Normal Multivariada

Graciela Boente

Definición 1

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
  Se dice que $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mu)\right\}$$

Definición 1

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p\times p}$ simétrica y definida positiva
  Se dice que $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mu) \right\}$$

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathrm{diag}(\lambda_1, \ldots, \lambda_p))$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{\prod_{j=1}^{p} \lambda_j^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} \sum_{j=1}^{p} \frac{x_j^2}{\lambda_j} \right\}$$

Por lo tanto, $x_1, \ldots, x_p$ son independientes $x_j \sim N(0, \lambda_j)$.

Definición 1

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
  Se dice que $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mu)\right\}$$

- Si $\mathbf{x} \sim N(\mathbf{0}, \operatorname{diag}(\lambda_1, \ldots, \lambda_p))$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{\prod_{j=1}^{p} \lambda_j^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} \sum_{j=1}^{p} \frac{x_j^2}{\lambda_j}\right\}$$

  Por lo tanto, $x_1, \ldots, x_p$ son independientes $x_j \sim N(0, \lambda_j)$.
- En particular, si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p)$, $x_1, \ldots, x_p$ son i.i.d. $N(0, 1)$.

Definición 1

- Sea $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ simétrica y definida positiva
  Se dice que $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ si su densidad está dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mu)\right\}$$

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathrm{diag}(\lambda_1, \ldots, \lambda_p))$

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \frac{1}{\prod_{j=1}^{p} \lambda_j^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} \sum_{j=1}^{p} \frac{x_j^2}{\lambda_j}\right\}$$

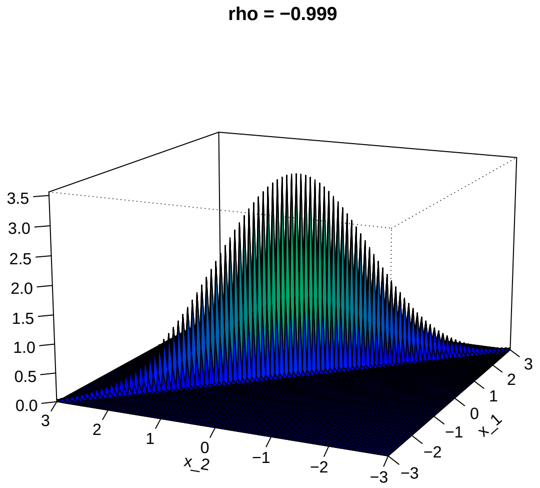
  Por lo tanto, $x_1, \ldots, x_p$ son independientes $x_j \sim N(0, \lambda_j)$.
- En particular, si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p)$, $x_1, \ldots, x_p$ son i.i.d. $N(0, 1)$.
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\mathbf{A} \in \mathbb{R}^{p \times p}$ es no singular $\Longrightarrow$
  $\mathbf{A}\mathbf{x} + \mathbf{b} \sim N(\mathbf{A}\mu + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}})$

Caso p=2

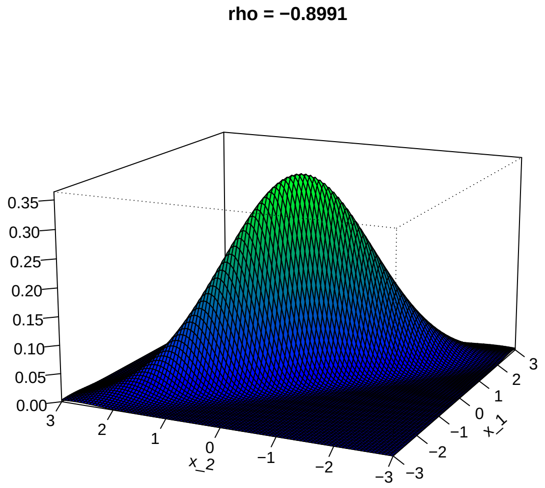
- Sea $\boldsymbol{\mu} \in \mathbb{R}^2$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$ definida positiva $(|\rho| \neq 1)$

$$f(\mathbf{x}) = \frac{1}{2\pi}\frac{1}{\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}$$

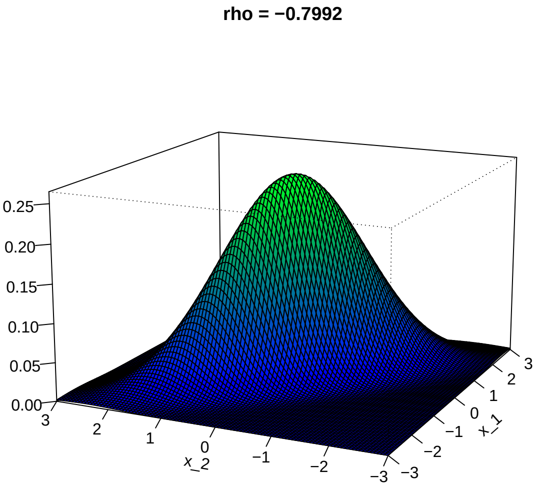
Caso p=2



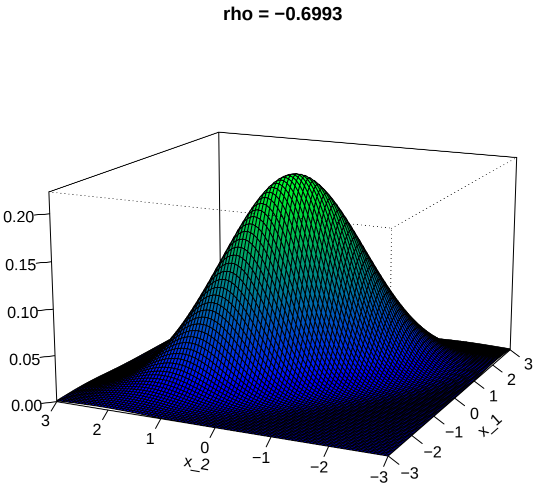
Caso p=2


rho = −0.8991
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
3
2
1
0
−1
−2
−3
x_2
3
2
1
0
−1
−2
−3
x_1

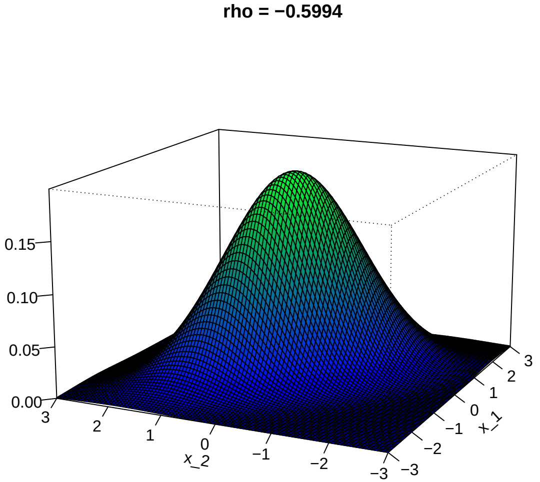
Caso p=2



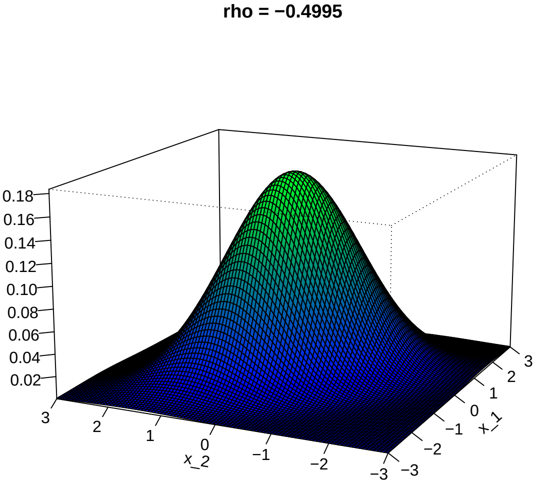
Caso p=2



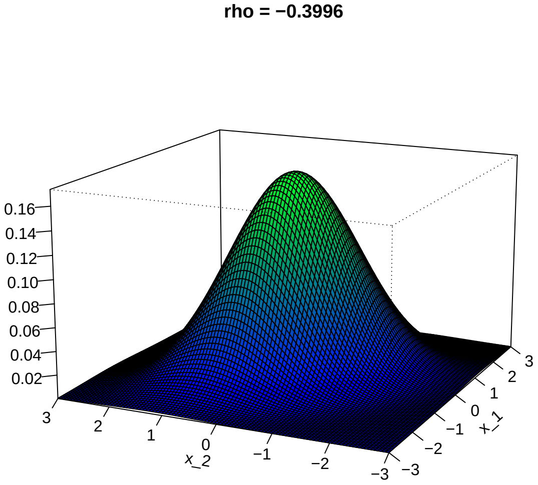
Caso p=2



Caso p=2

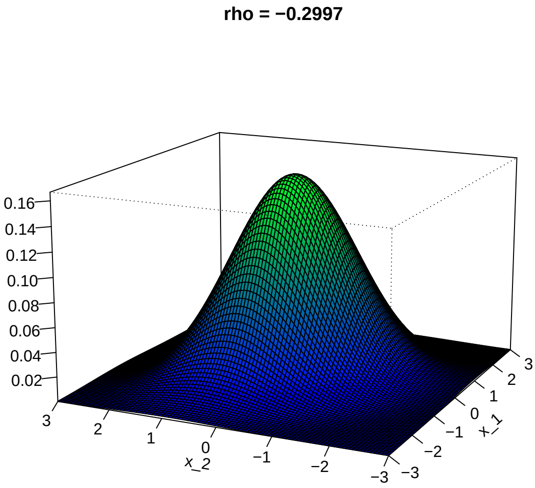

rho = −0.4995
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
−1
−2
−3
x_2
3
2
1
0
−1
−2
−3
x_1

