MAT-206: Sesión 4, Familia de contornos elípticos

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

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

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.

Observación:

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

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

Familia de posición-escala

Ejemplo:

Algunas familias de posición-escala corresponden a:

- Normal, $\mathsf{N}(a, b^2)$:

$$f(y; a, b) = \frac{1}{b}(2\pi)^{-1/2} \exp\Big\{ -\frac{1}{2}\Big(\frac{y-a}{b}\Big)^2 \Big\}.$$

- Laplace (doble exponencial), $\mathsf{Laplace}(a, b)$:

$$f(y; a, b) = \frac{1}{2b} \exp\Big\{ -\frac{|y-a|}{b} \Big\}.$$

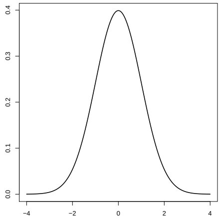
- Cauchy, $\mathsf{Cauchy}(a, b^2)$:

$$f(y; a, b) = \frac{b}{\pi} \frac{1}{b^2 + (y-a)^2}.$$

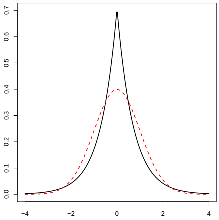
- Logística, $\mathsf{Logística}(a, b)$:

$$f(y; a, b) = \frac{1}{b} \frac{e^{-(y-a)/b}}{(1 + e^{-(y-a)/b})^2}.$$

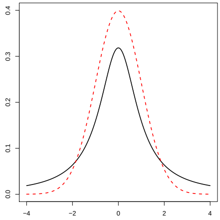
Familia de posición-escala



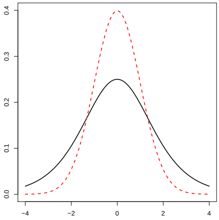
(a) Normal



(b) Laplace



(c) Cauchy



(d) Logística

Definición 2:

Un vector aleatorio p-dimensional, $\boldsymbol{X}$ tiene distribución normal con vector de medias $\boldsymbol{\mu} \in \mathbb{R}^p$ y matriz de covarianza $\operatorname{Cov}(\boldsymbol{X}) = \boldsymbol{\Sigma} \geq \boldsymbol{0}$ si y sólo si, para todo vector $\boldsymbol{t}$ la variable aleatoria (uni-dimensional) $\boldsymbol{t}^\top \boldsymbol{X}$ es normal y escribimos $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Observación:

Note que en la definición anterior no se ha hecho supuestos respecto de la independencia de los componentes de $\boldsymbol{X}$.

Definición 3:

La función característica de $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ está dada por

$$\varphi_X(\boldsymbol{t}) = \exp(i\boldsymbol{t}^\top\boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Sigma}\boldsymbol{t}).$$

En efecto, notando que $Y = \boldsymbol{t}^\top\boldsymbol{X}$ tiene media y varianza dadas por,

$$\lambda = \mathsf{E}(Y) = \boldsymbol{t}^\top\boldsymbol{\mu}, \qquad \sigma^2 = \mathsf{var}(Y) = \boldsymbol{t}^\top\boldsymbol{\Sigma}\boldsymbol{t} \geq 0,$$

sigue que

$$\varphi_Y(\boldsymbol{t}) = \varphi_{t^\top X}(1) = \exp(i\lambda - \tfrac{1}{2}\sigma^2) = \exp(i\boldsymbol{t}^\top\boldsymbol{\mu} - \tfrac{1}{2}\boldsymbol{t}^\top\boldsymbol{\Sigma}\boldsymbol{t})$$

Como un caso particular tenemos que la función característica para $\boldsymbol{Z} \sim \mathsf{N}_p(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_p)$, es

$$\varphi_Z(\boldsymbol{t}) = \exp(-\tfrac{1}{2}\sigma^2 \boldsymbol{t}^\top \boldsymbol{t}) = \prod_{i=1}^p \exp(-\tfrac{1}{2}\sigma^2 t_i^2) = \prod_{i=1}^p \varphi_{Z_i}(t_i)$$

y de este modo, obtenemos

$$\boldsymbol{Z} \sim \mathsf{N}_p(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_p) \quad \iff \quad Z_1, \dots, Z_p \text{ IID } \mathsf{N}(0, \sigma^2).$$

