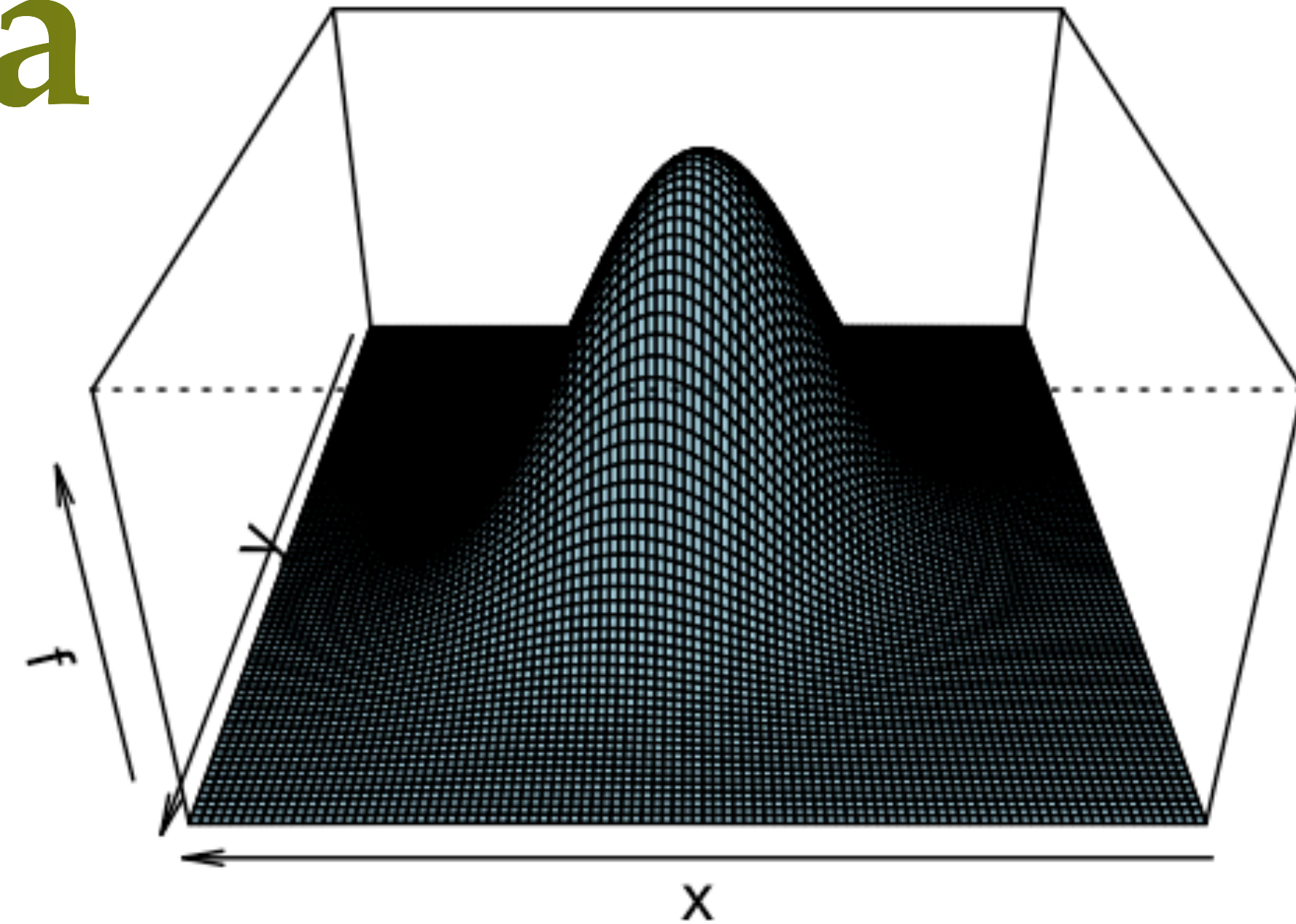


# Distribución normal multivariada



# Distribución normal multivariada

## Definición 1

Se dice que  $\mathbf{x}_{p \times 1}$  un vector aleatorio sigue una distribución normal multivariada no singular y denotado por  $\mathbf{x} \sim N_p(\mu, \Sigma)$ , si su densidad está dada por:

