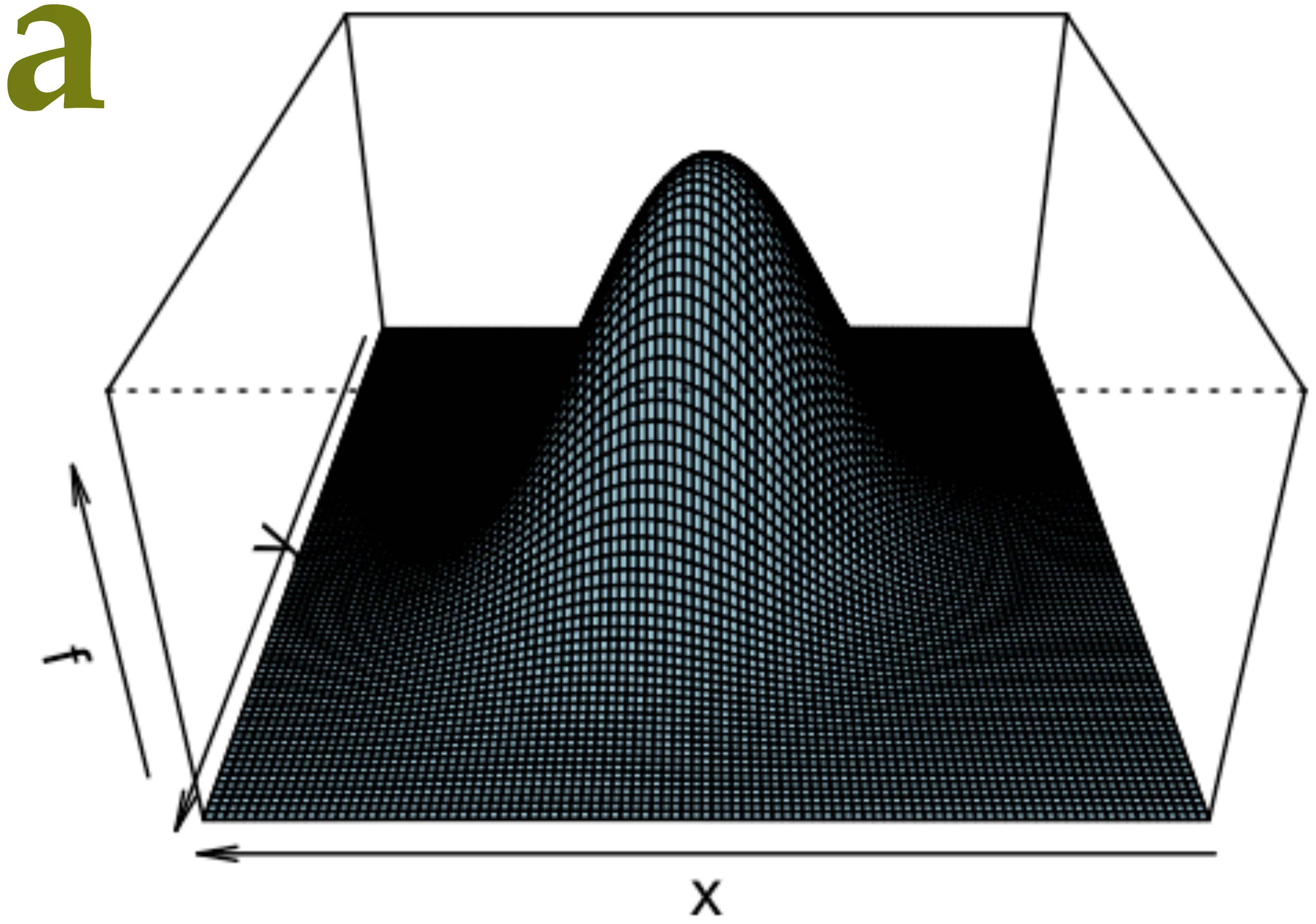
Distribución normal multivariada



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 $\mathbf{x} \sim N_p\left(\mu, \Sigma\right)$, si su densidad está dada por:

