IECD-325: Distribuciones de contornos elípticos

Felipe Osorio

felipe.osorio@uv.cl

Familia de posición-escala

Considere U una variable aleatoria con función de distribución F. De este modo, la variable

$$X = U + a,$$

tiene función de distribución

$$\mathsf{P}(X \leq x) = F(x - a),$$

para F fijo y $a \in \mathbb{R}$ tenemos que X corresponde a la familia de posición.

Análogamente la familia de escala es generada por la transformación

$$X = bU, \qquad b > 0,$$

en cuyo caso,

$$\mathsf{P}(X \leq x) = F(x/b), \qquad b > 0.$$

Familia de posición-escala

Definición 1:

Sea U una variable aleatoria con función de distribución acumulada fija F y considere la transformación

$$X = a + bU, \qquad a \in \mathbb{R}, b > 0.$$

Tenemos

$$\mathsf{P}(X \leq x) = F\Big(\frac{x-a}{b}\Big).$$

De este modo, X es conocida como familia de posición-escala.[1]

Observación:

Usualmente asociado a F tenemos una función de densidad f, dada por

$$f(x; a, b) = \frac{\mathsf{d}}{\mathsf{d}x} F\Big(\frac{x-a}{b}\Big) = \frac{1}{b} F'\Big(\frac{x-a}{b}\Big) = \frac{1}{b} f\Big(\frac{x-a}{b}\Big).$$

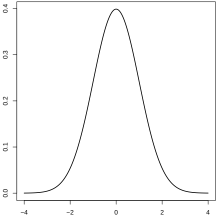
[1] a es el parámetro de posición, mientras que b es parámetro de escala.

Ejemplo (familias de posición-escala):

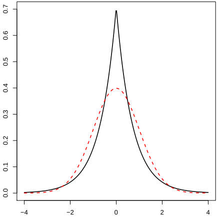
- $\mathsf{N}(a, b^2)$:
$$f(y; a, b) = \frac{1}{b}(2\pi)^{-1/2} \exp\Big\{ -\frac{1}{2}\Big(\frac{y-a}{b}\Big)^2 \Big\}.$$
- Laplace(a, b):[2]
$$f(y; a, b) = \frac{1}{2b} \exp\Big\{ -\frac{|y-a|}{b} \Big\}.$$
- Cauchy(a, b^2):
$$f(y; a, b) = \frac{b}{\pi} \frac{1}{b^2 + (y-a)^2}.$$
- Logística(a, b):
$$f(y; a, b) = \frac{1}{b} \frac{e^{-(y-a)/b}}{(1 + e^{-(y-a)/b})^2}.$$

[2] También es conocida como doble exponencial

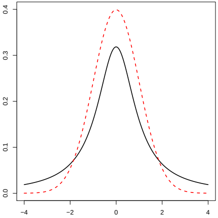
Familia de posición-escala



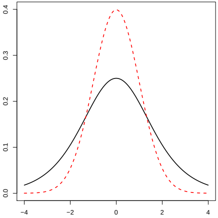
(a) Normal



(b) Laplace



(c) Cauchy



(d) Logística

Definición 2:

Sea $\boldsymbol{U}$ vector aleatorio $p \times 1$ con distribución uniforme sobre el conjunto

$$\mathcal{S}_p = \{\boldsymbol{x} \in \mathbb{R}^p : \|\boldsymbol{x}\| = 1\}, \tag{1}$$

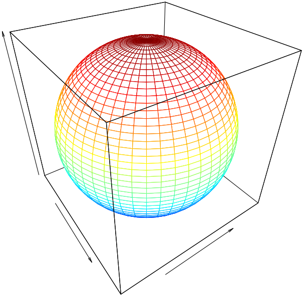
esto es $\mathcal{S}_p$ denota la superficie de la esfera unitaria en $\mathbb{R}^p$. En cuyo caso anotamos $\boldsymbol{U} \sim \mathsf{U}(\mathcal{S}_p)$.

Resultado 1:

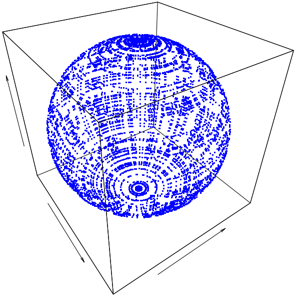
Si $Z_1, \dots, Z_p$ son variables aleatorias IID con distribución $\mathsf{N}(0,1)$, entonces $\boldsymbol{U} = (U_1, \dots, U_p)^\top$, definido como:

$$\boldsymbol{U} = \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|},$$

tiene distribución uniforme sobre la esfera unitaria.