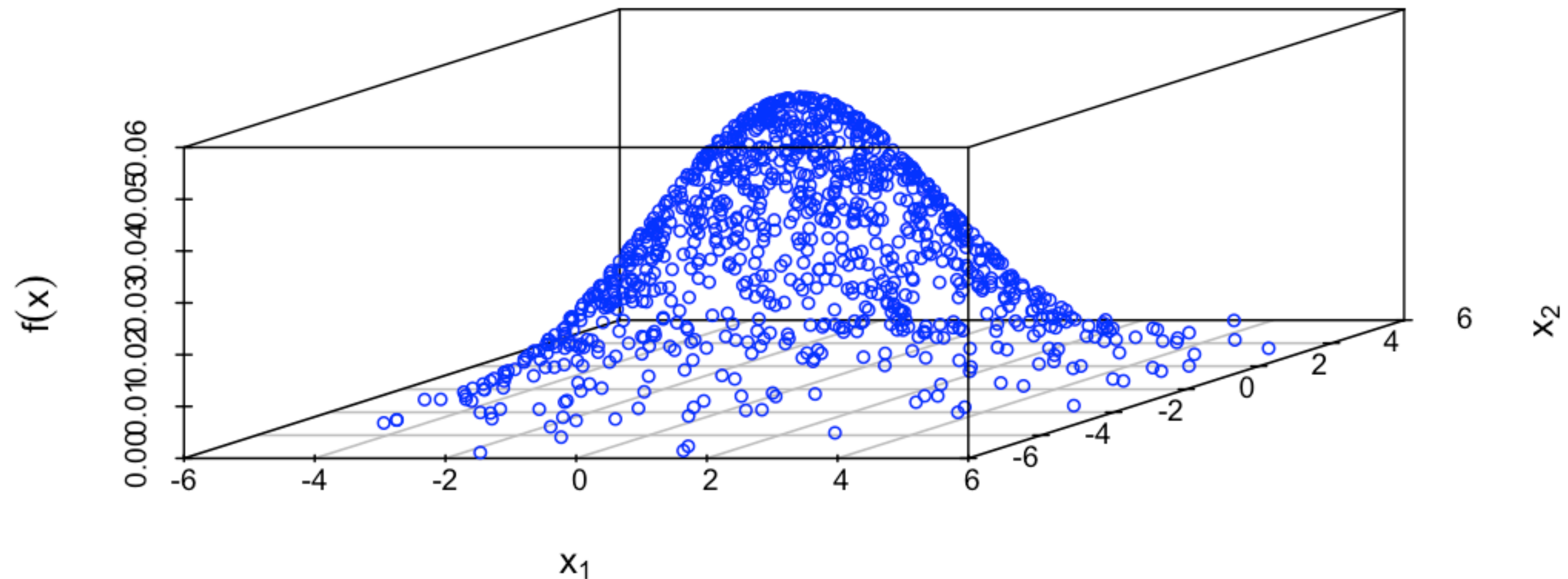
$$f(\mathbf{x}) = \frac{1}{|2\pi\Sigma|^{1/2}} \exp \left[ -\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu) \right],$$

donde

- $\mathbb{E}(\mathbf{x}) = \mu$
- $\text{Var}(\mathbf{x}) = \Sigma > 0$  (positiva definida)

# Diagrama de dispersión

Para vectores bivariados se puede obtener un diagrama de dispersión 3D con la librería `scatterplot3d`



# Distribución normal multivariada

## Observación 1

Si  $\text{ran}(\Sigma) = k < p$  se define la densidad de la distribución normal multivariada singular como

$$f(\mathbf{x}) = \frac{(2\pi)^{-\frac{k}{2}}}{(\lambda_1 \cdots \lambda_k)^{\frac{1}{2}}} \exp \left[ -\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^- (\mathbf{x} - \mu) \right],$$

donde

- $\mathbf{x}$  vive en el hiper-plano  $\mathbf{N}'(\mathbf{x} - \mu)$  y  $\mathbf{N}$  es una matriz de tamaño  $p \times (p - k)$  tal que:
  1.  $\mathbf{N}^T \Sigma = \mathbf{0}$
  2.  $\mathbf{N}^T \mathbf{N} = \mathbf{I}_{p-k}$
- $\Sigma^-$  es la inversa generalizada y  $\lambda_1, \dots, \lambda_k$  son los eigenvalores diferentes de cero.



