

José Antonio Perusquía Cortés

Análisis Multivariado, Semestre 2025 - Il



Definición 1

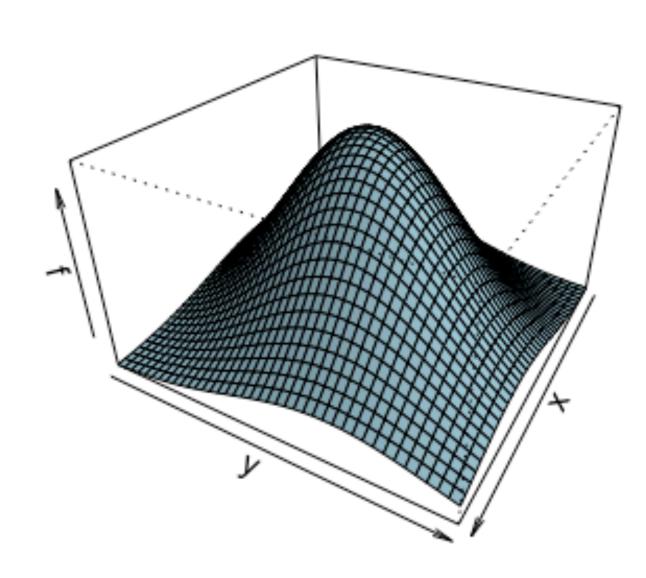
Se dice que ${\bf x}$ sigue una distribución normal multivariada p-variada no singular, denotado por ${\bf x} \sim N_p\left(\mu,\Sigma\right)$ si ${\bf x}$ tiene función de densidad

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

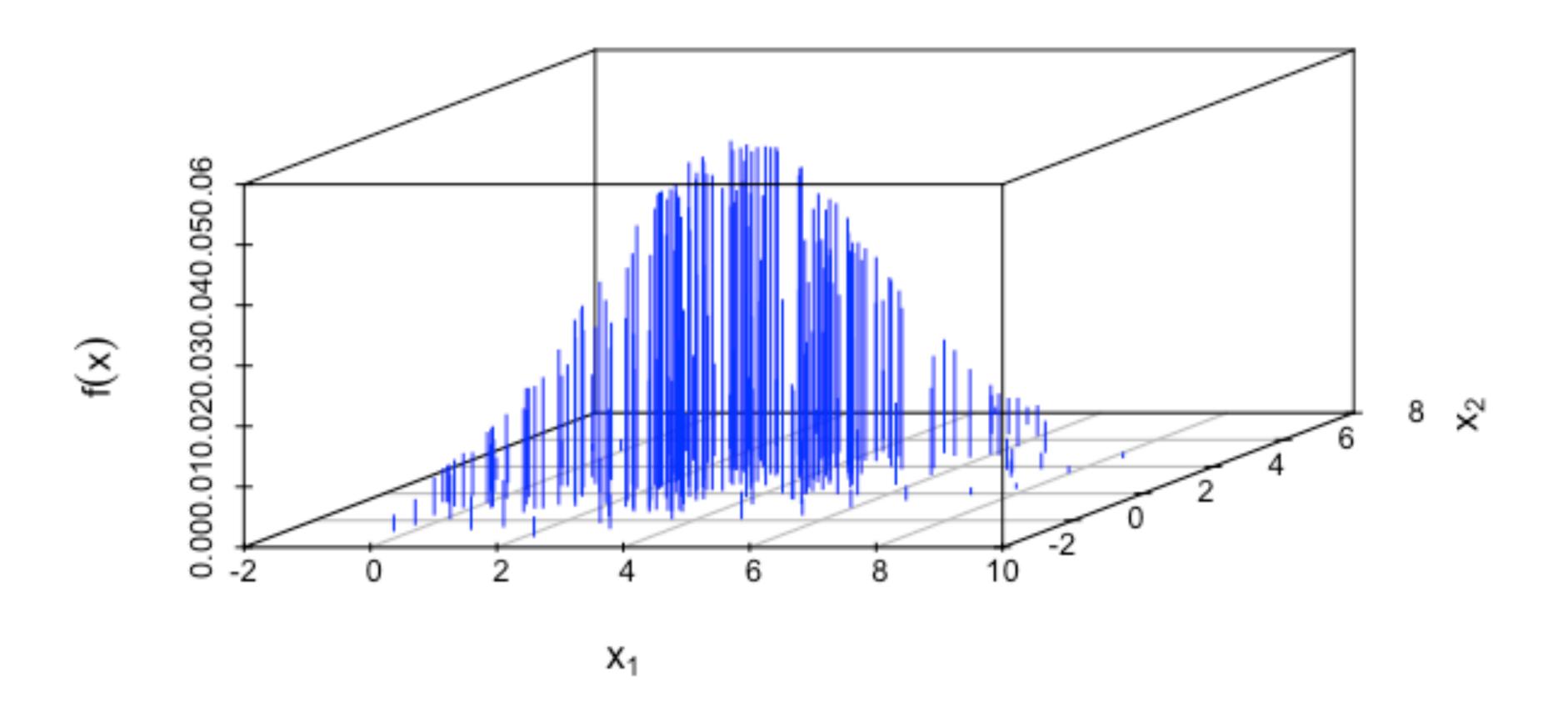
donde

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

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



Para datos bivariados también se puede crear un scatterplot en 3D con librería scatterplot3d