(e) esfera unitaria



(f) datos simulados

Definición 3:

Un vector aleatorio $p \times 1$, $\boldsymbol{X}$ se dice que tiene simetría esférica si para cualquier $\boldsymbol{Q} \in \mathcal{O}_p$, sigue que

$$\boldsymbol{Q}\boldsymbol{X} \stackrel{\mathsf{d}}{=} \boldsymbol{X}.$$

Ejemplo:

Sea $\boldsymbol{U} \sim \mathsf{U}(\mathcal{S}_p)$, entonces es directo que $\boldsymbol{Q}\boldsymbol{U} \stackrel{\mathsf{d}}{=} \boldsymbol{U}$.

Definición 4:

Un vector aleatorio p-dimensional tiene distribución esférica sólo si su función característica satisface

(a) $\varphi(\boldsymbol{Q}^\top \boldsymbol{t}) = \varphi(\boldsymbol{t})$, para todo $\boldsymbol{Q} \in \mathcal{O}_p$.

(b) Existe una función $\psi(\cdot)$ de una variable escalar tal que $\varphi(\boldsymbol{t}) = \psi(\boldsymbol{t}^\top \boldsymbol{t})$.

En este caso escribimos $\boldsymbol{X} \sim \mathsf{S}_p(\psi)$.

Ejemplo:

Sea $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$, tenemos que

$$\varphi(\boldsymbol{t}) = \exp\{-\tfrac{1}{2}(t_1^2 + \cdots + t_p^2)\} = \exp(-\tfrac{1}{2}\boldsymbol{t}^\top \boldsymbol{t}).$$

Resultado 2:

Suponga que $\boldsymbol{X} \sim \mathsf{S}_p(g)$. Entonces $\boldsymbol{X}$ tiene representación estocástica

$$\boldsymbol{X} \stackrel{\mathsf{d}}{=} R\boldsymbol{U}, \tag{2}$$

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

Resultado 3:

Suponga que $\boldsymbol{X} \stackrel{\mathsf{d}}{=} R\boldsymbol{U} \sim \mathsf{S}_p(g)$ ($\mathsf{P}(\boldsymbol{X} = \boldsymbol{0}) = 0$), entonces

$$\|\boldsymbol{X}\| \stackrel{\mathsf{d}}{=} R, \qquad \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|} \stackrel{\mathsf{d}}{=} \boldsymbol{U}.$$

Además $\|\boldsymbol{X}\|$ y $\boldsymbol{X}/\|\boldsymbol{X}\|$ son independientes.

Resultado 4:

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

$$\mathsf{E}(\boldsymbol{U}) = \mathbf{0}, \qquad \mathsf{Cov}(\boldsymbol{U}) = \frac{1}{p}\boldsymbol{I}_p,$$

respectivamente.

Demostración:

Sea $\boldsymbol{X} \sim \mathsf{N}_p(\mathbf{0}, \boldsymbol{I})$, tenemos que $\boldsymbol{X} \stackrel{\mathsf{d}}{=} \|\boldsymbol{X}\|\boldsymbol{U}$, con $\|\boldsymbol{X}\|$ independiente de $\boldsymbol{U}$. Sabemos que $\|\boldsymbol{X}\|^2 \sim \chi^2(p)$. Dado que

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

es resultado sigue.

Resultado 5:

Si $\boldsymbol{X} \overset{\mathsf{d}}{=} R\boldsymbol{U} \sim \mathsf{S}_p(g)$ y $\mathsf{E}(R^2) < \infty$. Entonces,

$$\mathsf{E}(\boldsymbol{X}) = \boldsymbol{0}, \qquad \mathsf{Cov}(\boldsymbol{X}) = \frac{\mathsf{E}(R^2)}{p}\boldsymbol{I}_p,$$

respectivamente.

Demostración:

En efecto, como R y $\boldsymbol{U}$ son independientes, sigue que

$$\mathsf{E}(\boldsymbol{X}) = \mathsf{E}(R)\,\mathsf{E}(\boldsymbol{U}) = \boldsymbol{0},$$

$$\mathsf{Cov}(\boldsymbol{X}) = \mathsf{E}(R^2)\,\mathsf{E}(\boldsymbol{U}\boldsymbol{U}^\top) = \mathsf{E}(R^2)\,\mathsf{Cov}(\boldsymbol{U}) = \frac{\mathsf{E}(R^2)}{p}\boldsymbol{I}_p,$$

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

Definición 5:

Un vector aleatorio $p \times 1$, $\boldsymbol{X}$ tiene distribución de contornos elípticos con parámetros $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \geq \boldsymbol{0}$ si

$$\boldsymbol{X} \stackrel{\mathsf{d}}{=} \boldsymbol{\mu} + \boldsymbol{B}\boldsymbol{Y}, \qquad \boldsymbol{Y} \sim \mathsf{S}_k(\psi),$$

donde $\boldsymbol{B} \in \mathbb{R}^{k \times p}$ es matriz de rango completo tal que, $\boldsymbol{B}\boldsymbol{B}^\top = \boldsymbol{\Sigma}$ con $\operatorname{rg}(\boldsymbol{\Sigma}) = k$ y escribimos $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \psi)$.