# Caracterizaciones

## Definición 2

Decimos que el vector aleatorio  $\mathbf{x}_{p \times 1}$  sigue una distribución normal  $p$ -variada si y solo si  $\mathbf{a}^T \mathbf{x}$  tiene una distribución normal univariada para todos los vectores  $p$ -variados (no triviales)  $\mathbf{a}$

## Proposición 1

Sea  $\mathbf{x}$  un vector normal  $p$ -variado. Si se define a  $\mathbf{y}_{q \times 1} = \mathbf{A}_{q \times p} \mathbf{x} + \mathbf{b}_{q \times 1}$  entonces el vector aleatorio  $\mathbf{y}$  tiene una distribución normal  $q$ -variada con esperanza y varianza dadas por:

$$\mathbb{E}(\mathbf{y}) = \mathbf{A}\boldsymbol{\mu} + \mathbf{b}$$

$$\text{Var}(\mathbf{y}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T$$

# Caracterizaciones

## Corolario 1

Sean  $\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{I}_p)$ ,  $\mu_{p \times 1}$  un vector real y  $\Sigma_{p \times p}$  una matriz semidefinida positiva y simétrica. Entonces  $\mathbf{y} = \Sigma^{\frac{1}{2}}\mathbf{x} + \mu \sim N_p(\mu, \Sigma)$

## Corolario 2

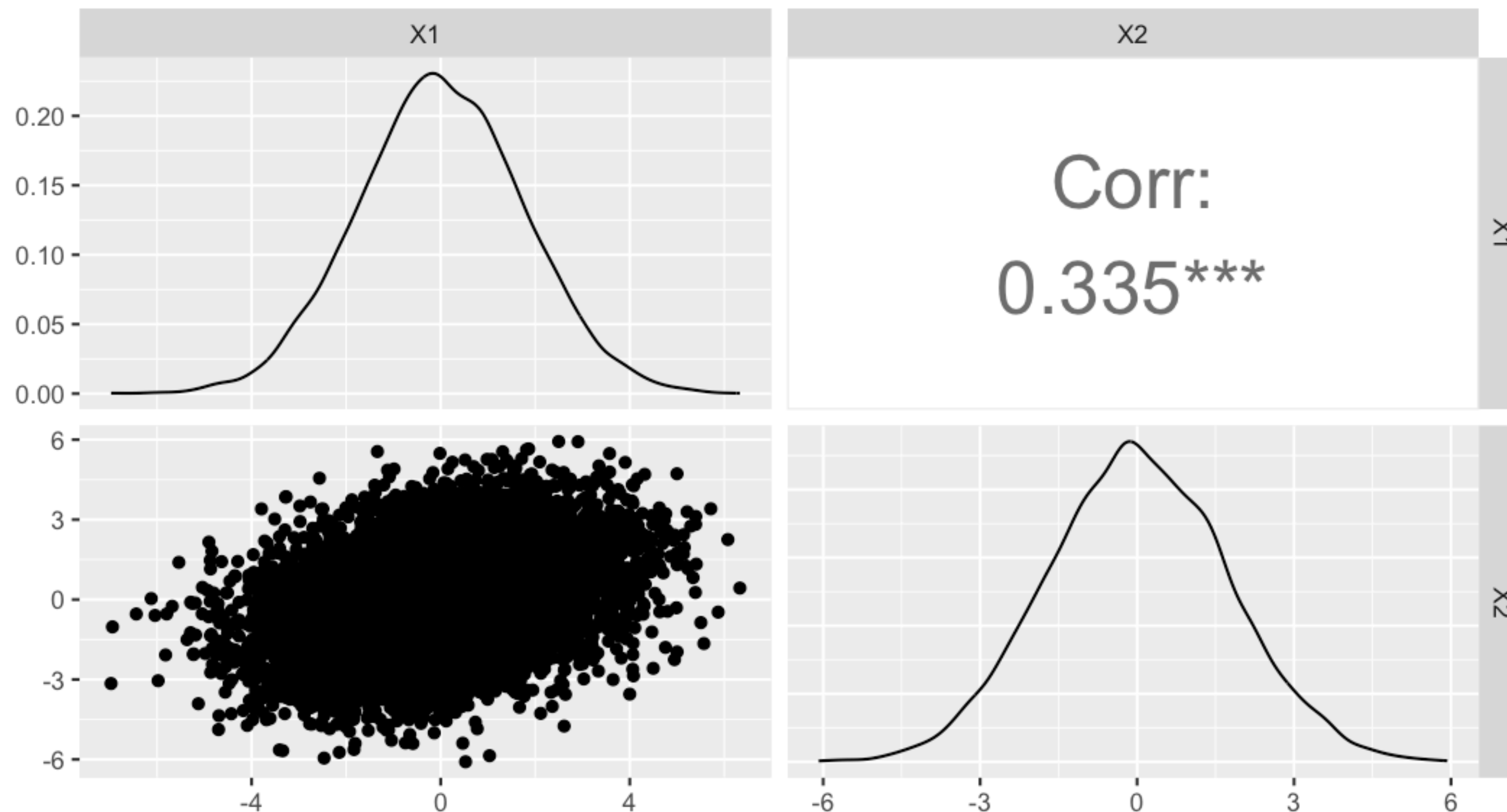
Sean  $\mathbf{x} \sim N_p(\mu, \Sigma)$  con  $\Sigma > 0$  y  $\mathbf{y} = \Sigma^{-\frac{1}{2}}(\mathbf{x} - \mu)$ , donde  $\Sigma^{-\frac{1}{2}}$  es la matriz raíz cuadrada de  $\Sigma^{-1}$ . Entonces,  $y_1, y_2, \dots, y_p$  son variables aleatorias iid  $N(0, 1)$ .