Observación 1

Si $ran(\Sigma) = k < p$ podemos definir 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

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

Caracterizaciones

Definición 2

Decimos que \mathbf{x} tiene 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 y definamos a $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ donde \mathbf{A} es una matriz de dimensión $q \times p$. Entonces \mathbf{y} tiene una distribución normal q-variada tal que:

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

Caracterizaciones

Corolario 1

Sea
$$\mathbf{x} \sim N_p(\mathbf{0}, \mathbf{I}_p)$$
 y definamos a $\mathbf{y} = \Sigma^{\frac{1}{2}}\mathbf{x} + \mu$, entonces, $\mathbf{y} \sim N_p(\mu, \Sigma)$

Corolario 2

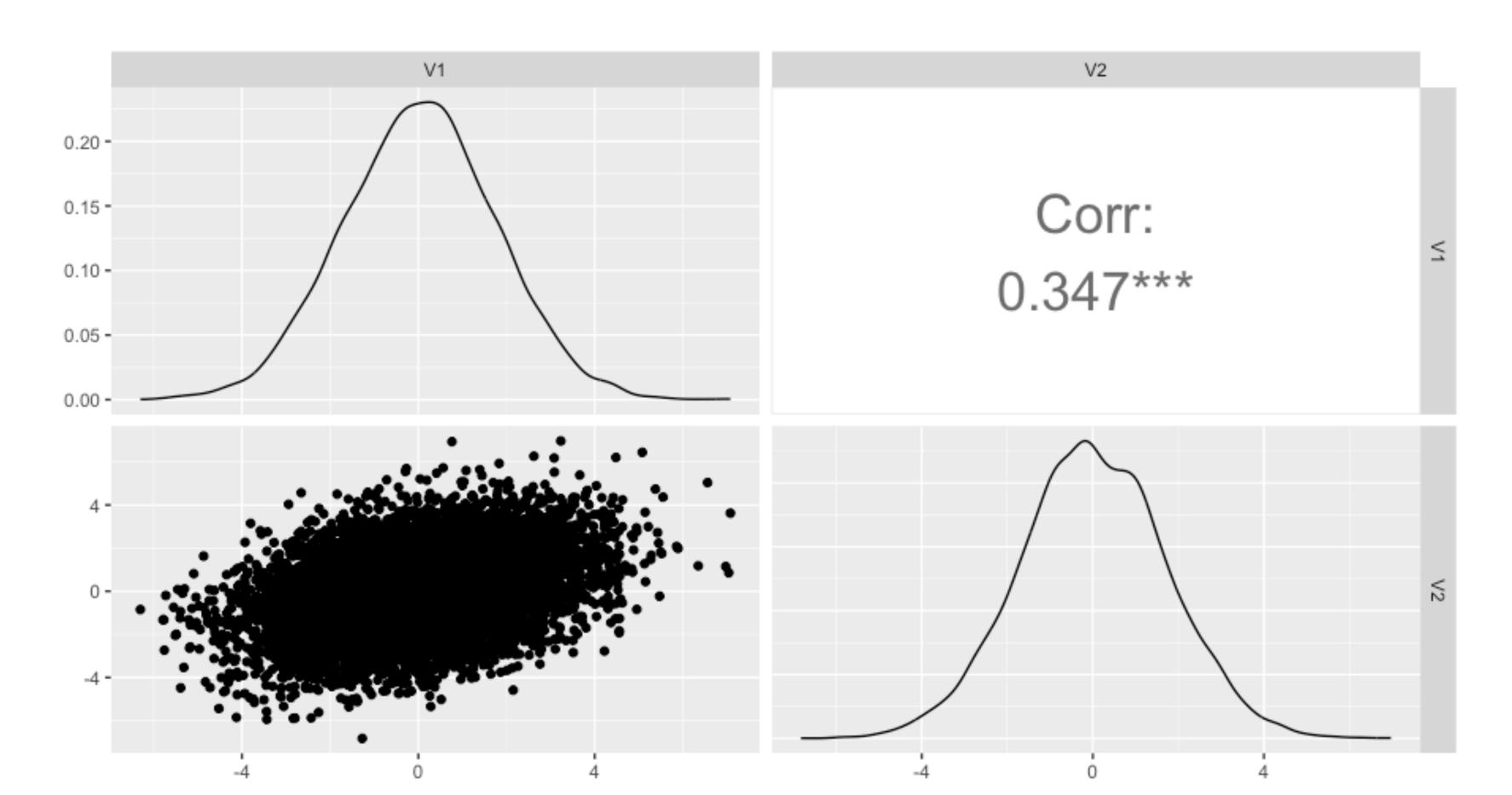
Sea $\mathbf{x} \sim N_p(\mu, \Sigma)$ con $\Sigma > 0$ y definamos a $\mathbf{y} = \Sigma^{-\frac{1}{2}}(\mathbf{x} - \mu)$, donde $\Sigma^{-\frac{1}{2}}$ es la matriz raíz cuadrada de Σ^{-1} . Entonces, y_1, y_2, \ldots, y_p son variables aleatorias iid N(0,1).