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

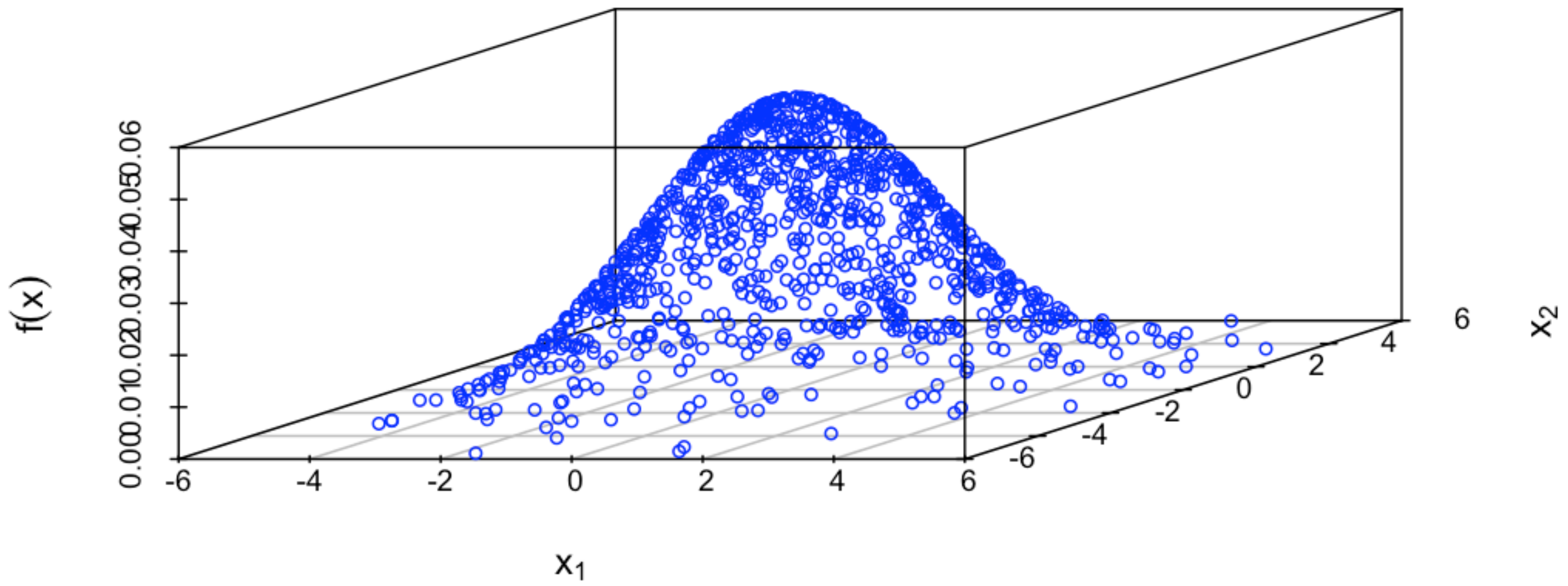
donde

- $\mathbb{E}(\mathbf{x}) = \mu$

- $\mathrm{Var}(\mathbf{x}) = \Sigma > 0$ (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 híper-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. $\mathbf{N}^T \mathbf{N} = \mathbf{I_{p-k}}$

- Σ^- es la inversa generalizada y $\lambda_1, \ldots, \lambda_k$ 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:

$$\mathbb{E}(\mathbf{y}) = \mathbf{A}\mu + \mathbf{b} \qquad \mathrm{Var}(\mathbf{y}) = \mathbf{A}\Sigma\mathbf{A}^{\mathbf{T}}$$

Caracterizaciones

Corolario 1

Sean $\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{I}_p)$, $\mu_{p\times1}$ un vector real y $\Sigma_{p\times p}$ una matriz semidefinda positiva y simétrica. Entonces $\mathbf{y} = \Sigma^{\frac{1}{2}}\mathbf{x} + \mu \sim N_p(\mu, \Sigma)$