- En **R** la librería **expm** proporciona la función requerida para obtener  $\Sigma^{-\frac{1}{2}}$  con **sqrtm**

# Simulación

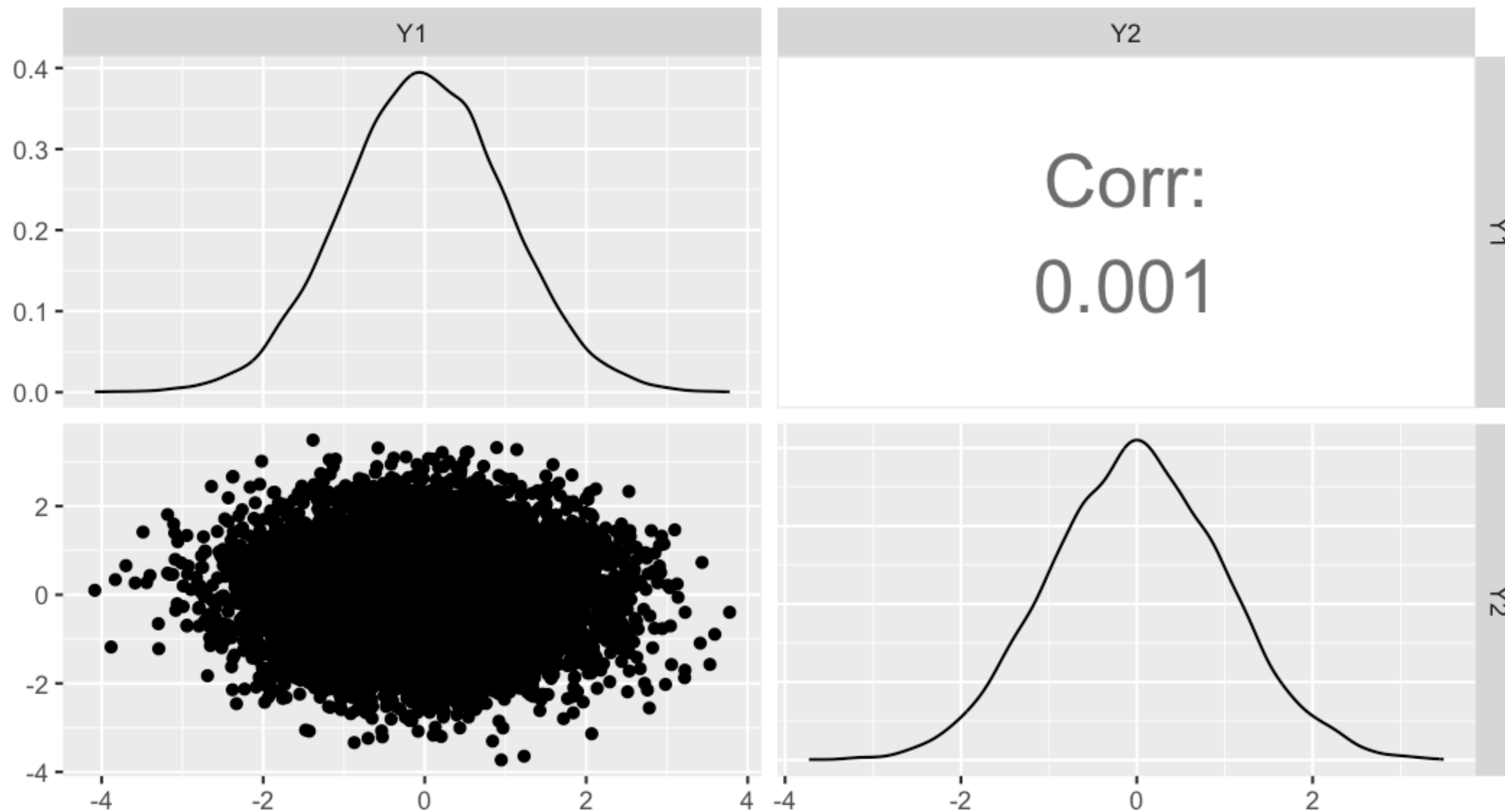
- ¿Cómo se ven las densidades univariadas y las curvas de nivel de  $\mathbf{x} \sim N_2(\mu, \Sigma)$ ? Donde