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

Simulaciones

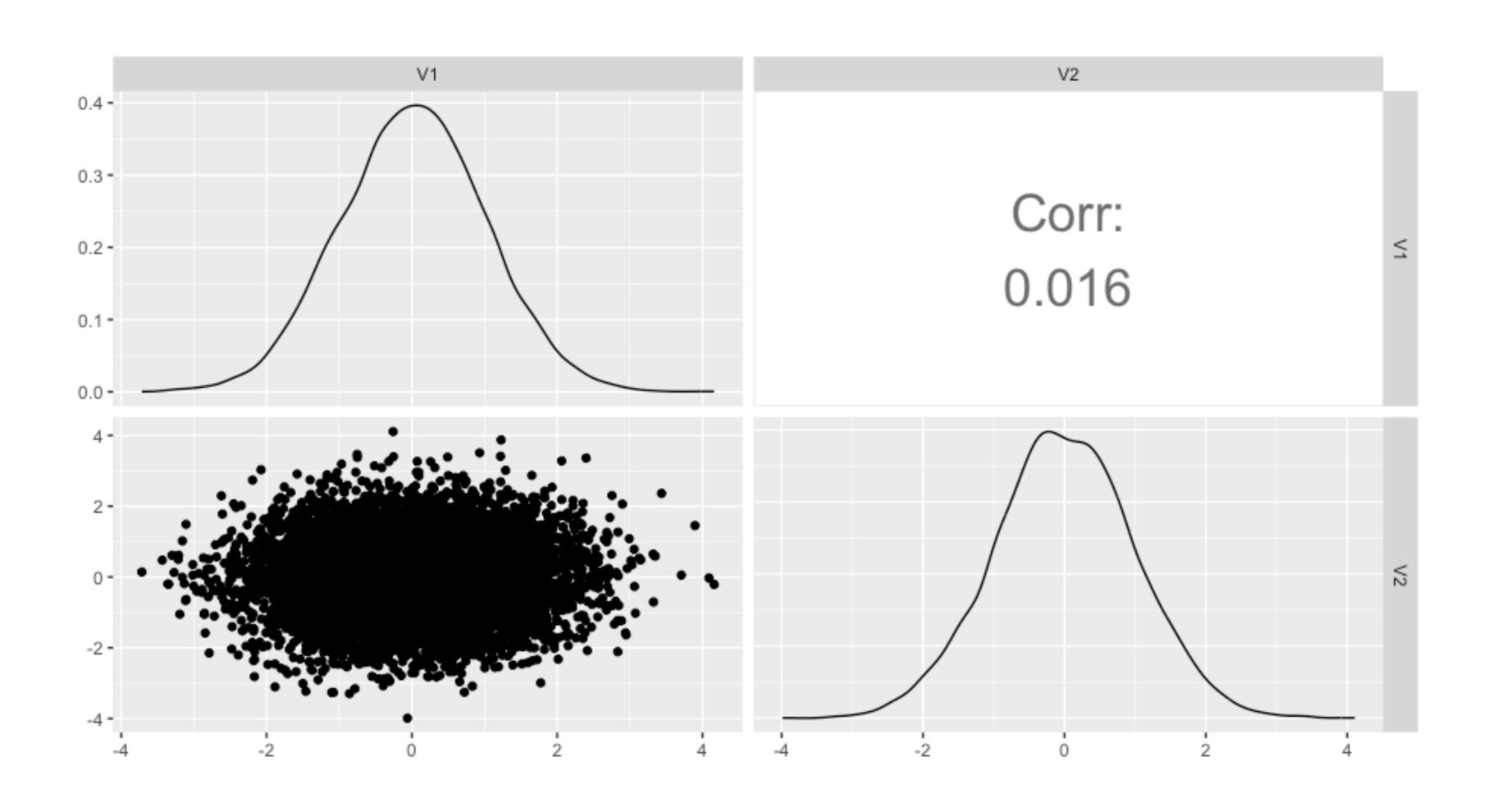
- ¿Cómo se ven las densidades univariadas y las curvas de nivel de una simulación de ${\bf x}\sim N_2(\mu,\Sigma)$? Donde