Caso p=2

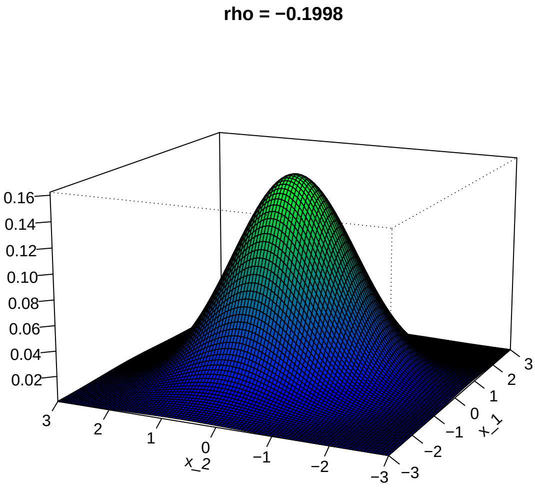


rho = −0.3996
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
x_2
−1
−2
−3
3
2
1
0
x_1
−1
−2
−3

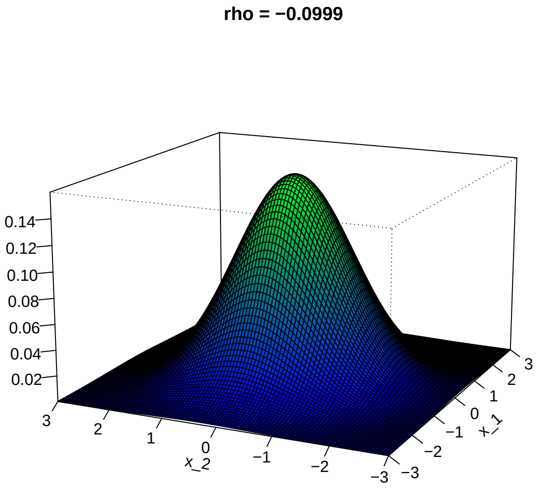
Caso p=2


rho = −0.2997
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
x_2
−1
−2
−3
−3
−2
−1
0
1
2
3
x_1

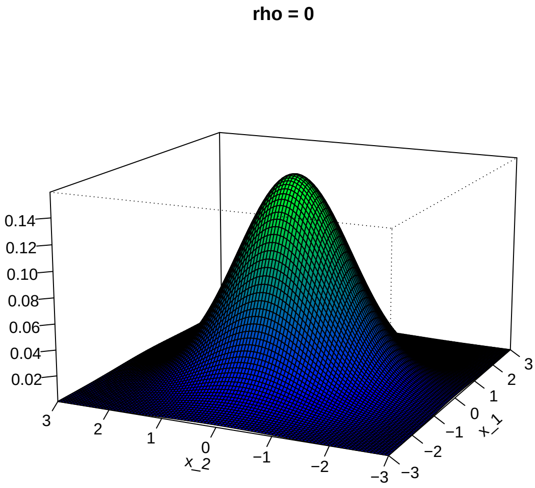
Caso p=2


rho = −0.1998
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
−1
−2
−3
x_2
3
2
1
0
−1
−2
−3
x_1

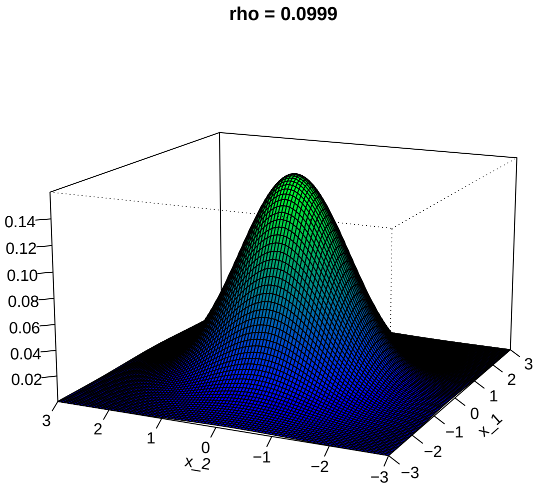
Caso p=2



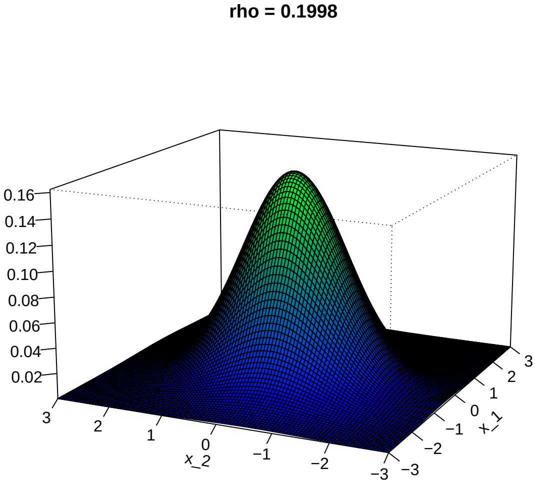
Caso p=2



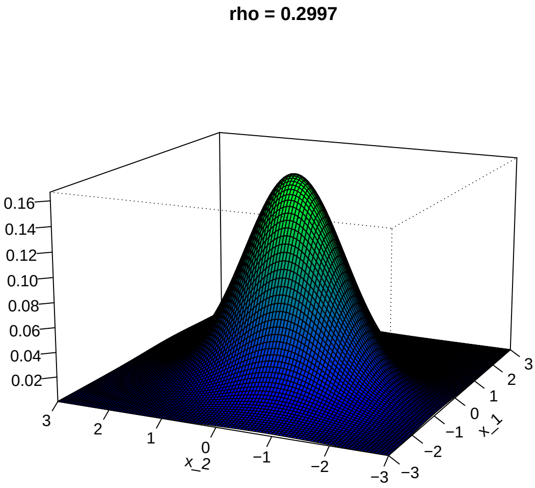
Caso p=2



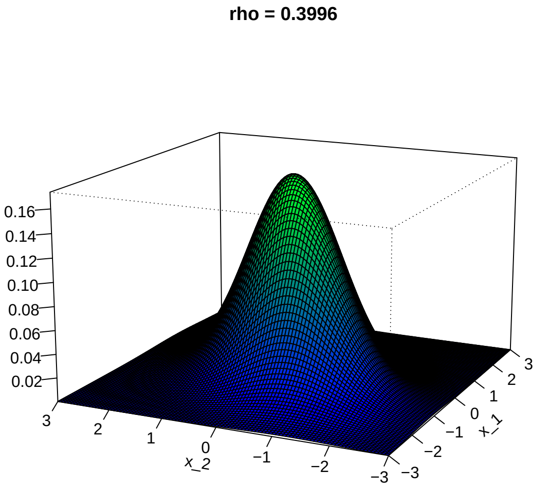
Caso p=2


rho = 0.1998
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
x_2
−1
−2
−3
3
2
1
0
x_1
−1
−2
−3

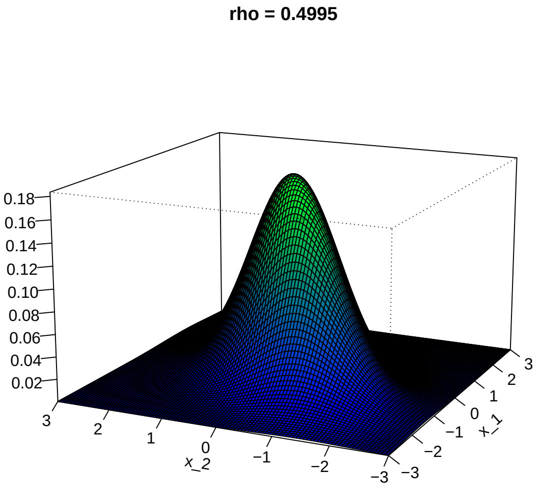
Caso p=2


rho = 0.2997
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
x_2
−1
−2
−3
3
2
1
0
x_1
−1
−2
−3

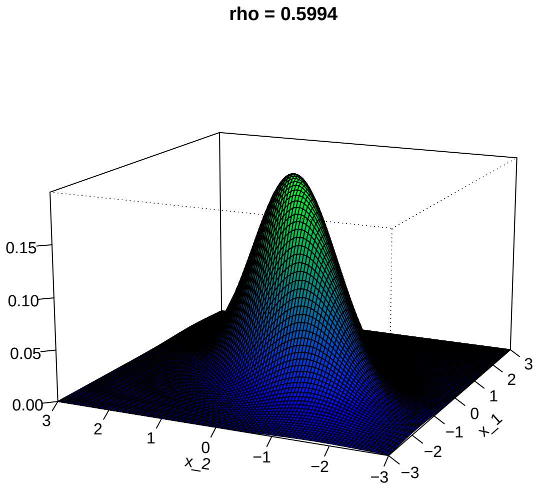
Caso p=2


rho = 0.3996
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
x_2
3
2
1
0
−1
−2
−3
x_1
3
2
1
0
−1
−2
−3