$$\mu = \begin{pmatrix} 3 \\ 3 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$



# Simulación

- ¿Qué pasa si hacemos  $\mathbf{y} = \Sigma^{-\frac{1}{2}}(\mathbf{x} - \mu)$ ?



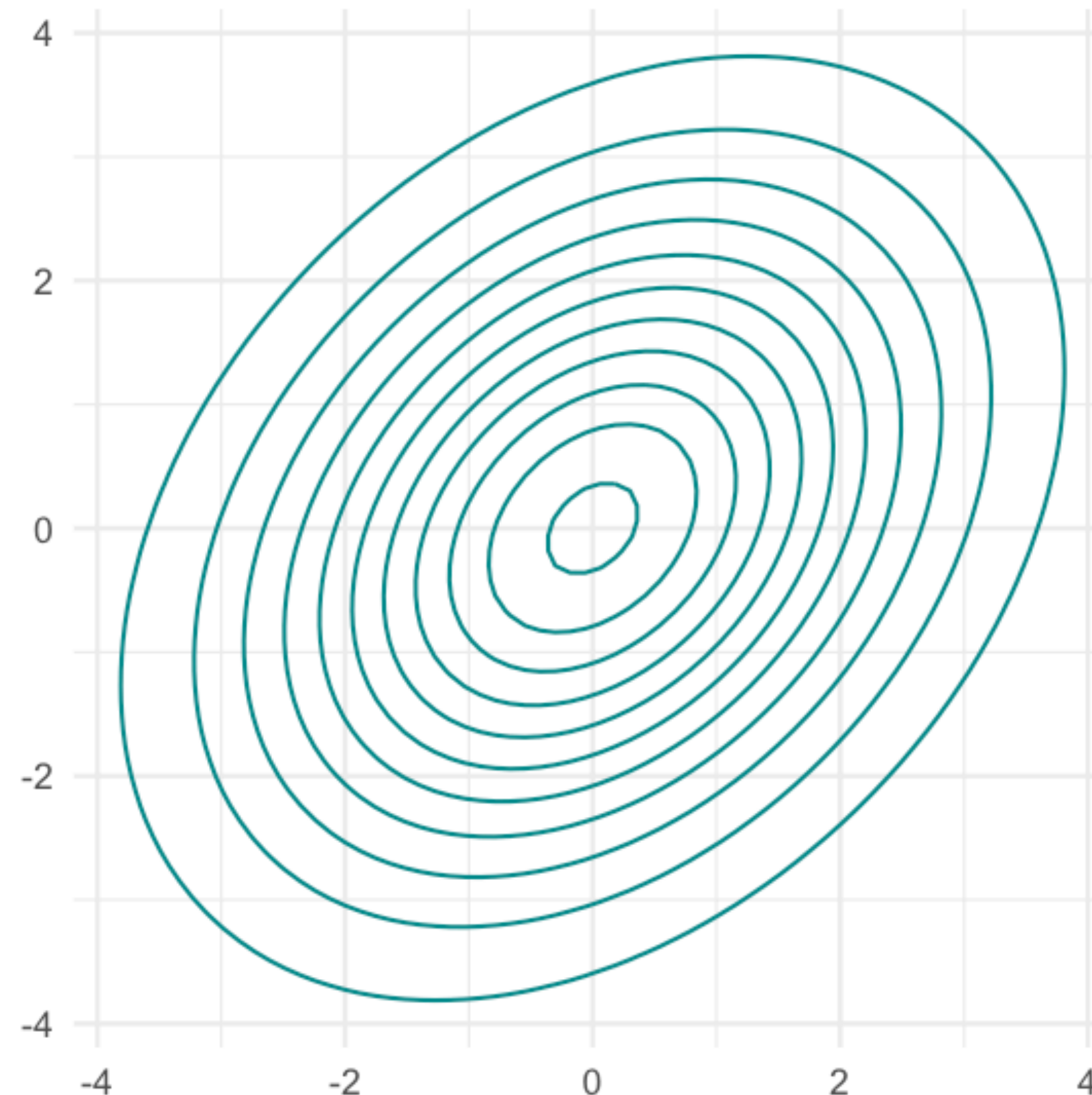


# Propiedades

## Observación 2

La distribución normal multivariada tiene densidad constante en elipses (elipsoides)

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



# Propiedades

## Proposición 2

Sea  $\mathbf{x} \sim N_p(\mu, \Sigma)$  entonces,  $U = (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \sim \chi_p^2$ .

## Observación 3

Se puede fácilmente evaluar la probabilidad de que  $\mathbf{x}$  este en un elipsoide, i.e.

$$\mathbb{P} \left[ (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) < k \right]$$

# Propiedades

## Proposición 3

Sea  $\mathbf{x} \sim N_p(\mu, \Sigma)$ , entonces los coeficientes de asimetría y curtosis están dados respectivamente por,

$$\beta_{1,p} = \mathbb{E} \left[ \left( (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right)^3 \right] = 0$$

$$\beta_{2,p} = \mathbb{E} \left[ \left( (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right)^2 \right] = p(p + 2)$$

## Proposición 4

Sea  $\mathbf{x} \sim N_p(\mu, \Sigma)$ , entonces la función característica de  $\mathbf{x}$  está dada por,

$$\phi(\mathbf{t}) = \exp \left( i \mathbf{t}^T \mu - \frac{1}{2} \mathbf{t}^T \Sigma \mathbf{t} \right)$$

# Propiedades

## Proposición 5

Sea  $\mathbf{x} \sim N_p(\mu, \Sigma)$  y considere la partición

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \quad \mu = \begin{pmatrix} \mu^{(1)} \\ \mu^{(2)} \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