Observación:

La función característica de $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \psi)$ es de la forma

$$\varphi(\boldsymbol{t}) = \exp(i\boldsymbol{t}^\top \boldsymbol{\mu})\psi(\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}).$$

Resultado 6:

$\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, $\boldsymbol{\Sigma} > \boldsymbol{0}$, si y sólo si

$$\boldsymbol{X} \stackrel{\mathsf{d}}{=} \boldsymbol{\mu} + R\,\boldsymbol{B}\boldsymbol{U},$$

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

Resultado 7:

Suponga que $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ y $\mathsf{E}(R^2) < \infty$. Entonces

$$\mathsf{E}(\boldsymbol{X}) = \boldsymbol{\mu}, \qquad \mathsf{Cov}(\boldsymbol{X}) = \frac{\mathsf{E}(R^2)}{p}\boldsymbol{\Sigma}.$$

Definición 6:

Un vector aleatorio $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$[3] con $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} > \mathbf{0}$ tiene función de densidad

$$f(\boldsymbol{x}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})], \qquad (3)$$

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

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

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

Observación:

Asuma que $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \psi)$ con $\operatorname{rg}(\boldsymbol{\Sigma}) = k$. Entonces,

$$U = (\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-}(\boldsymbol{X} - \boldsymbol{\mu}) \stackrel{\mathsf{d}}{=} R^2,$$

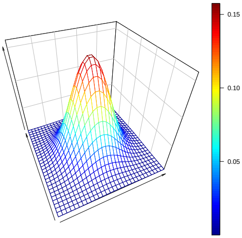
donde $\boldsymbol{\Sigma}^{-}$ es una inversa generalizada de $\boldsymbol{\Sigma}$.

[3] Cuando $\boldsymbol{X} \sim \mathsf{EC}_p(\mathbf{0}, \boldsymbol{I}; g)$ es usual escribir $\boldsymbol{X} \sim \mathsf{S}_p(g)$.

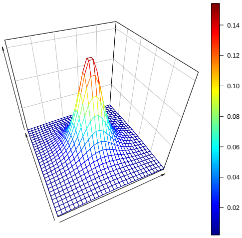
Distribuciones de contornos elípticos

- Normal: $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_1 \exp(-u/2).$$
- t-Student: $\boldsymbol{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, donde
$$g(u) = c_2(1 + u/\nu)^{-(\nu+p)/2}, \qquad \nu > 0.$$
- Normal contaminada: $\boldsymbol{X} \sim \mathsf{CN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$, con $\epsilon \in [0, 1)$ y $\gamma > 0$,
$$g(u) = c_1 \big\{(1-\epsilon)\exp(-u/2) + \epsilon\gamma^{-p/2}\exp(-u/(2\gamma))\big\}.$$
- Cauchy: $\boldsymbol{X} \sim \mathsf{Cauchy}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_3(1 + u)^{-(p+1)/2}.$$
- Logística: $\boldsymbol{X} \sim \mathsf{L}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_4 \exp(-u)/\{1 + \exp(-u)\}^2.$$
- Exponencial Potencia: $\boldsymbol{X} \sim \mathsf{PE}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$, donde
$$g(u) = c_5 \exp(-u^\lambda/2), \qquad \lambda > 0.$$

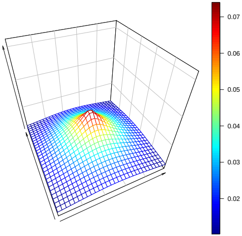
Distribuciones de contornos elípticos



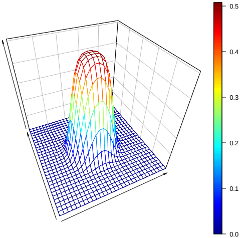
(a) normal



(b) Cauchy



(c) Laplace



(d) Exp. Potencia

Definición 7:

Sea $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Sigma}$ matriz $p \times p$ definida positiva y H función de distribución de una variable aleatoria positiva, W. Entonces, se dice que el vector aleatorio $\boldsymbol{X}$ sigue una distribución de mezcla de escala normal si su función de densidad asume la forma:

$$f(\boldsymbol{x}) = |2\pi\boldsymbol{\Sigma}|^{-1/2} \int_0^\infty \omega^{p/2} \exp(-\omega u/2) \, \mathrm{dH}(\omega),$$

donde $u = (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$ y anotamos $\boldsymbol{X} \sim \mathrm{SMN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathrm{H})$.