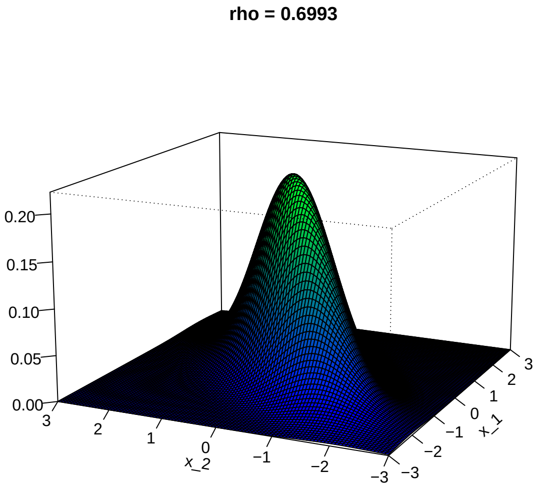
Caso p=2


rho = 0.4995
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
3
2
1
0
−1
−2
−3
x_2
3
2
1
0
−1
−2
−3
x_1

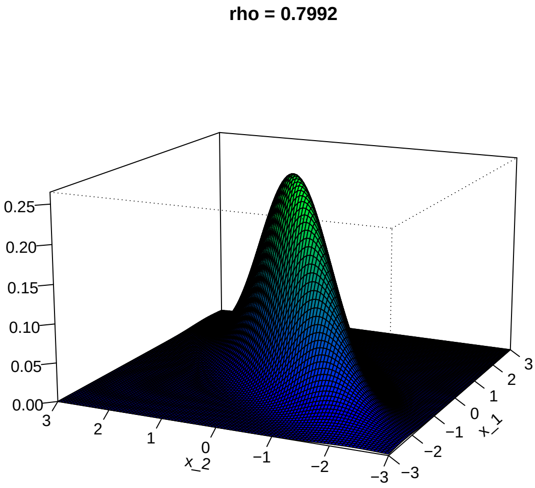
Caso p=2



Caso p=2

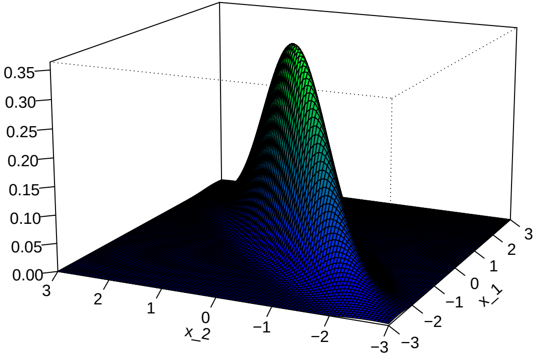


Caso p=2

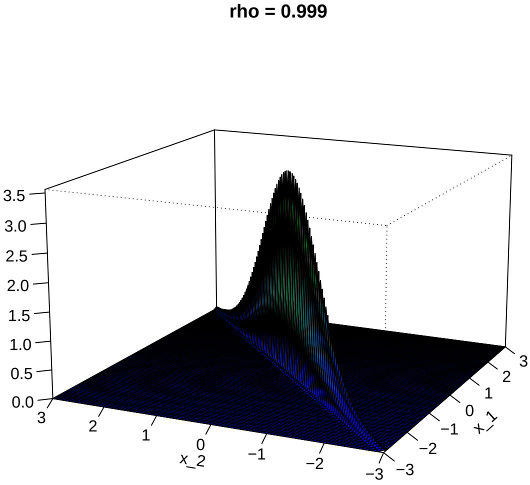


Caso p=2

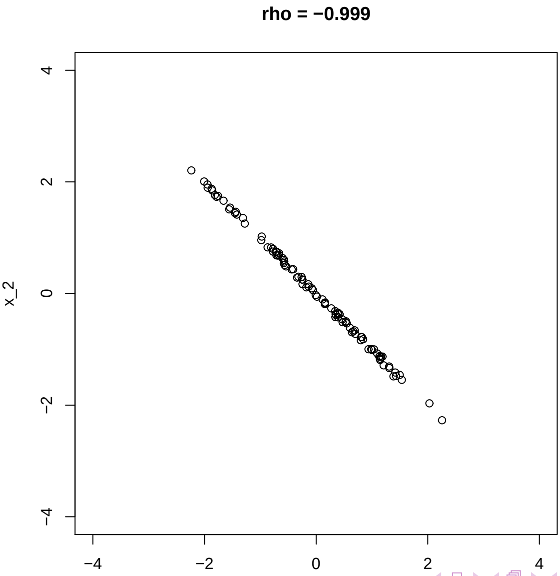
rho = 0.8991



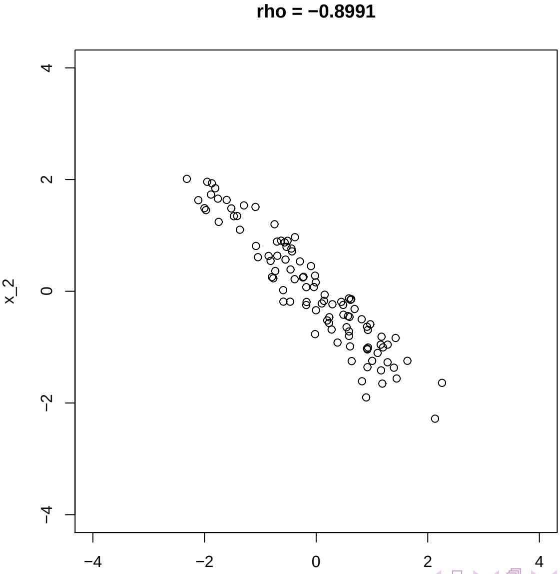
Caso p=2



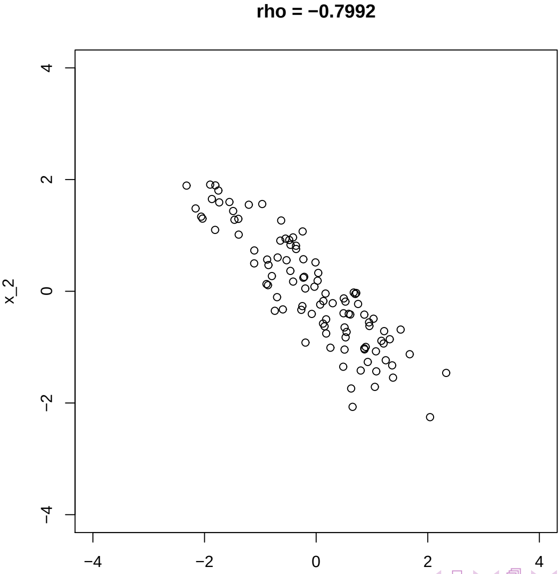
Caso p=2



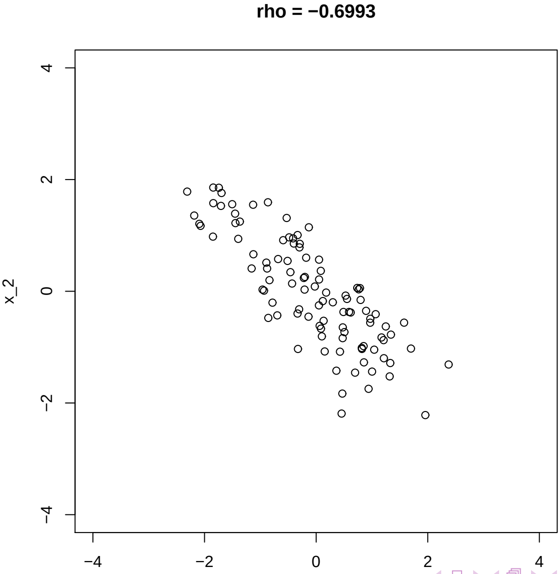
Caso p=2



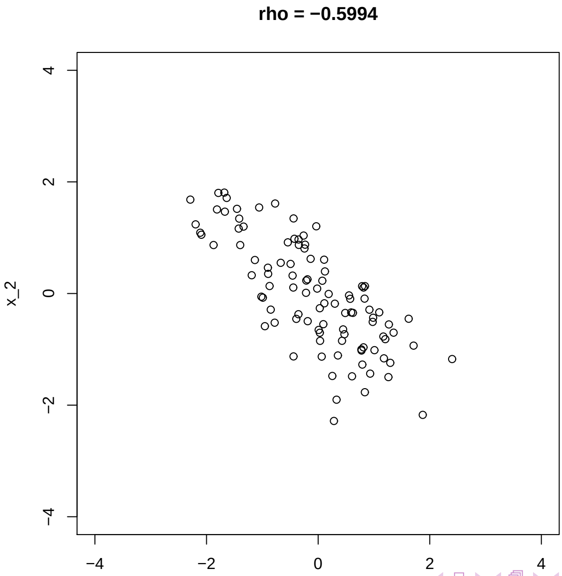
Caso p=2



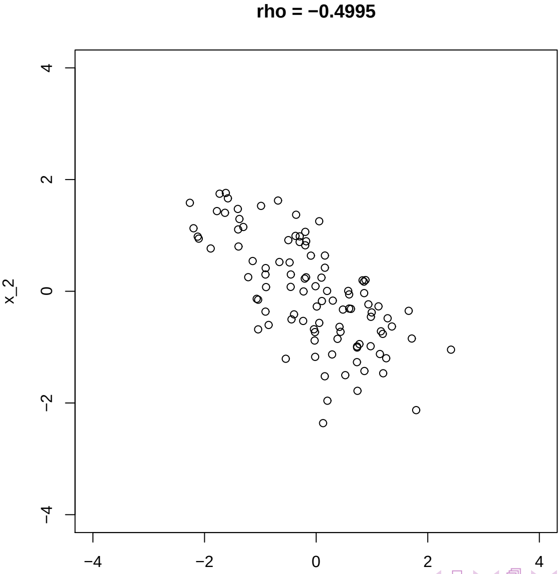