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



Simulaciones

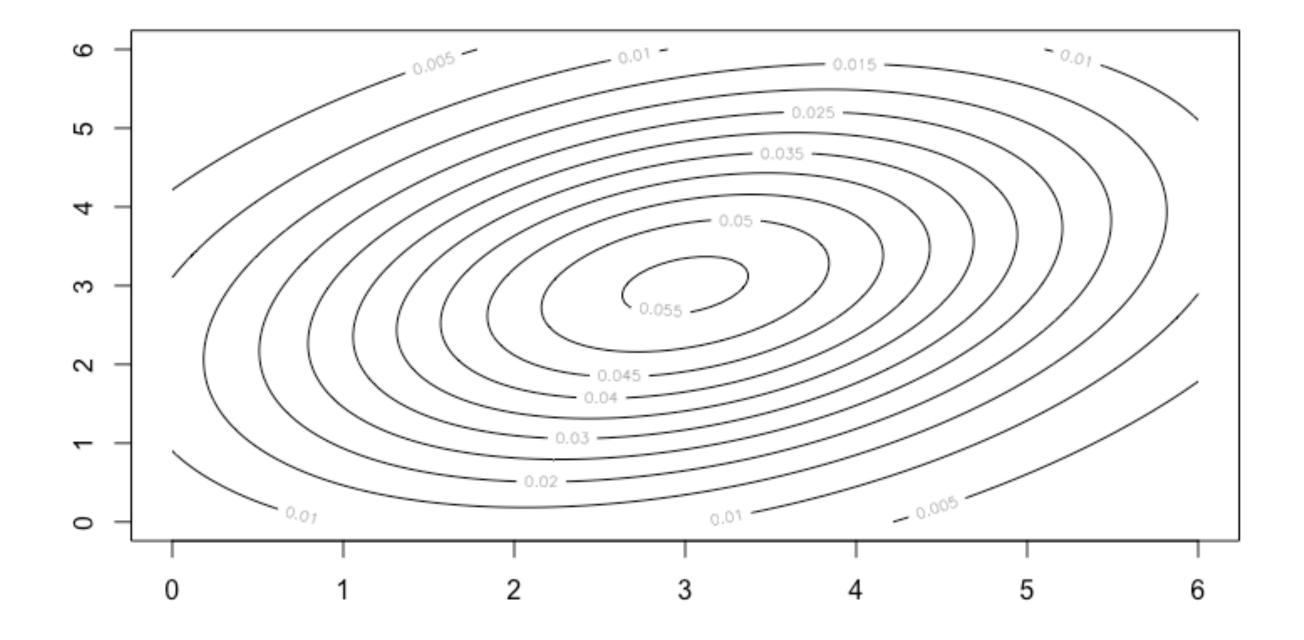
• ¿Qué pasa si hacemos $\mathbf{y} = \Sigma^{-\frac{1}{2}}(\mathbf{x} - \mu)$?