Corolario 2

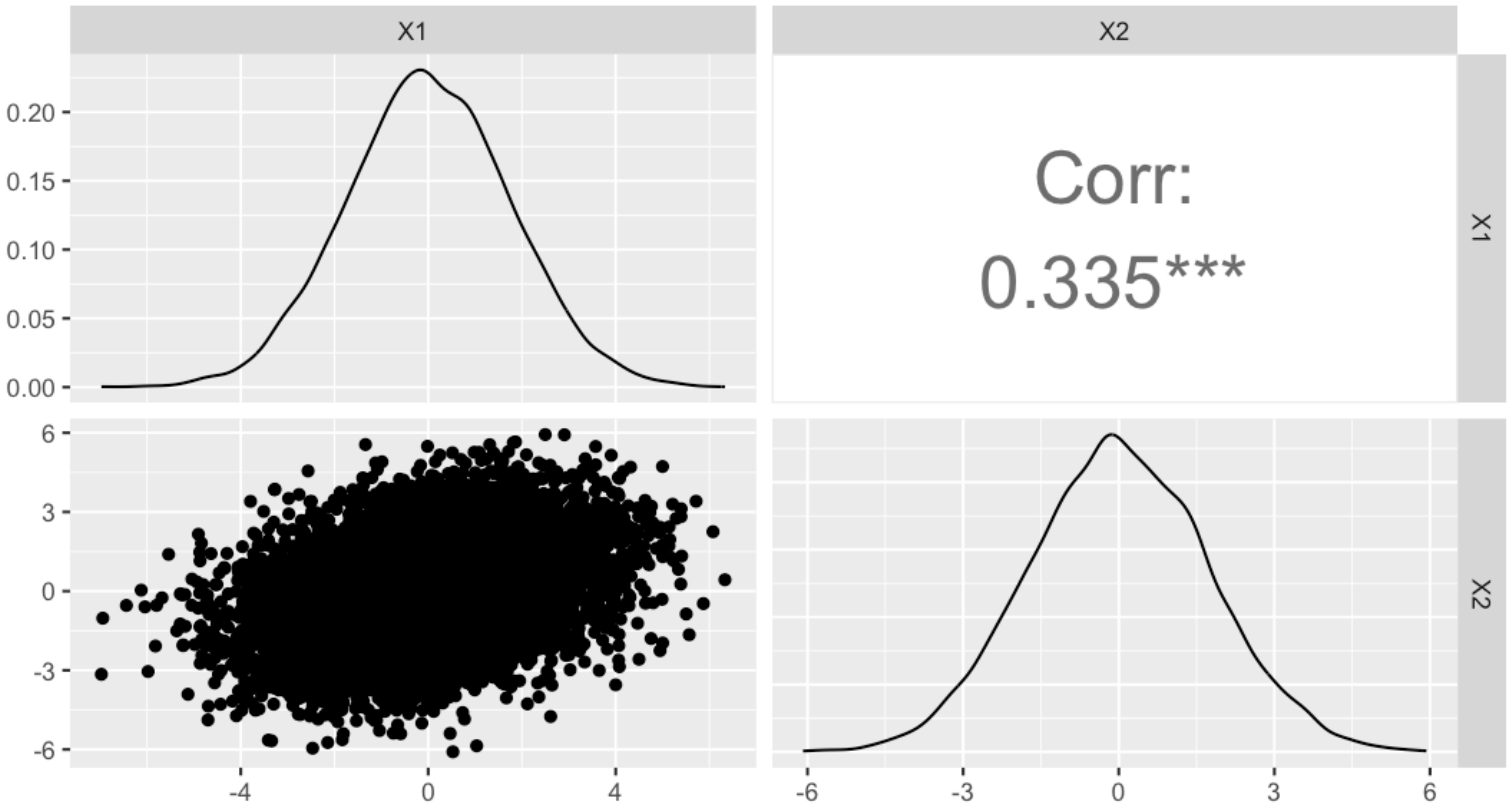
Sean $\mathbf{x} \sim N_p(\mu, \Sigma)$ con $\Sigma > 0$ y $\mathbf{y} = \Sigma^{-\frac{1}{2}}(\mathbf{x} - \mu)$, donde $\Sigma^{-\frac{1}{2}}$ es la matriz raíz cuadrada de Σ^{-1}. Entonces, $y_1, y_2, \ldots, y_p$ son variables aleatorias iid $N(0,1)$.