Caso p=2



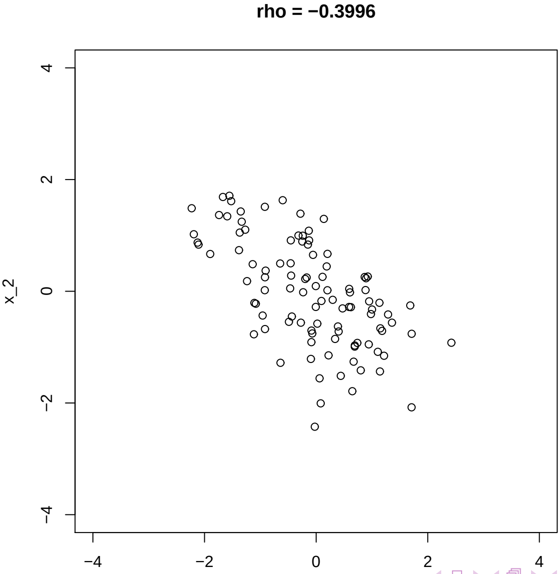
Caso p=2



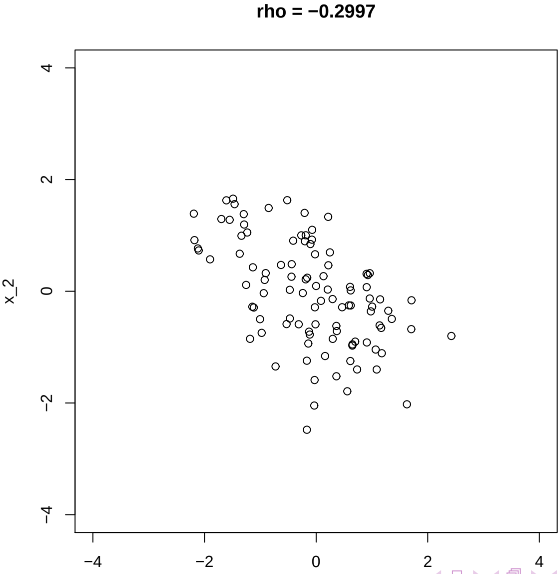
Caso p=2



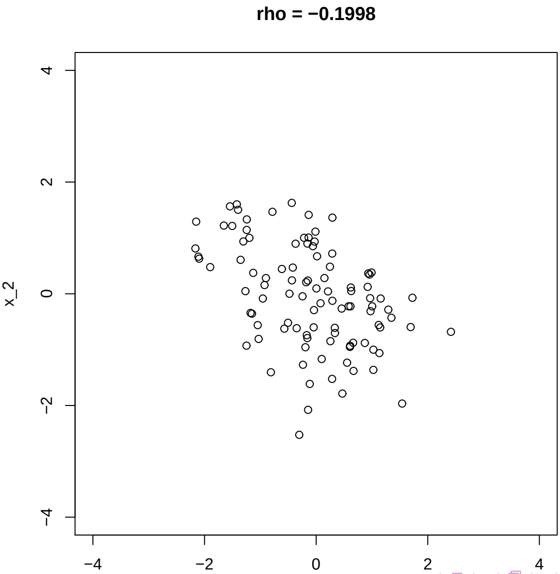
Caso p=2



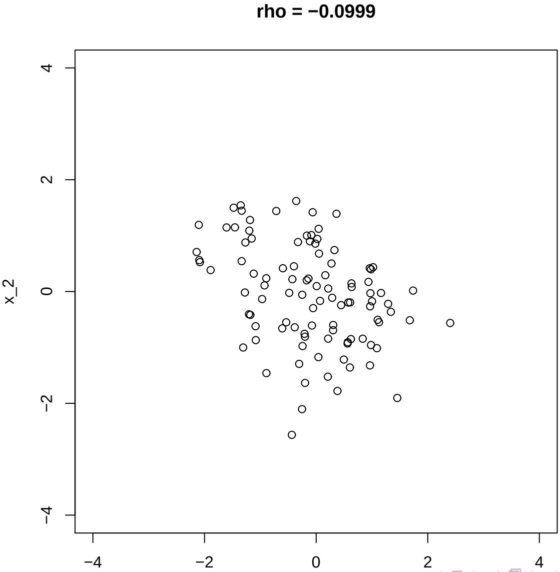
Caso p=2



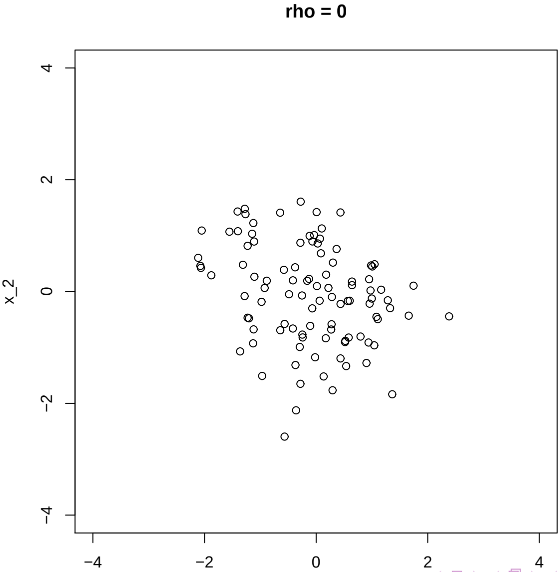
Caso p=2



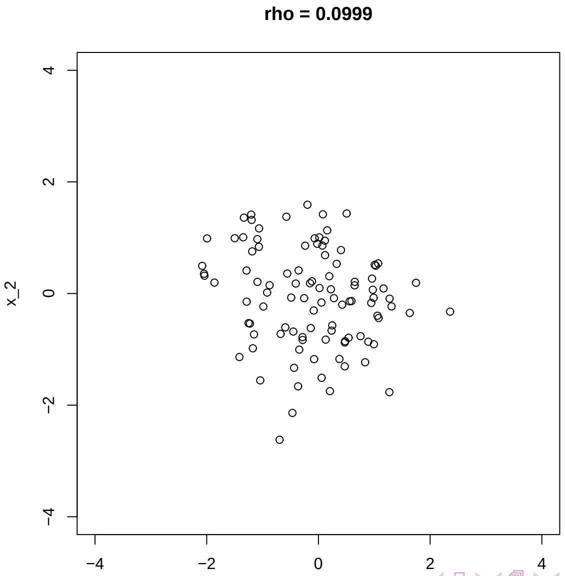
Caso p=2



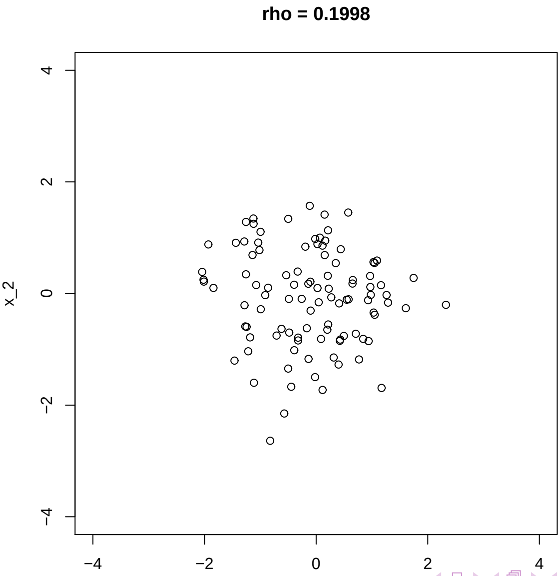
Caso p=2



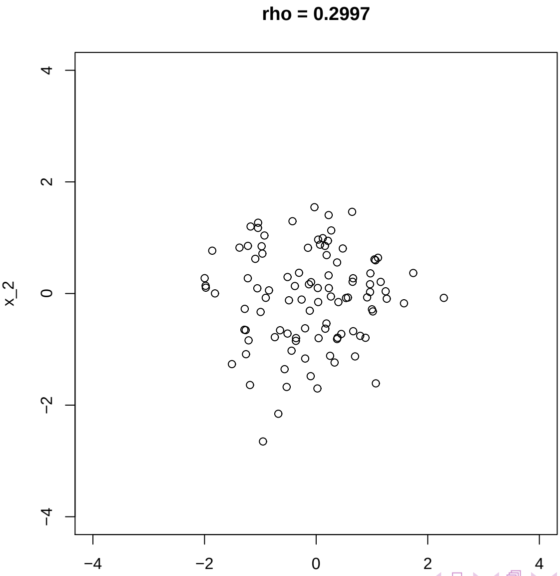
Caso p=2



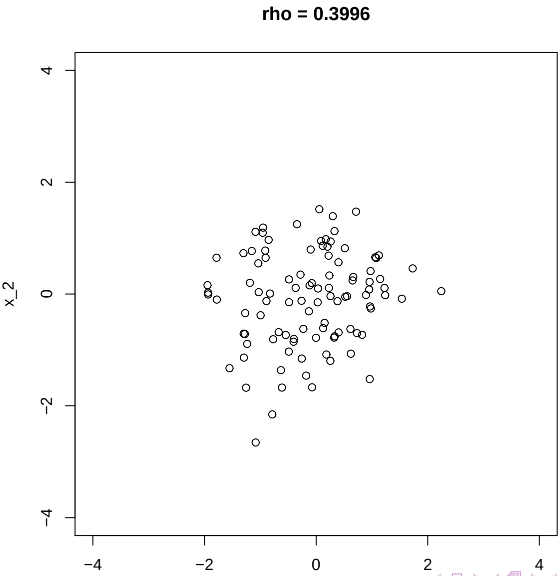
Caso p=2



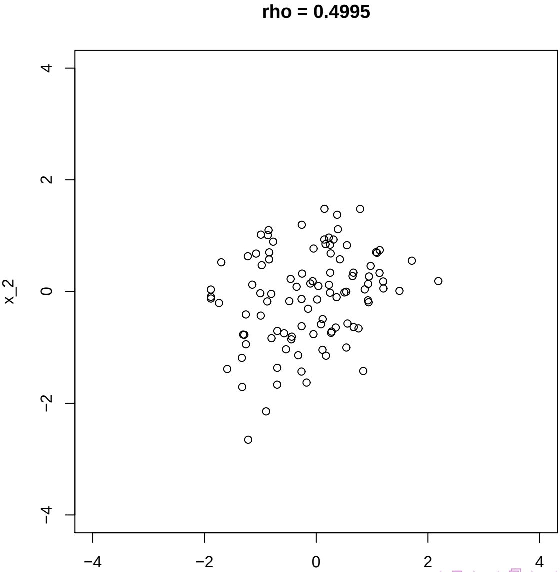
Caso p=2



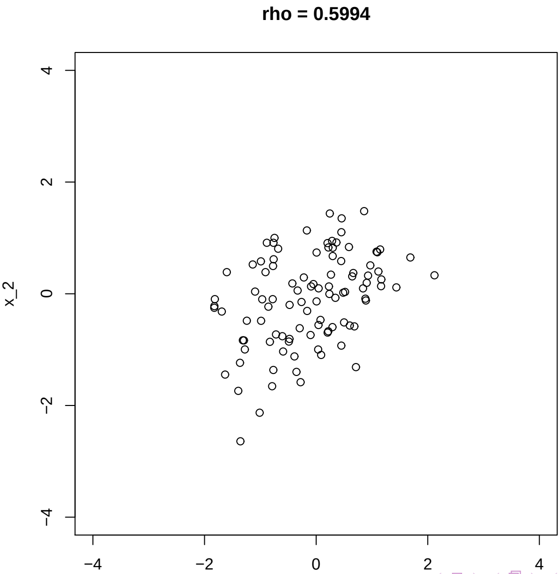
Caso p=2



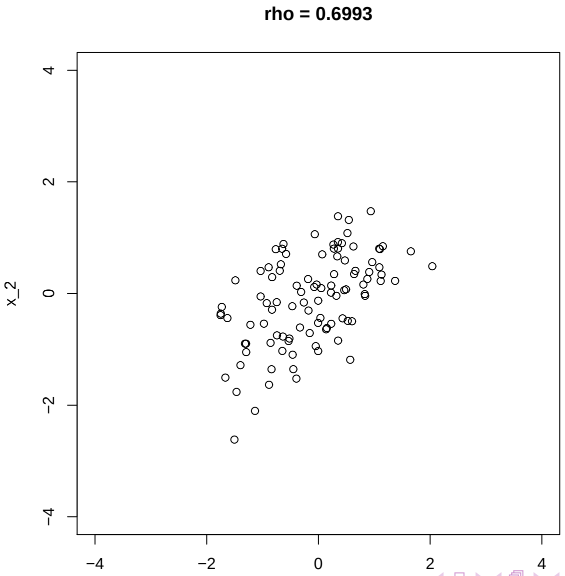
Caso p=2