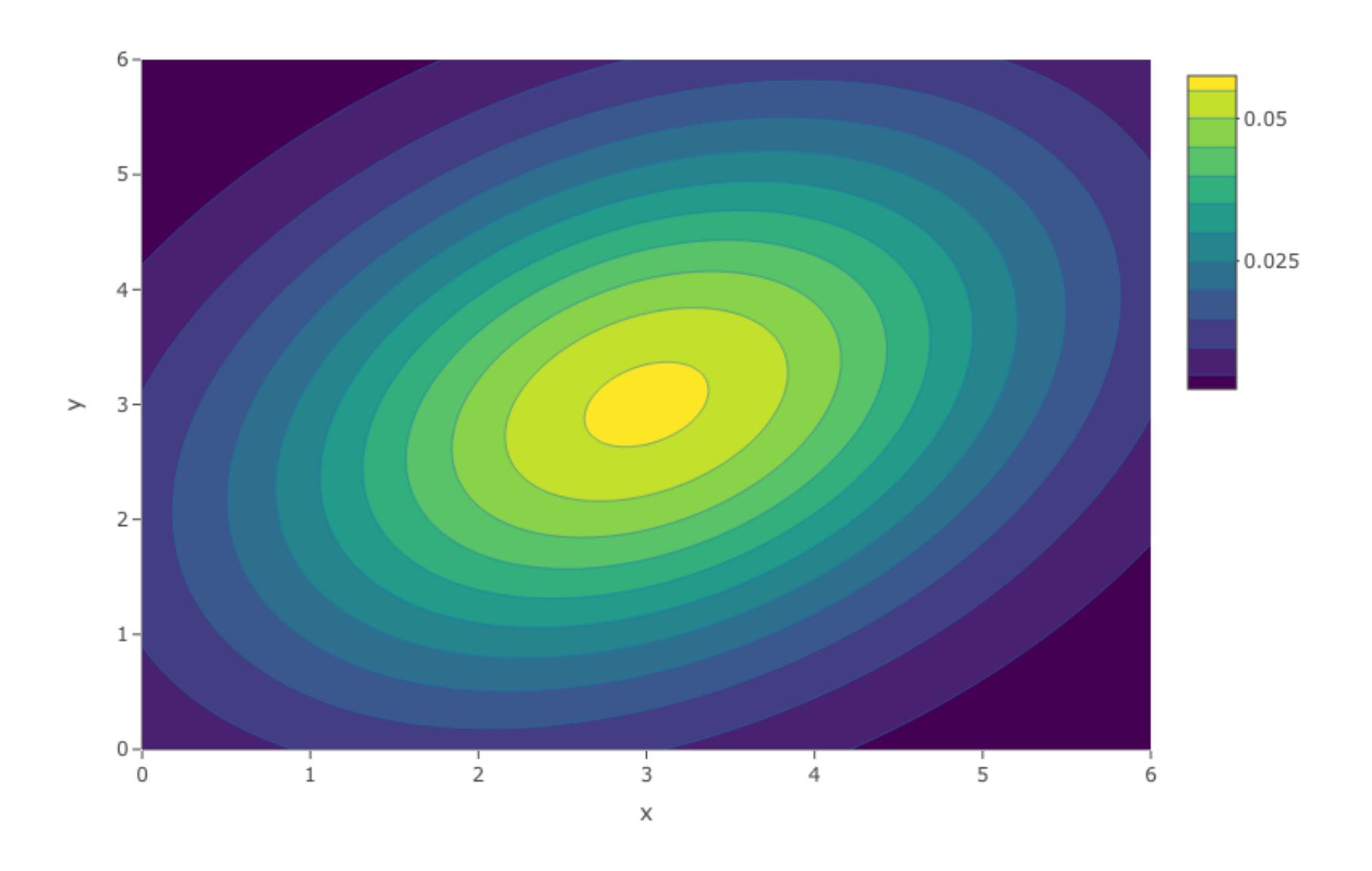
Observación 2

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

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



► En **R**: la librería plotly para una gráfica más interactiva



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

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

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

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,

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

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

- 2. $\mathbf{x}^{(1)}$ y $\mathbf{x}^{(2)}$ son independientes si y solo si Cov $\left(\mathbf{x}^{(1)},\mathbf{x}^{(2)}\right)=\Sigma_{12}=\mathbf{0}$
- 3. $\mathbf{x}^T \mathbf{\Sigma}^{-1} \mathbf{x} \sim \chi_p^2 \left(\mu^T \mathbf{\Sigma}^{-1} \mu \right)$
- 4. $\mathbf{x}^{(2)} | \mathbf{x}^{(1)} \sim N_{p-k} \left(\mu^{(2)} + \Sigma_{21} \Sigma_{11}^{-1} [\mathbf{x}^{(1)} \mu^{(1)}], \Sigma_{22} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \right)$

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

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

Teorema 2 (Teorema de Cramér-Wold)

Para $\mathbf{x}_n = (x_{n1}, ..., x_{np})$ y $\mathbf{x} = (x_1, ..., x_p)$ 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:

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

donde

- etr es el operador exp^{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 $\mathbf{x}_1, ..., \mathbf{x}_n$ 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 $n \ge p$, entonces se puede probar que $\mathbb{P}(\mathbf{M} > 0) = 1$. De lo contrario, se tiene que $\mathbf{M} \ge 0$, por lo que la densidad no existe y se dice que \mathbf{M} tiene una distribución singular

Teorema 3

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

Corolario 3

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

Corolario 4

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

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 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 ran $(\mathbf{A}) = q$, entonces $\left(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^{\mathsf{T}}\right)^{-1} \sim W_q\left(n-p+q,\left(\mathbf{A}\mathbf{\Sigma}^{-1}\mathbf{A}^{\mathsf{T}}\right)^{-1}\right)$
- 2. Sea $\mathbf{y}_{p \times 1}$ independiente de \mathbf{M} y tal que $\mathbb{P}(\mathbf{y} = \mathbf{0}) = 0$, entonces $\frac{\mathbf{y}^{\mathsf{T}} \mathbf{M} \mathbf{y}}{\mathbf{y}^{\mathsf{T}} \mathbf{\Sigma} \mathbf{y}} \sim \chi_n^2 \, \mathbf{y} \, \frac{\mathbf{y}^{\mathsf{T}} \mathbf{\Sigma}^{-1} \mathbf{y}}{\mathbf{y}^{\mathsf{T}} \mathbf{M}^{-1} \mathbf{y}} \sim \chi_{n-p+1}^2$
- 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
 - $\mathbf{X}^{\mathbf{T}}\mathbf{A}\mathbf{X} \sim W_p(r, \Sigma)$ si y solo si $\mathbf{y}^{\mathbf{T}}\mathbf{A}\mathbf{y} \sim \sigma_a^2 \cdot \chi_r^2$
 - $\mathbf{X^TAX} \sim W_p(r, \Sigma) \mathbf{y} \mathbf{X^TBX} \sim W_p(s, \Sigma)$ son ind. sii $\mathbf{y^TAy} \sim \sigma_a^2 \cdot \chi_r^2 \mathbf{y} \mathbf{y^TBy} \sim \sigma_a^2 \cdot \chi_r^2$ son ind.
 - $\mathbf{X^Tb} \sim N_p \mathbf{y} \mathbf{X^TAX} \sim W_p(r, \Sigma)$ son ind. $\sin \mathbf{y^Tb} \sim N_1 \mathbf{y} \mathbf{y^TAy} \sim \sigma_a^2 \cdot \chi_r^2$ son ind.