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

Simulación

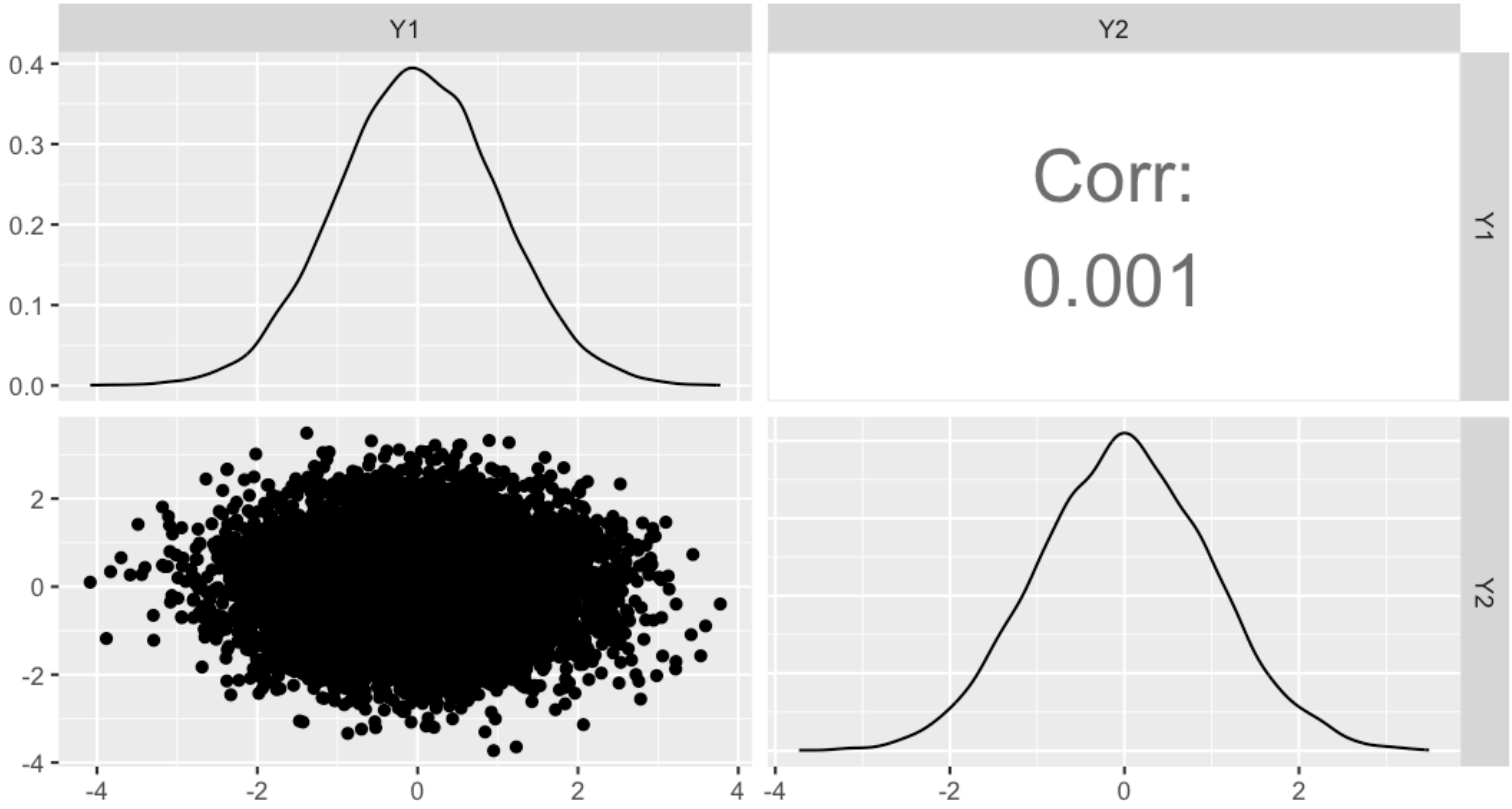
- ¿Cómo se ven las densidades univariadas y las curvas de nivel de $\mathbf{x} \sim N_2(\mu, \Sigma)$? Donde