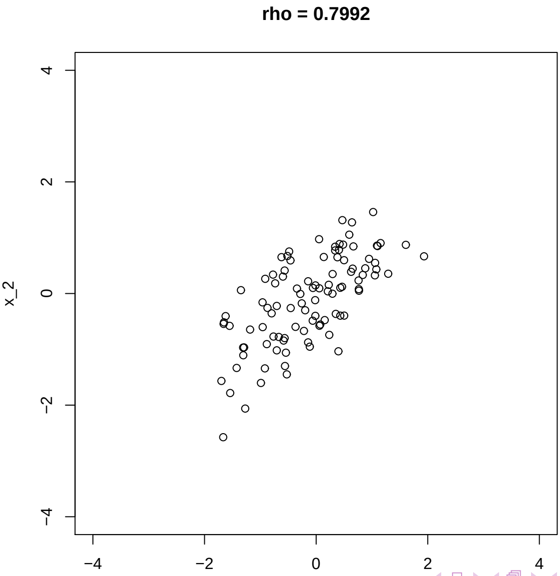
Caso p=2



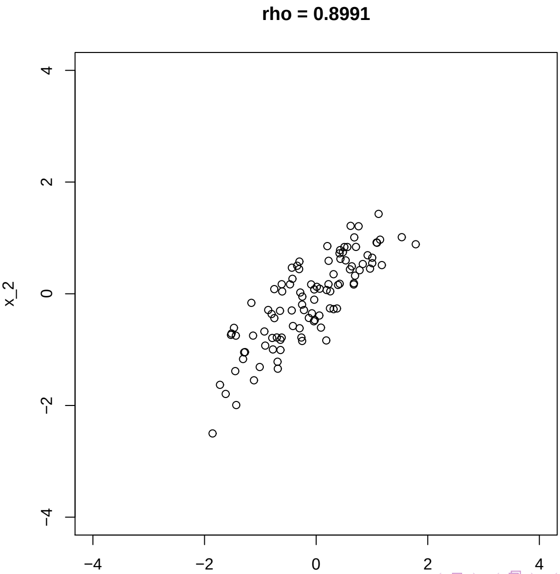
Caso p=2



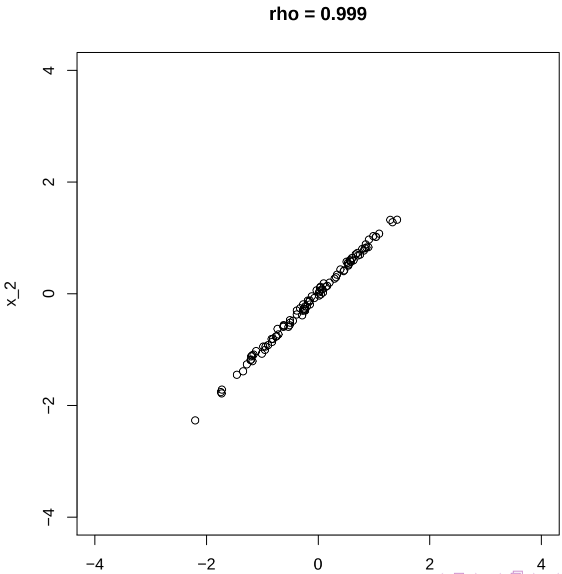
Caso p=2



Caso p=2



Caso p=2



Definición 2

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$

Definición 2

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$

Definición 2

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\}\exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$

Definición 2

Se dice que $\mathbf{x}$ es normal multivariada si y sólo si $\forall\, \mathbf{t} \in \mathbb{R}^p$ se tiene que $\mathbf{t}^{\mathrm{T}}\mathbf{x}$ es normal univariada.