Resultado 1:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\boldsymbol{\Sigma}$ es definida positiva, entonces la densidad de $\boldsymbol{X}$ es

$$f(\boldsymbol{x}) = |2\pi \boldsymbol{\Sigma}|^{-1/2} \exp\big\{ -\tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\big\}.$$

Demostración:

Sea $Z_1, \dots, Z_p$ variables aleatorias IID $\mathsf{N}(0,1)$. Entonces la densidad conjunta de $\boldsymbol{Z} = (Z_1, \dots, Z_p)^\top$ es

$$f(\boldsymbol{z}) = \prod_{i=1}^{p} (2\pi)^{-1/2} \exp(-z_i^2/2) = (2\pi)^{-p/2} \exp(-\tfrac{1}{2}\|\boldsymbol{z}\|^2).$$

Considere $\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{B}\boldsymbol{Z}$ con $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$, con $\boldsymbol{B}$ matriz de rango completo. Entonces, tenemos la transformación inversa

$$\boldsymbol{Z} = \boldsymbol{g}^{-1}(\boldsymbol{X}) = \boldsymbol{B}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}),$$

y $\mathrm{d}\,\boldsymbol{Z} = \mathrm{d}\,\boldsymbol{g}^{-1}(\boldsymbol{X}) = \boldsymbol{B}^{-1}\,\mathrm{d}\,\boldsymbol{X}$, con matriz jacobiana $\mathrm{D}\,\boldsymbol{g}^{-1}(\boldsymbol{X}) = \boldsymbol{B}^{-1}$, como

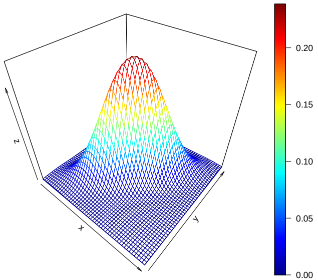
$$|\,\mathrm{D}\,\boldsymbol{g}^{-1}(\boldsymbol{X})|_+ = |\boldsymbol{B}|^{-1} = |\boldsymbol{B}\boldsymbol{B}^\top|^{-1/2},$$

obtenemos

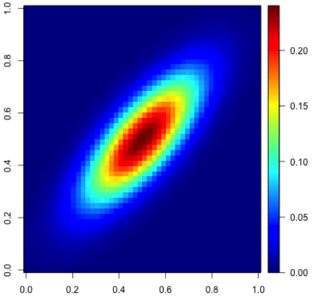
$$\begin{aligned} f(\boldsymbol{x}) &= |\,\mathrm{D}\,\boldsymbol{g}^{-1}(\boldsymbol{x})|_+ f_{\boldsymbol{Z}}(\boldsymbol{g}^{-1}(\boldsymbol{x})) \\ &= (2\pi)^{-p/2}|\boldsymbol{B}\boldsymbol{B}^\top|^{-1/2} \exp\{-\tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{B}^{-\top}\boldsymbol{B}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\}, \end{aligned}$$

notando que $\boldsymbol{\Sigma}^{-1} = \boldsymbol{B}^{-\top}\boldsymbol{B}^{-1}$ sigue el resultado deseado.

Distribución normal multivariada



(a) densidad



(b) contornos

Definición 4:

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

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

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

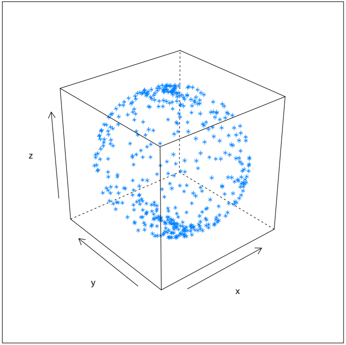
Resultado 2:

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:

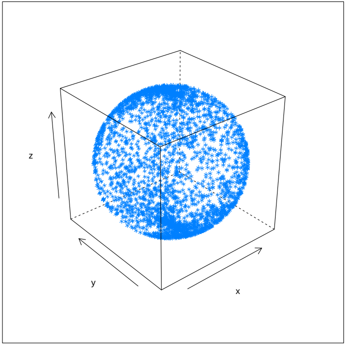
$$\boldsymbol{U} = \frac{\boldsymbol{Z}}{\|\boldsymbol{Z}\|},$$

tiene distribución uniforme sobre la esfera unitaria.