donde  $\mathbf{x}^{(1)}$  es de dimensión  $k$  y  $\mathbf{x}^{(2)}$  es de dimensión  $p - k$ , entonces:

1.  $\mathbf{x}^{(1)} \sim N_k(\mu^{(1)}, \Sigma_{11})$

2.  $\mathbf{x}^{(1)}$  y  $\mathbf{x}^{(2)}$  son independientes si y solo si  $\text{Cov}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = \Sigma_{12} = \mathbf{0}$

3.  $\mathbf{x}^T \Sigma^{-1} \mathbf{x} \sim \chi_p^2(\mu^T \Sigma^{-1} \mu)$

4.  $\mathbf{x}^{(2)} | \mathbf{x}^{(1)} \sim N_{p-k}(\mu^{(2)} + \Sigma_{21} \Sigma_{11}^{-1} [\mathbf{x}^{(1)} - \mu^{(1)}], \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12})$

# Pruebas de normalidad

- Todas las distribuciones univariadas son normales
  - qqplot
  - histogramas
  - Pruebas de normalidad (e.g. Anderson-Darling, Shapiro-Wilk, Lilliefors, etc.)
- $(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \sim \chi_p^2$ 
  - qqplot
- Prueba de Mardia (1970) basada en los coeficientes de asimetría y curtosis multivariados
- Otras pruebas (e.g. Henze-Zirkler (1990), Royston (1982))
- En **R**: librería **MVN**



# Convergencia

## Teorema 1 (Teorema Central del Límite)

Sean  $\mathbf{X}_n = (x_{n1}, \dots, x_{np})$  una colección de vectores aleatorios independientes e idénticamente distribuidos, con vector de medias  $\mu$  y matriz (finita) de covarianza  $\Sigma$ . Entonces,

$$\sqrt{n} (\bar{\mathbf{x}} - \mu) \rightarrow N_p(\mathbf{0}_p, \Sigma)$$

## Teorema 2 (Teorema de Cramér-Wold)

Para  $\mathbf{X}_n = (x_{n1}, \dots, x_{np})$  y  $\mathbf{X} = (x_1, \dots, x_p)$  dos vectores aleatorios y  $\mathbf{t} \in \mathbb{R}^p$ , entonces

$$\mathbf{X}_n \xrightarrow{d} \mathbf{X} \quad \Leftrightarrow \quad \sum_{i=1}^p t_i x_{ni} \xrightarrow{d} \sum_{i=1}^p t_i x_i$$

