.
The Annals of Statistics 4, 51-67.



Shih, W.J., Weisberg, S. (1986).

Assessing influence in multiple linear regression with incomplete data.
Technometrics 28, 231-239.



Xie, F.C., Wei, B.C., Lin, J.G. (207).

Case-deletion influence measures for the data from multivariate t distributions.
Journal of Applied Statistics 34, 907-921.