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



Simulación

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

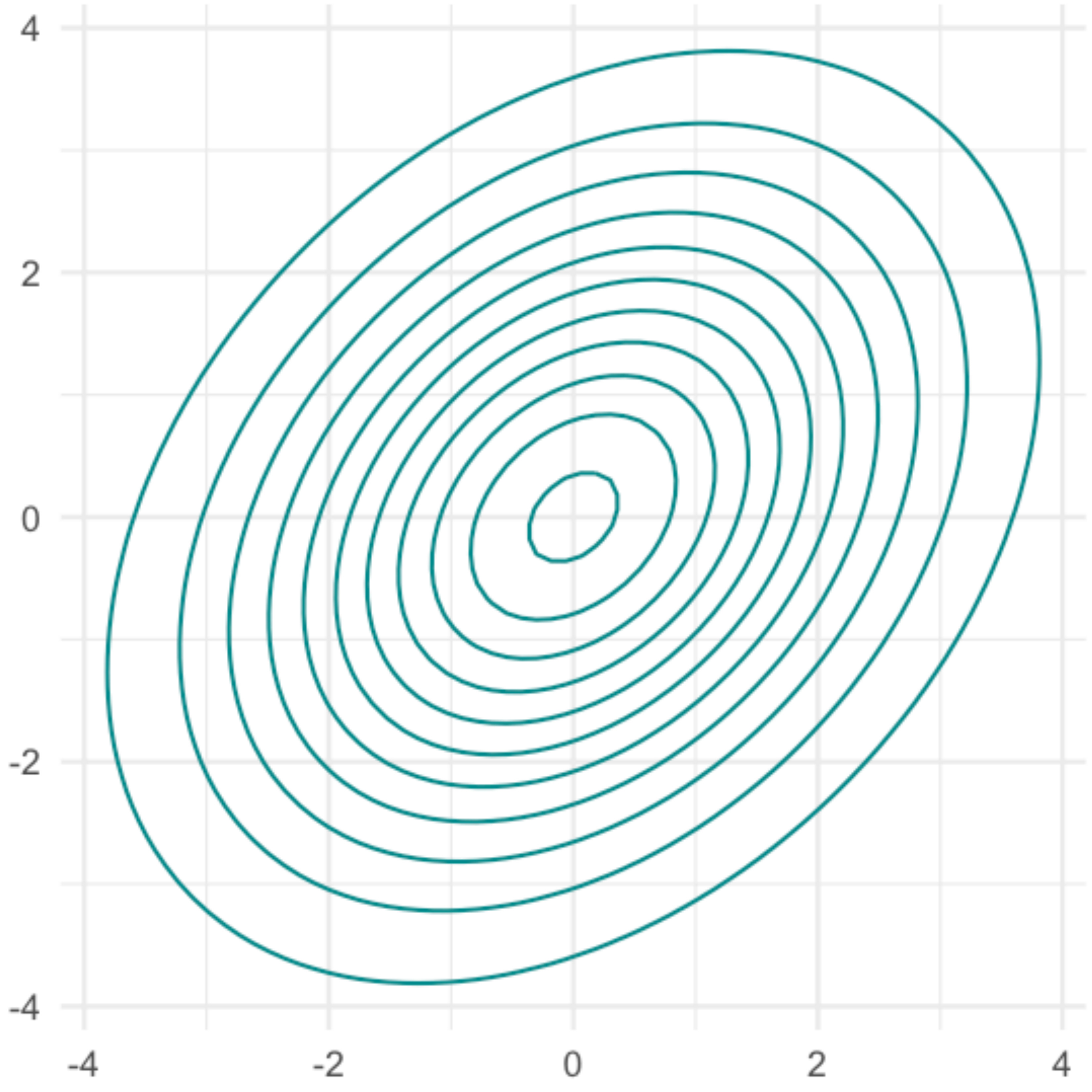


Propiedades

Observación 2

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

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



Propiedades

Proposición 2

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

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,

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

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

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

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

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

1. $\mathbf{x}^{(1)} \sim N_k\left(\mu^{(1)}, \Sigma_{11}\right)$
2. $\mathbf{x}^{(1)}$ y $\mathbf{x}^{(2)}$ son independientes si y solo si $\mathrm{Cov}\left(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}\right) = \Sigma_{12} = \mathbf{O}$
3. $\mathbf{x}^T\Sigma^{-1}\mathbf{x} \sim \chi_p^2\left(\mu^T\Sigma^{-1}\mu\right)$
4. $\mathbf{x}^{(2)} \mid \mathbf{x}^{(1)} \sim N_{p-k}\left(\mu^{(2)} + \Sigma_{21}\Sigma_{11}^{-1}[\mathbf{x}^{(1)} - \mu^{(1)}], \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)$

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 $\mathbf{x}_n = (x_{n1}, \ldots, x_{np})$ una colección de vectores aleatorios independientes e idénticamente distribuidos, con vector de medias μ y matriz (finita) de covarianza Σ. Entonces,

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

Teorema 2 (Teorema de Cramér-Wold)

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

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

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:

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

donde

- 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 $\mathbf{x}_1, \ldots, \mathbf{x}_n$ 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 $\mathbf{CMC^T} \sim W_q\left(n, \mathbf{C}\Sigma\mathbf{C}^T\right)$