Formas cuadráticas

Lema 1

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

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

- 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 $\{\mathbf{a}_1, ..., \mathbf{a}_r\}$, $r \le n$, 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 $\mathbf{X}\mathbf{b} \sim N_n\left(\mathbf{0}, \sigma_b^2 \mathbf{I}\right)$ donde $\sigma_b^2 = \mathbf{b}^T \Sigma \mathbf{b}$

Formas cuadráticas

Lema 2

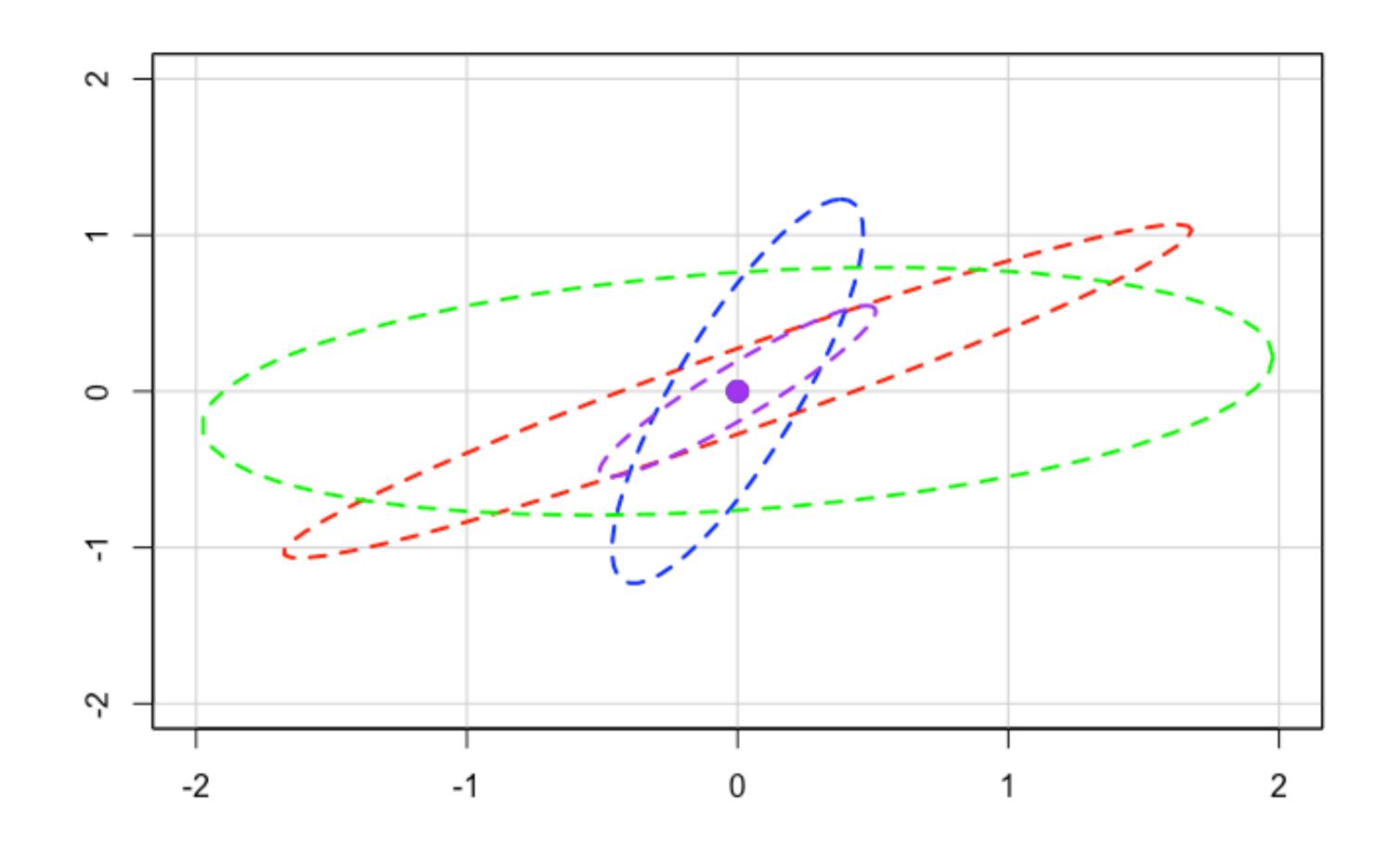
Sea $\mathbf{x} \sim N_p\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$ y $\mathbf{A}_{p\times q}$ una matriz simétrica entonces $\mathbf{x}^T\mathbf{A}\mathbf{x} \sim \sigma^2 \cdot \chi_r^2$ si y solo si \mathbf{A} es idempotente y con ran $(\mathbf{A}) = r$

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 $\mathbf{P}_1 \mathbf{P}_2 = \mathbf{O}$

Simulaciones

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



Distribución Wishart no centrada

Definición 5 (Distribución Wishart no centrada)

Sean $\mathbf{x}_1, \dots, \mathbf{x}_n$ 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

$$-\Lambda = (\mu_1, \dots, \mu_n)^T$$

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,

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

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,

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

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,

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



Estimación

Función de verosimilitud

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

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

Estimación

Teorema 7

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

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

2.
$$(n-1)$$
S ~ $W_p(n-1,\Sigma)$

3. x y S son independientes

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

Prueba de hipótesis para µ

Prueba para μ (Σ conocida)