# Distribución Wishart

# Distribución Wishart

## Definición 3

Sea  $\mathbf{M}_{p \times p}$  una matriz simétrica de variables aleatorias, tal que  $\mathbb{P}(\mathbf{M} > 0) = 1$ , y sea  $\Sigma_{p \times p}$  una matriz definida positiva. Si  $n \in \mathbb{N}$ , tal que  $n \geq p$ , entonces  $\mathbf{M}_{p \times p}$  tiene una distribución Wishart,  $\mathbf{M} \sim W_p(n, \Sigma)$ , no singular con  $n$  grados de libertad si la función de densidad de los  $\frac{p(p+1)}{2}$  distintos elementos de  $\mathbf{M}_{p \times p}$  está dada por:

$$f(m_{11}, m_{12}, \dots, m_{pp}) = c^{-1} |\mathbf{M}|^{(n-p-1)/2} \text{etr} \left( -\frac{\Sigma^{-1} \mathbf{M}}{2} \right)$$

donde

- $\text{etr}$  es el operador  $\exp^{\text{trace}}$

- $c = 2^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}} \Gamma_p \left( \frac{n}{2} \right)$  y  $\Gamma_p(\cdot)$  la función gamma multivariada

# Distribución Wishart

## Definición 4

Sean  $\mathbf{x}_1, \dots, \mathbf{x}_n$  vectores aleatorios iid distribuidos como  $N_p(\mathbf{0}, \Sigma)$  entonces  $\mathbf{M}_{p \times p} = \mathbf{X}^T \mathbf{X}$  tiene una distribución Wishart con  $n$  grados de libertad

## Observación 4

Si  $\Sigma > 0$  y  $n \geq p$ , entonces se puede probar que  $\mathbb{P}(\mathbf{M} > 0) = 1$ . De lo contrario, se tiene que  $\mathbf{M} \geq 0$ , por lo que la densidad no existe y se dice que  $\mathbf{M}$  tiene una distribución singular

# Propiedades

## Teorema 3

Sea  $\mathbf{M} \sim W_p(n, \Sigma)$  entonces, si  $\mathbf{C}_{q \times p}$  tal que  $\text{ran}(\mathbf{C}) = q$ , se tiene que  $\mathbf{CMCT} \sim W_q(n, \mathbf{C}\Sigma\mathbf{C}^T)$

## Corolario 3

Si  $\mathbf{M} \sim W_p(n, \Sigma)$  y  $\mathbf{a}$  es un vector de constantes, entonces  $\mathbf{a}^T\mathbf{M}\mathbf{a} \sim \sigma_{\mathbf{a}}^2 \cdot \chi_n^2$ , donde  $\sigma_{\mathbf{a}}^2 = \mathbf{a}^T\mathbf{\Sigma}\mathbf{a}$

## Corolario 4

Si  $\mathbf{M} \sim W_p(n, \Sigma)$  entonces  $m_{ii} \sim \Sigma_{ii} \cdot \chi_n^2$



# Propiedades

## Proposición 6

Sea  $\mathbf{M} \sim W_p(n, \Sigma)$  entonces:

1.  $\mathbb{E}(\mathbf{M}) = n\Sigma$

2. (Aditividad) Si  $\mathbf{M}_i \sim W_p(n_i, \Sigma)$  independientes entonces,  $\sum_{i=1}^m \mathbf{M}_i \sim W_p\left(\sum_{i=1}^m n_i, \Sigma\right)$

3. Si partimos a  $\mathbf{M}$  y a  $\Sigma$  como,

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

entonces,  $\mathbf{M}_{11} \sim W_k(n, \Sigma_{11})$  y  $\mathbf{M}_{22} \sim W_{p-k}(n, \Sigma_{22})$ . Más aún si  $\Sigma_{12} = 0$ , entonces  $\mathbf{M}_{11}$  y  $\mathbf{M}_{22}$  son independientes.