Corolario 3

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

Corolario 4

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

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 Σ como,

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

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 $\left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^{\mathbf{T}}\right)^{-1} \sim W_q\left(n - p + q, \left(\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{A}^{T}\right)^{-1}\right)$

2. Sea $\mathbf{y}_{p\times 1}$ independiente de $\mathbf{M}$ y tal que $\mathbb{P}(\mathbf{y} = \mathbf{0}) = 0$, entonces $\dfrac{\mathbf{y}^{\mathbf{T}}\mathbf{M}\mathbf{y}}{\mathbf{y}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{y}} \sim \chi^2_n$ y $\dfrac{\mathbf{y}^{\mathbf{T}}\boldsymbol{\Sigma}^{-1}\mathbf{y}}{\mathbf{y}^{\mathbf{T}}\mathbf{M}^{-1}\mathbf{y}} \sim \chi^2_{n-p+1}$

3. Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ 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

   - $\mathbf{X}^{\mathbf{T}}\mathbf{A}\mathbf{X} \sim W_p(r, \Sigma)$ si y solo si $\mathbf{y}^{\mathbf{T}}\mathbf{A}\mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$
   - $\mathbf{X}^{\mathbf{T}}\mathbf{A}\mathbf{X} \sim W_p(r, \Sigma)$ y $\mathbf{X}^{\mathbf{T}}\mathbf{B}\mathbf{X} \sim W_p(s, \Sigma)$ son ind. sii $\mathbf{y}^{\mathbf{T}}\mathbf{A}\mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$ y $\mathbf{y}^{\mathbf{T}}\mathbf{B}\mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$ son ind.
   - $\mathbf{X}^{\mathbf{T}}\mathbf{b} \sim N_p$ y $\mathbf{X}^{\mathbf{T}}\mathbf{A}\mathbf{X} \sim W_p\,(r, \Sigma)$ son ind. sii $\mathbf{y}^{\mathbf{T}}\mathbf{b} \sim N_1$ y $\mathbf{y}^{\mathbf{T}}\mathbf{A}\mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$ son ind.