Distribuciones de contornos elípticos



(a) $n = 750$



(b) $n = 3000$

Definition 5:

Un vector aleatorio p-dimensional tiene distribución esférica si y 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 $\phi(\cdot)$ de una variable escalar tal que $\varphi(\boldsymbol{t}) = \phi(\boldsymbol{t}^\top \boldsymbol{t})$.

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

Ejemplo:

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

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

Resultado 3

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

$$\boldsymbol{X} \stackrel{\mathsf{d}}{=} R\boldsymbol{U}, \tag{1}$$

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

Resultado 4

Suponga que $\boldsymbol{X} \stackrel{\mathsf{d}}{=} R\boldsymbol{U} \sim \mathsf{S}_p(\phi)$ $(\mathsf{P}(\boldsymbol{X} = \mathbf{0}) = 0)$, 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 5

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

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

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

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

el resultado sigue.

Definición 6:

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

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

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

Observación:

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

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

Note además que la representación estocástica de $\boldsymbol{X}$ es dada por

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

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

Definición 7:

Se dice que el vector $\boldsymbol{X}$ tiene distribución de contornos elípticos si su función de densidad es de la forma

$$f(\boldsymbol{x}) = |\boldsymbol{\Sigma}|^{-1/2} g((\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})), \qquad \boldsymbol{x} \in \mathbb{R}^p,$$

donde $g : \mathbb{R} \to [0, \infty)$ es función decreciente, llamada función generadora de densidad, tal que:

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

En cuyo caso escribimos $\boldsymbol{X} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$.

Ejemplo:

La función generadora de densidad de un vector aleatorio con distribución *t* multivariada asume la forma

$$g(u) = \frac{\Gamma(\frac{\nu+p}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{p/2}}\Big(1 + \frac{u}{\nu}\Big)^{-(\nu+p)/2}, \qquad \nu > 0.$$

En cuyo caso escribimos, $\boldsymbol{X} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$.

En este caso, tenemos que $R^2/p \sim F_{p,\nu}$. Además, su función característica es dada por

$$\varphi(\boldsymbol{t}) = \frac{\|\sqrt{\nu}\boldsymbol{\Sigma}^{1/2}\boldsymbol{t}\|^{\nu/2}}{2^{\nu/2-1}\Gamma(\nu/2)}\ \exp(i\boldsymbol{t}^\top\boldsymbol{\mu})K_{\nu/2}(\|\sqrt{\nu}\boldsymbol{\Sigma}^{1/2}\boldsymbol{t}\|), \qquad \boldsymbol{t} \in \mathbb{R}^p,$$

donde $K_\nu(x)$ denota la función de Bessel modificada de segundo tipo.

Ejemplo:

Para la distribución Exponencial Potencia, la función generadora de densidades es dada por

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

y es usual utilizar la notación $\boldsymbol{X} \sim \mathsf{PE}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$. En este caso tenemos que la variable aleatoria positiva R tiene densidad

$$h(r) = \frac{p}{\Gamma(1+\frac{p}{2\lambda})2^{\frac{p}{2\lambda}}}\, r^{p-1}\exp(-r^{2\lambda}/2), \qquad r > 0.$$

Note también que $R^{2\lambda} \sim \mathsf{Gama}(\frac{1}{2}, \frac{p}{2\lambda})$.

Definición 8:

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:

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

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

Observación:

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

$$\boldsymbol{X} \stackrel{\mathsf{d}}{=} \boldsymbol{\mu} + W^{-1/2}\boldsymbol{Z}, \qquad (2)$$

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

Ejemplo:

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:

Se dice que un vector aleatorio $\boldsymbol{X}$ tiene distribución Exponencial-Potencia[1], si su función de densidad es dada por:

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

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

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

Distribuciones de mezcla de escala normal

Observación:

La representación estocástica en (2), puede ser escrita de forma equivalente, como:

$$\boldsymbol{X}|W \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/\omega), \qquad W \sim \mathsf{H}(\boldsymbol{\delta}). \tag{3}$$

Esta representación permite, por ejemplo

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

Además, la formulación condicional en (3) es muy útil para:

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

Ejemplo:

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

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

es decir,

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

Ejemplo:

Considere $\boldsymbol{X} \sim \mathsf{CN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$ 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 $\boldsymbol{\delta} = (\epsilon, \gamma)^\top$.