Propiedades

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$

Propiedades

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$

Propiedades

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\}\exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$
- $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longleftrightarrow \mathbf{a}^{\mathrm{T}}\mathbf{x} \sim N(\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu}, \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{a}),\ \forall \mathbf{a} \in \mathbb{R}^p$

Propiedades

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\} \exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$
- $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longleftrightarrow \mathbf{a}^{\mathrm{T}}\mathbf{x} \sim N(\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu}, \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{a}), \ \forall \mathbf{a} \in \mathbb{R}^p$
- $\mathbf{x} = (x_1, \ldots, x_p)^{\mathrm{T}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \mathbb{E}(\mathbf{x}) = \boldsymbol{\mu}$ y $\mathrm{COV}(\mathbf{x}) = (\mathrm{COV}(x_i, x_j))_{1 \leq i,j \leq p} = \boldsymbol{\Sigma}$

Propiedades

- Si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p) \implies \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{-\frac{1}{2}\|\mathbf{t}\|^2\}$
- Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \varphi_{\mathbf{x}}(\mathbf{t}) = \exp\{i\mathbf{t}^{\mathrm{T}}\boldsymbol{\mu}\}\exp\{-\frac{1}{2}\mathbf{t}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{t}\}$
- $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \iff \mathbf{a}^{\mathrm{T}}\mathbf{x} \sim N(\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu}, \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{a})$, $\forall \mathbf{a} \in \mathbb{R}^p$
- $\mathbf{x} = (x_1, \ldots, x_p)^{\mathrm{T}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \implies \mathbb{E}(\mathbf{x}) = \boldsymbol{\mu}$ y $\mathrm{COV}(\mathbf{x}) = (\mathrm{COV}(x_i, x_j))_{1 \leq i,j \leq p} = \boldsymbol{\Sigma}$
- Sea $\mathbf{x} = (x_1, \ldots, x_p)^{\mathrm{T}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces, $x_1, \ldots, x_p$ son independientes $\iff \boldsymbol{\Sigma}$ es diagonal.

Propiedades

Vimos que si $\mathbf{x} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, entonces

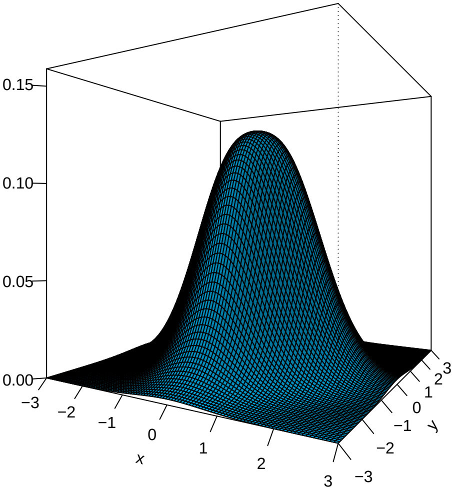
- $x_1, \ldots, x_p$ son independientes $\Longleftrightarrow$ $\boldsymbol{\Sigma}$ es diagonal.
- $x_j \sim N(0, \sigma_{jj})$ donde $\boldsymbol{\Sigma} = (\sigma_{ij})_{1 \leq i,j \leq p}$

El siguiente ejemplo muestra que existen vectores aleatorios $\mathbf{x} = (x_1, x_2)$ tales que

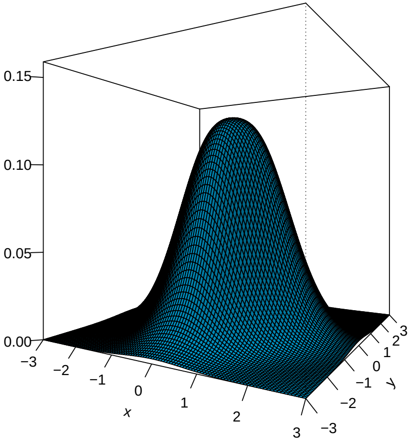
- $\mathrm{COV}(\mathbf{x}) = \mathbf{I}_p$.
- $x_j \sim N(0, 1)$
- x_1 y x_2 no son independientes.

Sea **x** con densidad

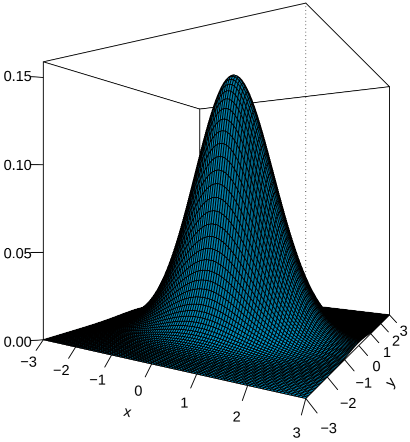
$$h(x,y) = \frac{1}{2\pi}\left\{\left(\sqrt{2}\,e^{-\frac{x^2}{2}} - e^{-x^2}\right)e^{-y^2} + \left(\sqrt{2}\,e^{-\frac{y^2}{2}} - e^{-y^2}\right)e^{-x^2}\right\}$$



Densidad h



Densidad $N(\mathbf{0}, \mathbf{I}_2)$



Propiedades

Sea $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} > 0$. Definamos $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix}$,
$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$ con $\mathbf{x}^{(i)}, \boldsymbol{\mu}^{(i)} \in \mathbb{R}^{p_i}$,
$\boldsymbol{\Sigma}_{ii} \in \mathbb{R}^{p_i \times p_i}$, $p_1 + p_2 = p$.

Entonces,

a) $\mathbf{x}^{(1)} \sim N(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}_{11})$ y $\mathbf{x}^{(2)} \sim N(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$.

b) Más aún, $\mathbf{x}^{(1)}$ y $\mathbf{x}^{(2)}$ son independientes $\iff \boldsymbol{\Sigma}_{21} = 0$.

c) Dada $\mathbf{A} \in \mathbb{R}^{q \times p}$, $rg(\mathbf{A}) = q \implies \mathbf{A}\mathbf{x} \sim N(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathrm{T}})$

En particular, si $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_p)$ y $\mathbf{H} = (\mathbf{h}_1, \ldots, \mathbf{h}_q) \in \mathbb{R}^{p \times q}$, es ortogonal incompleta, o sea, $\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{I}_q$, entonces
$\mathbf{y} = \mathbf{H}^{\mathrm{T}}\mathbf{x} \sim N(\mathbf{0}, \mathbf{I}_q)$.

Propiedades

d) Sea $\mathbf{\Sigma} = \mathbf{H}\mathbf{\Lambda}\mathbf{H}^{\mathrm{T}}$, con $\mathbf{H}$ ortogonal y $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\lambda_1 \geq \ldots, \lambda_p$.
Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma}) \Longrightarrow \mathbf{H}^{\mathrm{T}}(\mathbf{x} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \mathbf{\Lambda})$.

Propiedades

d) Sea $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^{\mathrm{T}}$, con $\mathbf{H}$ ortogonal y $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\lambda_1 \geq \ldots, \lambda_p$.
Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \mathbf{H}^{\mathrm{T}}(\mathbf{x} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Lambda})$.

e) Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longleftrightarrow \mathbf{x} = \mathbf{A}\mathbf{z} + \boldsymbol{\mu}$ con $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_p)$ y $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \boldsymbol{\Sigma}$.

Propiedades

d) Sea $\boldsymbol{\Sigma} = \mathbf{H}\boldsymbol{\Lambda}\mathbf{H}^{\mathrm{T}}$, con $\mathbf{H}$ ortogonal y $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, $\lambda_1 \geq \ldots, \lambda_p$.
Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \mathbf{H}^{\mathrm{T}}(\mathbf{x} - \boldsymbol{\mu}) \sim N(\mathbf{0}, \boldsymbol{\Lambda})$.

e) Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longleftrightarrow \mathbf{x} = \mathbf{A}\mathbf{z} + \boldsymbol{\mu}$ con $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I}_p)$ y $\mathbf{A}\mathbf{A}^{\mathrm{T}} = \boldsymbol{\Sigma}$.

f) Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow (\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \mu) \sim \chi^2_p$

g) Si $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \Longrightarrow \mathbf{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{x} \sim \chi^2_p(\delta^2)$ con $\delta^2 = \boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$

Propiedades

Sea $\mathbf{x} \sim N(\mathbf{0}, \sigma^2 \mathbf{I_p})$

y sea $\mathbf{H}_1 = (\mathbf{h}_1, \ldots, \mathbf{h}_q) \in \mathbb{R}^{p\times q}$ ortogonal incompleta, o sea, $\mathbf{H}_1^{\mathrm{T}} \mathbf{H}_1 = \mathbf{I}_q$.