• Sean $\mathbf{x}_1, \dots, \mathbf{x}_n \sim N_p(\mu, \Sigma)$ (iid) queremos hacer el siguiente contraste

$$H_0: \mu = \mu_0$$
 vs $H_a: \mu \neq \mu_0$

Usamos el estadístico de prueba

$$\xi^2 = n(\bar{\mathbf{x}} - \mu_0)^T \Sigma^{-1}(\bar{\mathbf{x}} - \mu_0)$$

- Bajo H_0 se tiene que $\xi^2 \sim \chi_p^2$
- Región de confianza $100(1-\alpha)\%$ son las elipsoides

$$\left\{\mathbf{x}: \boldsymbol{\xi}^2 \leq \chi_{p,1-\alpha}^2\right\}$$

Prueba para μ (Σ conocida)

Ejemplo

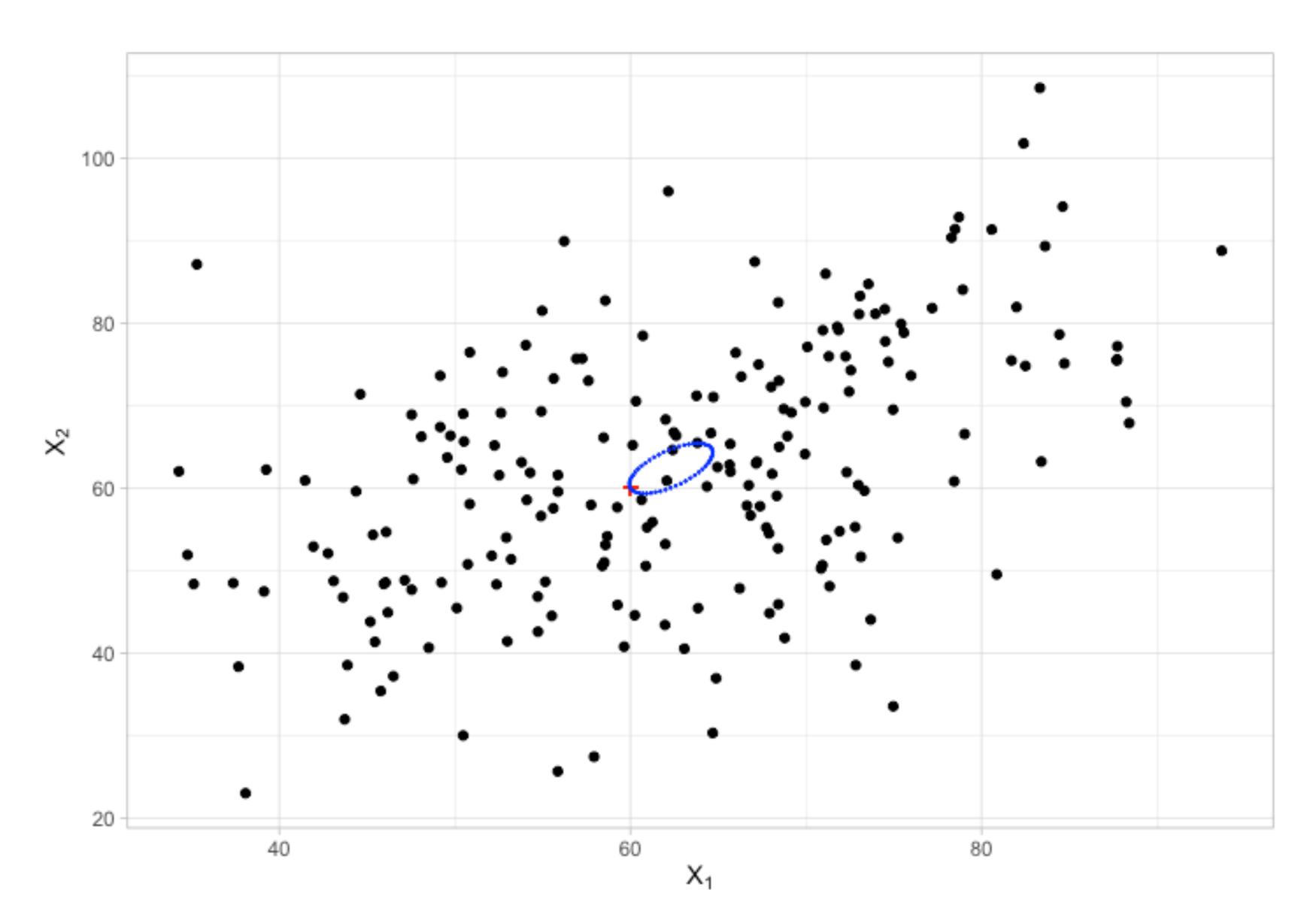
Dados $\mathbf{x}_1, \dots, \mathbf{x}_{203} \sim N_2(\mu, \Sigma)$ (iid) con

$$\mu = \begin{pmatrix} 64.1 \\ 64.7 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} 191 & 155.6 \\ 155.6 & 313.5 \end{pmatrix}.$$