Observación:

Un vector aleatorio $\boldsymbol{X} \sim \mathrm{SMN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathrm{H})$ admite la representación:

$$\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + W^{-1/2}\boldsymbol{Z}, \tag{4}$$

donde $\boldsymbol{Z} \sim \mathrm{N}_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ y $W \sim \mathrm{H}(\boldsymbol{\delta})$ son independientes.

Distribuciones de mezcla de escala normal

Ejemplo (distribución Slash):

Un vector aleatorio $\boldsymbol{X}$ tiene distribución Slash si su función de densidad es de la forma:

$$f(\boldsymbol{x}) = \nu|2\pi\boldsymbol{\Sigma}|^{-1/2}\int_0^1 \omega^{p/2+\nu-1}\exp(-\omega u/2)\,\mathrm{d}\omega.$$

Tenemos que $h(\omega) = \nu\omega^{\nu-1}$, para $\omega \in (0,1)$ y $\nu > 0$. Es decir, $W \sim \text{Beta}(\nu, 1)$.

Ejemplo (distribución Exponencial-Potencia):

Se dice que un vector aleatorio $\boldsymbol{X}$ tiene distribución Exponencial-Potencia (Gómez, Gómez-Villegas y Marín, 1988)[4], si su función de densidad es dada por:

$$f(\boldsymbol{x}) = \frac{p\Gamma(\frac{p}{2})\pi^{-p/2}}{\Gamma(1+\frac{p}{2\lambda})2^{1+\frac{p}{2\lambda}}}|\boldsymbol{\Sigma}|^{-1/2}\exp(-u^{\lambda}/2), \qquad \lambda > 0.$$

en cuyo caso anotamos $\boldsymbol{X} \sim \mathsf{PE}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$. Debemos destacar que la distribución de la variable mezcladora W tiene una representación en series y es de poco interés práctico.

[4] Esta familia pertenece a la clase SMN cuando $\lambda \in (0,1]$.

Distribuciones de mezcla de escala normal

Observación:

La representación estocástica en (4), puede ser escrita de forma equivalente, como:

$$\boldsymbol{X}|W \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\omega), \qquad W \sim \mathsf{H}(\boldsymbol{\delta}). \tag{5}$$

Esta representación permite, por ejemplo

$$\begin{aligned} \mathsf{E}(\boldsymbol{X}) &= \mathsf{E}(\mathsf{E}(\boldsymbol{X}|W)) = \boldsymbol{\mu} \\ \mathsf{Cov}(\boldsymbol{X}) &= \mathsf{E}(\mathsf{Cov}(\boldsymbol{X}|W)) + \mathsf{Cov}(\mathsf{E}(\boldsymbol{X}|W)) = \mathsf{E}(W^{-1})\boldsymbol{\Sigma}. \end{aligned}$$

Además, la formulación condicional en (5) es muy útil para:

- Generación de dígitos pseudo-aleatorios.
- Estimación ML usando el algoritmo EM.

Ejemplo (distribución t multivariada):

Para $\boldsymbol{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, con $\nu > 0$, podemos escribir

$$\boldsymbol{X}|W \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\omega), \qquad W \sim \mathsf{Gamma}(\nu/2, \nu/2),$$

es decir,

$$h(\omega; \nu) = \frac{(\nu/2)^{\nu/2}\omega^{\nu/2-1}}{\Gamma(\nu/2)} \exp(-\nu\omega/2).$$

Ejemplo (distribución normal contaminada):

Considere $\boldsymbol{X} \sim \mathsf{CN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$ (Little, 1988) donde $0 \leq \epsilon \leq 1$ denota el porcentaje de contaminación y $0 < \gamma < 1$ corresponde a un factor de inflación de escala. En este caso,

$$h(\omega; \boldsymbol{\delta}) = \begin{cases} \epsilon, & \omega = \gamma \\ 1 - \epsilon & \omega = 1 \end{cases},$$

con $\boldsymbol{\delta} = (\epsilon, \gamma)^\top$.

Referencias bibliográficas

Andrews, D.F., Mallows, C.L. (1974).
Scale mixtures of normal distributions
Journal of the Royal Statistical Society, Series B **36**, 99-102.

Fang, K.T., Kotz, S., Ng, K.W. (1990).
Symmetric Multivariate and Related Distributions.
Chapman & Hall, London.

Gómez, E., Gómez-Villegas, M.A., Marín, J.M. (1988).
A multivariate generalization of the power exponential family of distributions.
Communication in Statistics: Theory and Methods **27**, 589-600.

Lange, K., Sinsheimer, J.S. (1993).
Normal/independent distributions and their applications in robust regression.
Journal of Computational and Graphical Statistics **2**, 175-198.

Little, R.J.A. (1988).
Robust estimation of the mean and covariance matrix from data with missing values.
Applied Statistics **37**, 23-38.