Sea $\mathbf{H} = (\mathbf{H}_1, \mathbf{H}_2)$ ortogonal, o sea, $\mathbf{H}^{\mathrm{T}}\mathbf{H} = \mathbf{H}\mathbf{H}^{\mathrm{T}} = \mathbf{I}_p$

Entonces

a) $\mathbf{z} = \mathbf{H}_1^{\mathrm{T}}\mathbf{x} \sim N(\mathbf{0}, \sigma^2 \mathbf{I_q})$

b) $\mathbf{z}$ es independiente de $\mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{z}^{\mathrm{T}}\mathbf{z}$

c) $\dfrac{\mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{z}^{\mathrm{T}}\mathbf{z}}{\sigma^2} \sim \chi^2_{p-q}$

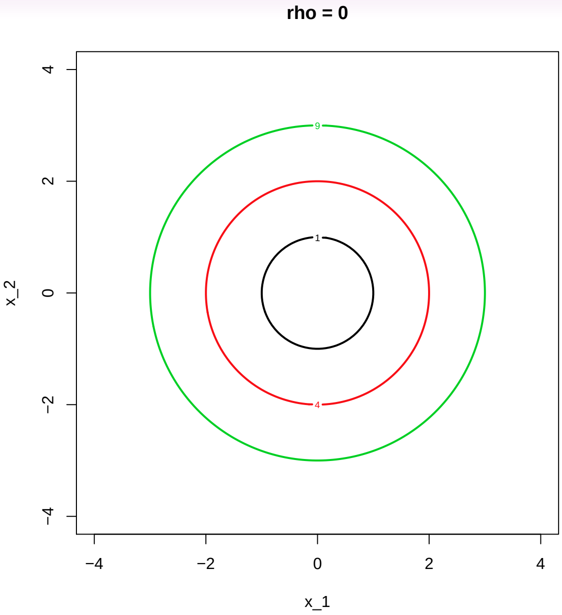
Propiedades

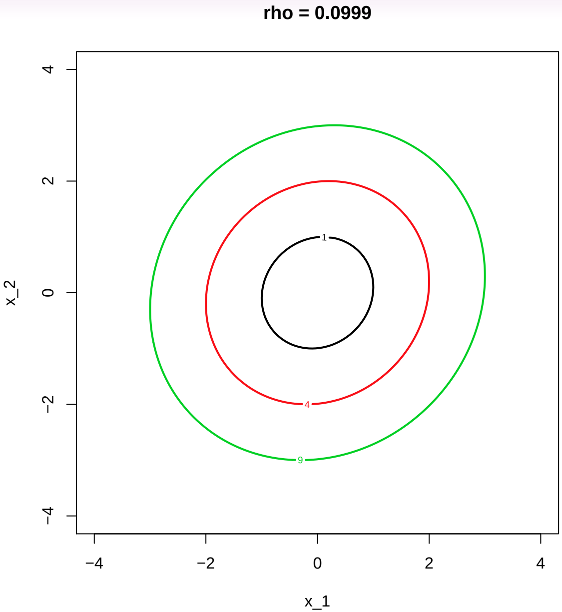
- Las regiones de densidad constante son los elipsoides

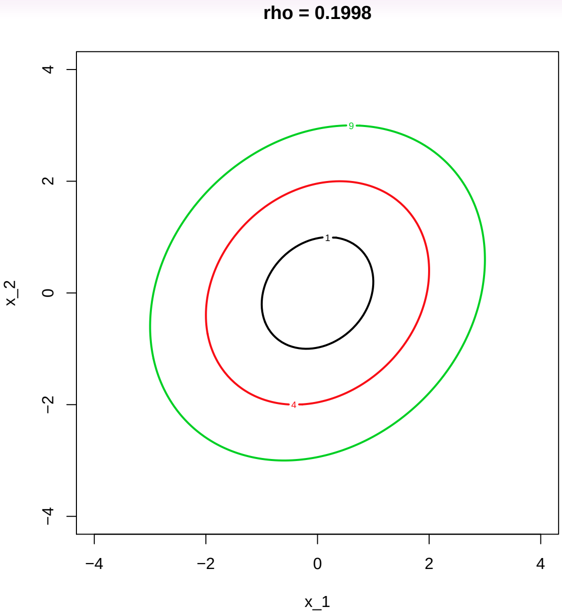
$$(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mu) = c^2$$

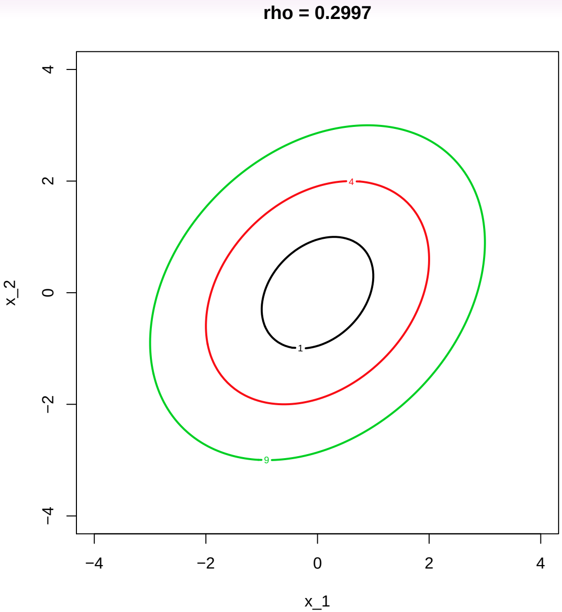
- Si $\boldsymbol{\Sigma} = \mathbf{H} \boldsymbol{\Lambda} \mathbf{H}^{\mathrm{T}}$, con $\mathbf{H}$ ortogonal y $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$, definamos $\mathbf{y} = \mathbf{H}^{\mathrm{T}} \mathbf{x}$ y $\boldsymbol{\nu} = \mathbf{H}^{\mathrm{T}} \boldsymbol{\mu}$.
  Los contornos de densidad constante son los elipsoides centrados en $\boldsymbol{\nu}$ con ejes principales de longitud $2c\lambda_j^{\frac{1}{2}}$ soportados en los autovectores, $\mathbf{h}_1, \ldots, \mathbf{h}_p$, o sea,

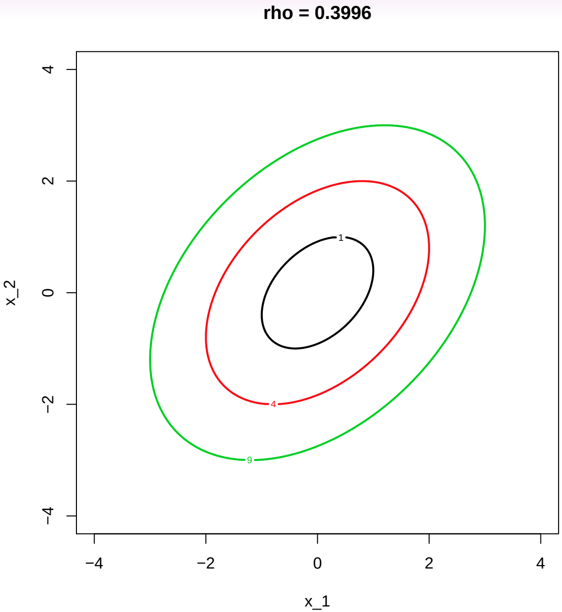
$$\sum_{j=1}^{p} \frac{(y_j - \nu_j)^2}{\lambda_j} = c^2$$