Se busca contrastar,

$$H_0: \mu = 60$$
 vs $H_a: \mu \neq 60$

Prueba para μ (Σ conocida)



$$\xi^2 = 5.971581 < 5.991465 = \chi^2_{2,.95}$$

No rechazamos H_0

Prueba para μ (Σ desconocida)

▶ Para una muestra, utilizamos S para construir el estadístico de prueba

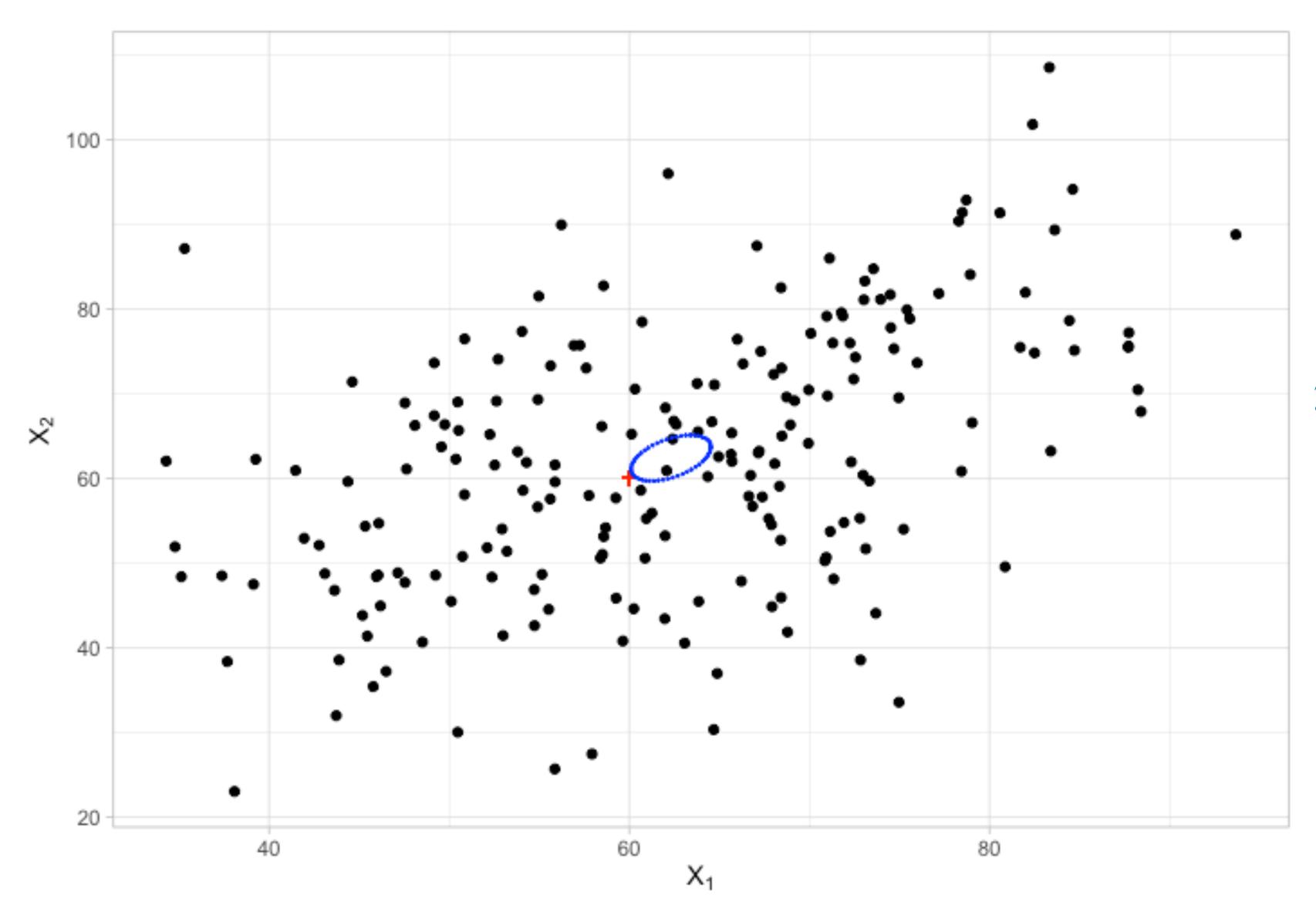
$$\gamma^2 = \frac{n(n-p)}{(n-1)p} (\bar{\mathbf{x}} - \mu_0)^T \mathbf{S}^{-1} (\bar{\mathbf{x}} - \mu_0)$$

• Bajo H_0 se tiene $\gamma^2 \sim F_{p,n-p}$

- Región de confianza $100(1-\alpha)$ % son las elipsoides

$$\left\{\mathbf{x}: \gamma^2 \le F_{p,n-p,1-\alpha}\right\}$$

Prueba para μ (Σ desconocida)



 $\gamma^2 = 3.870381 > 3.013826 = F_{2,201,.95}$ Rechazamos H_0