Formas cuadráticas

Lema 1

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

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

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

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

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

Formas cuadráticas

Lema 2

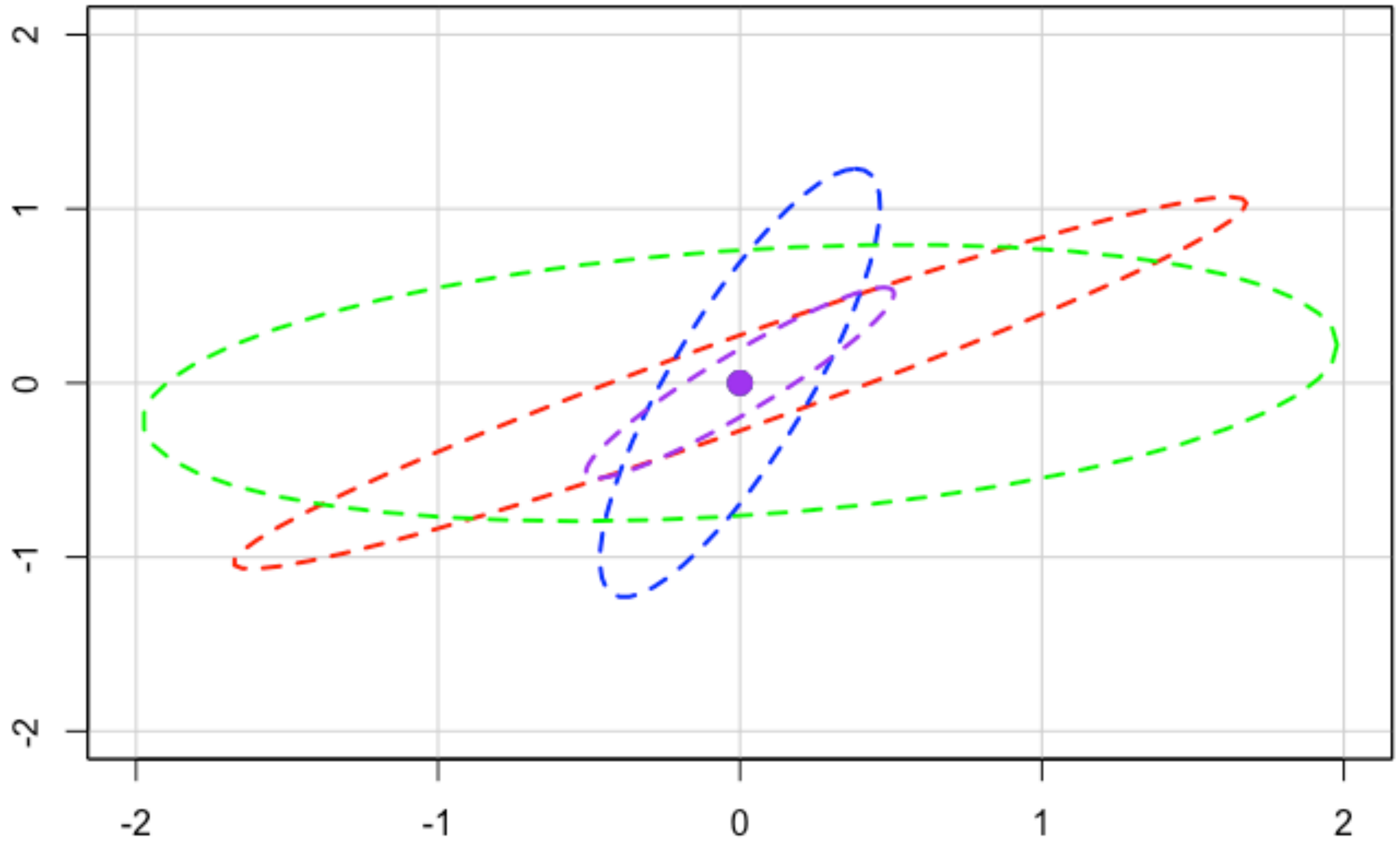
Sea $\mathbf{x} \sim N_p\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ y $\mathbf{A}_{p\times q}$ una matriz simétrica entonces $\mathbf{x}^T\mathbf{A}\mathbf{x} \sim \sigma^2 \cdot \chi^2_r$ si y solo si $\mathbf{A}$ es idempotente y con $\operatorname{ran}(\mathbf{A}) = r$