# Formas cuadráticas

## Teorema 4

Sea  $\mathbf{M} \sim W_p(n, \Sigma)$ , entonces

1. Sea  $\mathbf{A}_{q \times p}$  una matriz tal que  $\text{ran}(\mathbf{A}) = q$ , entonces  $(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)^{-1} \sim W_q\left(n - p + q, (\mathbf{A}\Sigma^{-1}\mathbf{A}^T)^{-1}\right)$
2. Sea  $\mathbf{y}_{p \times 1}$  independiente de  $\mathbf{M}$  y tal que  $\mathbb{P}(\mathbf{y} = \mathbf{0}) = 0$ , entonces  $\frac{\mathbf{y}^T \mathbf{M} \mathbf{y}}{\mathbf{y}^T \Sigma \mathbf{y}} \sim \chi_n^2$  y  $\frac{\mathbf{y}^T \Sigma^{-1} \mathbf{y}}{\mathbf{y}^T \mathbf{M}^{-1} \mathbf{y}} \sim \chi_{n-p+1}^2$
3. Sean  $\mathbf{x}_1, \dots, \mathbf{x}_n$  vectores aleatorios iid  $N_p(\mathbf{0}, \Sigma)$ . Entonces si consideramos a  $\mathbf{y} = \mathbf{X}\mathbf{a}$  con  $\mathbf{a}_{p \times 1}$ ,  $\mathbf{A}_{n \times n}$ ,  $\mathbf{B}_{n \times n}$  matrices simétricas de rango  $r, s$  respectivamente y  $\mathbf{b}_{n \times 1}$  un vector de constantes entonces
  - $\mathbf{X}^T \mathbf{A} \mathbf{X} \sim W_p(r, \Sigma)$  si y solo si  $\mathbf{y}^T \mathbf{A} \mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$
  - $\mathbf{X}^T \mathbf{A} \mathbf{X} \sim W_p(r, \Sigma)$  y  $\mathbf{X}^T \mathbf{B} \mathbf{X} \sim W_p(s, \Sigma)$  son ind. sii  $\mathbf{y}^T \mathbf{A} \mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$  y  $\mathbf{y}^T \mathbf{B} \mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$  son ind.
  - $\mathbf{X}^T \mathbf{b} \sim N_p$  y  $\mathbf{X}^T \mathbf{A} \mathbf{X} \sim W_p(r, \Sigma)$  son ind. sii  $\mathbf{y}^T \mathbf{b} \sim N_1$  y  $\mathbf{y}^T \mathbf{A} \mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$  son ind.

# Formas cuadráticas

## Lema 1

Sean  $\mathbf{x}_1, \dots, \mathbf{x}_n \sim N_p(\mathbf{0}, \Sigma)$  (iid) entonces se cumple lo siguiente

1.  $\mathbf{x}^{(j)} \sim N_n(\mathbf{0}, \sigma_{jj}\mathbf{I})$

2. Si  $\mathbf{a}_{n \times 1}$  es un vector de constantes entonces  $\mathbf{X}^T \mathbf{a} \sim N_p(\mathbf{0}, ||\mathbf{a}||^2 \Sigma)$

3. Si  $\{\mathbf{a}_1, \dots, \mathbf{a}_r\}$ ,  $r \leq n$ , es un conjunto de vectores mutuamente ortogonal entonces, los vectores aleatorios dados por  $\mathbf{X}^T \mathbf{a}_i$  son mutuamente independientes

4. Si  $\mathbf{b}_{p \times 1}$  es un vector de constantes, entonces  $\mathbf{X}\mathbf{b} \sim N_n(\mathbf{0}, \sigma_b^2 \mathbf{I})$  donde  $\sigma_b^2 = \mathbf{b}^T \Sigma \mathbf{b}$

# Formas cuadráticas

## Lema 2

Sea  $\mathbf{x} \sim N_p(\mathbf{0}, \sigma^2 \mathbf{I})$  y  $\mathbf{A}_{p \times q}$  una matriz simétrica entonces  $\mathbf{x}^T \mathbf{A} \mathbf{x} \sim \sigma^2 \cdot \chi_r^2$  si y solo si  $\mathbf{A}$  es idempotente y con  $\text{ran}(\mathbf{A}) = r$

## Lema 3

Sea  $\mathbf{x} \sim N_p(\mathbf{0}, \sigma^2 \mathbf{I})$  y sean  $\mathbf{Q}_i = \mathbf{x}^T \mathbf{P}_i \mathbf{x} \sim \sigma^2 \cdot \chi_{r_i}^2$  ( $i = 1, 2$ ) dos formas cuadráticas,  
Entonces  $\mathbf{Q}_1$  y  $\mathbf{Q}_2$  son independientes si y solo si  $\mathbf{P}_1 \mathbf{P}_2 = \mathbf{0}$