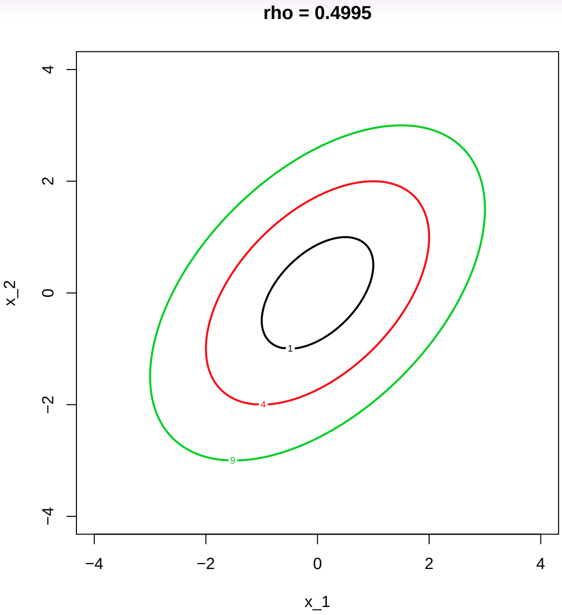

rho = 0
9
4
1
x_1
x_2
−4
−2
0
2
4

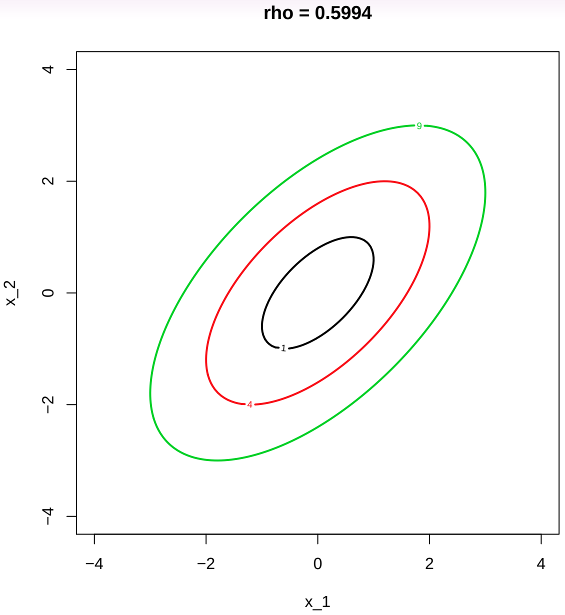

rho = 0.0999
x_2
x_1
4
2
0
−2
−4
−4
−2
0
2
4
1
4
9

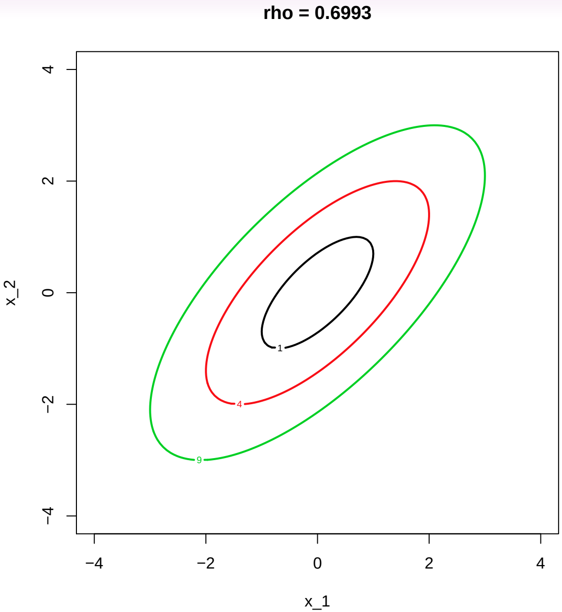

rho = 0.1998
x_1
x_2
−4
−2
0
2
4
9
4
1


rho = 0.2997
x_1
x_2
−4
−2
0
2
4
1
4
9

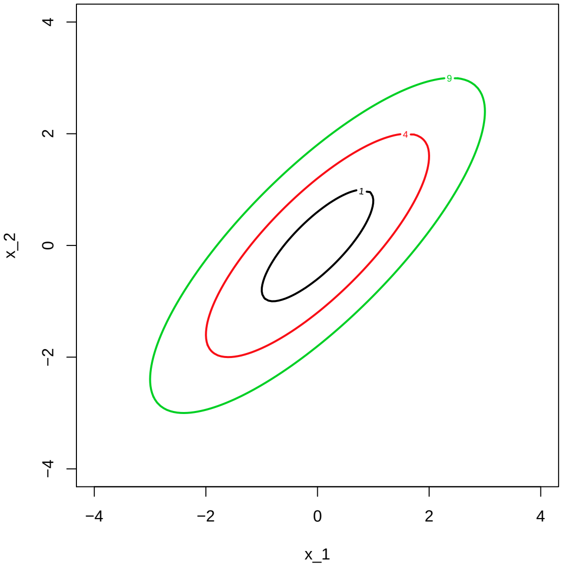

rho = 0.3996
x_2
x_1
4
2
0
−2
−4
−4
−2
0
2
4
1
4
9

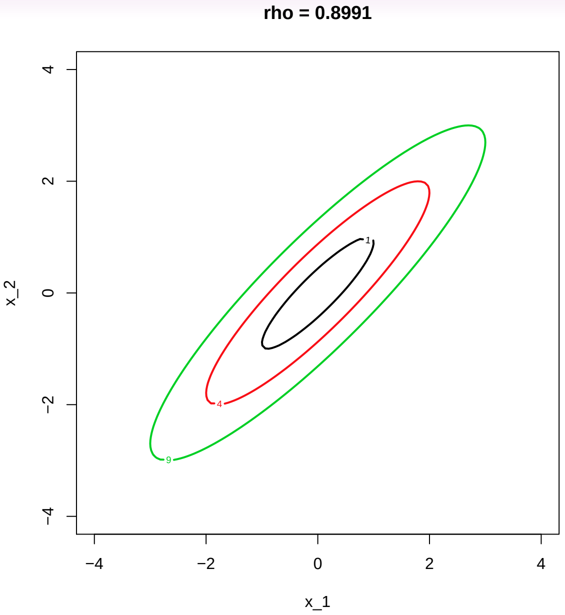

rho = 0.4995
x_2
x_1
4
2
0
−2
−4
1
4
9


rho = 0.5994
x_1
x_2
−4
−2
0
2
4
9
4
1

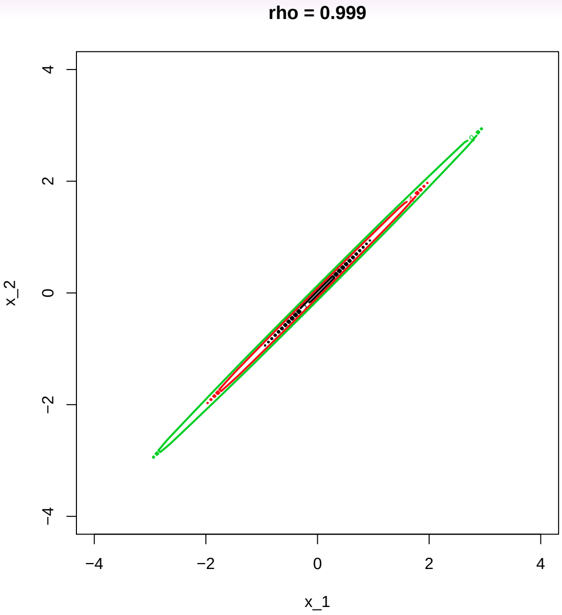

rho = 0.6993
x_2
x_1
4
2
0
−2
−4
−4
−2
0
2
4
1
4
9

rho = 0.7992





rho = 0.8991
x_1
x_2
−4
−2
0
2
4
1
4
9


rho = 0.999
x_2
4
2
0
−2
−4
x_1
−4
−2
0
2
4

Teorema

Sea $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} > 0$. Definamos $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix}$,
$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$ con $\mathbf{x}^{(i)}, \boldsymbol{\mu}^{(i)} \in \mathbb{R}^{p_i}$,
$\boldsymbol{\Sigma}_{ii} \in \mathbb{R}^{p_i \times p_i}$, $p_1 + p_2 = p$.

Entonces,

$$\mathbf{x}^{(1)} | \mathbf{x}^{(2)} = \mathbf{x}_0 \sim N\left(\boldsymbol{\mu}^{(1)} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_0 - \boldsymbol{\mu}^{(2)}), \boldsymbol{\Sigma}_{11.2}\right)$$

donde $\boldsymbol{\Sigma}_{11.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$.

Observación

- $\mathbf{\Sigma}_{11.2} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$
- $\mathbf{\Sigma}_{22.1} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}$

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} \mathbf{\Sigma}_{11.2}^{-1} & -\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22.1}^{-1} \\ -\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11.2}^{-1} & \mathbf{\Sigma}_{22.1}^{-1} \end{pmatrix}$$

- $\boldsymbol{\beta} = \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}$: **la matriz de regresión de $\mathbf{x}^{(1)}$ en $\mathbf{x}^{(2)}$.**
- $\mathbf{\Sigma}_{11.2}$: **matriz de covarianza parcial de $\mathbf{x}^{(1)}$ dado $\mathbf{x}^{(2)}$**
- $\boldsymbol{\mu}_{1.2} = \boldsymbol{\mu}^{(1)} - \boldsymbol{\beta}\boldsymbol{\mu}^{(2)}$
- $\mathbf{x}_{1.2} = \mathbf{x}^{(1)} - \mathbb{E}\left(\mathbf{x}^{(1)}|\mathbf{x}^{(2)}\right) = \mathbf{x}^{(1)} - \boldsymbol{\beta}\mathbf{x}^{(2)} - \boldsymbol{\mu}_{1.2}$ es el residuo y

$$\operatorname{COV}(\mathbf{x}_{1.2}) = \mathbf{\Sigma}_{11.2}$$

Caso $q = 1$

Sea $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} > 0$. Definamos $\mathbf{x} = \begin{pmatrix} x_1 \\ \mathbf{x}^{(2)} \end{pmatrix}$,

$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \boldsymbol{\sigma}_{12} \\ \boldsymbol{\sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$ con $\mathbf{x}^{(2)}, \boldsymbol{\mu}^{(2)} \in \mathbb{R}^{p-1}$.

Entonces,

$$x_1 | \mathbf{x}^{(2)} = \mathbf{x}_0 \sim N\left(\mu_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_0 - \boldsymbol{\mu}^{(2)}), 1/\sigma^{11}\right)$$

donde $\boldsymbol{\Sigma}^{-1} = (\sigma^{ij})_{1 \leq i,j \leq p}$.

Teorema

Sea $\mathbf{x}_1, \ldots \mathbf{x}_n$ i.i.d. $\mathbf{x}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} > 0$. Definamos

$$\mathbf{X}^{\mathrm{T}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n), \text{ o sea, } \mathbf{X} = \begin{pmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{pmatrix} = \left(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(p)}\right).$$

Se tiene,

a) $\mathbf{x}^{(j)} \sim N(\mathbf{0}, \sigma_{\mathbf{jj}}\, \mathbf{I_n})$

b) Dado $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{X}^{\mathrm{T}}\mathbf{a} = \sum_{i=1}^n a_i \mathbf{x}_i \sim N(\mathbf{0}, \|\mathbf{a}\|^2 \boldsymbol{\Sigma})$

c) Dados $\mathbf{a}_\ell = (a_{\ell,1}, \ldots, a_{\ell,n})^{\mathrm{T}} \in \mathbb{R}^n$, $1 \leq \ell \leq r$ con $r \leq n$ ortogonales, entonces $\mathbf{X}^{\mathrm{T}}\mathbf{a}_\ell = \sum_{i=1}^n a_{\ell,i} \mathbf{x}_i$ son independientes.

d) Dado $\mathbf{b} \in \mathbb{R}^p$, $\mathbf{X}\mathbf{b} = \sum_{j=1}^p b_j \mathbf{x}^{(j)} \sim N(\mathbf{0}, \sigma_{\mathbf{b}}^2 \mathbf{I_n})$