Lema 3

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

Simulaciones

- En **R:** rWishart
- Para entender su aleatoriedad podemos graficar las elipses generadas: $\mathbf{a}^{\mathbf{T}}\mathbf{M}_{\mathbf{i}}\mathbf{a} = c$



Distribución Wishart no centrada

Definición 5 (Distribución Wishart no centrada)

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ vectores aleatorios independientes y distribuidos como $N_p(\mu_{\mathbf{i}}, \Sigma)$, entonces $\mathbf{M}_{p\times p} = \mathbf{X}^T\mathbf{X}$ tiene una distribución Wishart no centrada, $\mathbf{M} \sim W_p(n, \Sigma, \Delta)$, con n grados de libertad y matriz de no centralidad Δ definida como

$$\Delta = \sum_{i=1}^{n} (\Sigma^{-\frac{1}{2}}\mu_i)(\Sigma^{-\frac{1}{2}}\mu_i)^T = \Sigma^{-\frac{1}{2}}\Lambda^T\Lambda\Sigma^{-\frac{1}{2}}$$

donde

- $\Lambda = \left(\mu_1, \ldots, \mu_n\right)^T$

Distribución T^2 de Hotelling

Distribución T^2 de Hotelling

Teorema 5 (Distribución centrada)

Sean $\mathbf{x} \sim N_p(\mu, \Sigma)$ y $\mathbf{M} \sim W_p(n, \Sigma)$ independientes y no singulares, entonces,

$$T^2 = n(\mathbf{x} - \mu)^T \mathbf{M}^{-1} (\mathbf{x} - \mu) \sim \left(\frac{np}{n - p + 1} \right) F_{p,n-p+1} = T^2_{p,n}$$

Corolario 5

Sean $\mathbf{x} \sim N_p(\mu, \lambda^{-1}\Sigma)$ y $\mathbf{M} \sim W_p(n, \Sigma)$ independientes y no singulares, entonces,

$$\lambda \left(\mathbf{x} - \mu\right)^T \left(\frac{M}{n} \right)^{-1} \left(\mathbf{x} - \mu\right) \sim T^2_{p,n}$$

Distribución T^2 de Hotelling

Teorema 6 (Distribución no centrada)

Sean $\mathbf{x} \sim N_p(\mu, \Sigma)$ y $\mathbf{M} \sim W_p(n, \Sigma)$ independientes y no singulares, y denotemos por $\delta = \mu^T \Sigma^{-1} \mu$ (parámetro de no centralidad), entonces,

$$T^2 = n\mathbf{x}^T \mathbf{M}^{-1} \mathbf{x} \sim \left(\frac{np}{n-p+1} \right) F_{p,n-p+1,\delta} = T^2_{p,n,\delta}$$

Estimación para la distribución normal multivariada

Estimación

Función de verosimilitud

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n \sim N_p(\mu, \Sigma)$ (iid) , entonces la log-verosimilitud

$$\log(L(\mu, \Sigma)) = -\frac{n}{2}\log(|2\pi\Sigma|) - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \mu)^T\Sigma^{-1}(\mathbf{x}_i - \mu)$$

Proposición 7

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n \sim N_p(\mu, \Sigma)$ (iid), con $n \geq p + 1$, entonces los estimadores máximo verosímiles están dados por

$$\hat{\mu} = \bar{x} \qquad \hat{\Sigma} = \frac{(n-1)}{n}\mathbf{S}$$

Estimación

Teorema 7

Sean $\bar{\mathbf{x}}$ y $\mathbf{S}$ la media y matriz de varianzas muestrales de una distribución normal multivariada $N_p(\mu, \Sigma)$ con $(n-1) \geq p$ entonces,

1. $\bar{\mathbf{x}} \sim N_p(\mu, n^{-1}\Sigma)$

2. $(n-1)\mathbf{S} \sim W_p(n-1, \Sigma)$

3. $\bar{\mathbf{x}}$ y $\mathbf{S}$ son independientes

4. $n(\bar{\mathbf{x}} - \mu)^T \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu) \sim T^2(p, n-1)$