Tema 3

Normalidad multivariante

Aurea Grané
Departamento de Estadística
Universidad Carlos III de Madrid

3 Normalidad multivariante

1. Distribuciones de probabilidad multivariantes
2. La distribución normal multivariante
3. Distribuciones relacionadas con la distribución normal multivariante
4. Aplicación a la inferencia multivariante

3.1 Distribuciones de probabilidad multivariantes

Funciones de distribución y de densidad conjuntas

Sea $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$ un vector aleatorio $p \times 1$.

Se define la **función de distribución conjunta** de $\mathbf{X}$ (o función de probabilidad acumulada) en el punto $\mathbf{x} = (x_1, x_2, \ldots, x_p)' \in \mathbb{R}^p$ como

$$F(\mathbf{x}) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_p} f(\mathbf{t}) dt_1 dt_2 \ldots dt_p,$$

donde $\mathbf{t} = (t_1, t_2, \ldots, t_p) \in \mathbb{R}^p$.

La función $f(\mathbf{t})$ se denomina **función de densidad conjunta** de $\mathbf{X}$ y verifica que:

- $f(\mathbf{t}) = f(t_1, t_2, \ldots, t_p) \geq 0$, $\forall \mathbf{t} \in \mathbb{R}^p$,
- $\int_{\mathbb{R}^p} f(\mathbf{t}) dt_1 dt_2 \ldots dt_p = 1$.

Distribuciones marginales

Dado el vector aleatorio $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$, la **función de densidad marginal** de la variable aleatoria X_j, para $j = 1, \ldots, p$, se obtiene integrando la densidad conjunta:

$$f_{X_j}(x_j) = \int_{\mathbb{R}^{p-1}} f(t_1, t_2, \ldots, t_p) dt_1 \ldots dt_{j-1} dt_{j+1} \ldots dt_p.$$

La función $f_{X_j} : \mathbb{R} \to [0, +\infty)$ y cumple que $\int_{\mathbb{R}} f_{X_j}(t) dt = 1$.

De forma más general, dada una partición del vector $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, donde $\mathbf{X}_1 = (X_1, \ldots, X_{p_1})'$ y $\mathbf{X}_2 = (X_{p_1+1}, \ldots, X_p)'$, la **función de densidad marginal** de $\mathbf{X}_1$ se obtiene:

$$f_{\mathbf{X}_1}(x_1, \ldots, x_{p_1}) = \int_{\mathbb{R}^{p_2}} f(t_1, t_2, \ldots, t_p) dt_{p_1+1} \ldots dt_p, \quad \text{donde } p_2 = p - p_1.$$

La función $f_{\mathbf{X}_1} : \mathbb{R}^{p_1} \to [0, +\infty)$ y $\int_{\mathbb{R}^{p_1}} f_{\mathbf{X}_1}(t_1, \ldots, t_{p_1}) dt_1 \ldots dt_{p_1} = 1$.

Análogamente se obtiene la función de densidad marginal de $\mathbf{X}_2$, dando lugar a una función $f_{\mathbf{X}_2} : \mathbb{R}^{p_2} \to [0, +\infty)$.

Distribuciones condicionadas

Consideremos la misma partición del vector $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, donde $\mathbf{X}_1 = (X_1, \ldots, X_{p_1})'$, $\mathbf{X}_2 = (X_{p_1+1}, \ldots, X_p)'$ y $p_2 = p - p_1$.

Supongamos que se conoce un valor determinado del vector $\mathbf{X}_2$, es decir, $\mathbf{X}_2 = \mathbf{x}_2^0 = (x_{p_1+1}^0, \ldots, x_p^0) \in \mathbb{R}^{p_2}$. Entonces, la **función de densidad** de $\mathbf{X}_1$ **condicionada** a que $\mathbf{X}_2 = \mathbf{x}_2^0$ se obtiene como

$$f_{\mathbf{X}_1|\mathbf{X}_2=\mathbf{x}_2^0}(\mathbf{x}_1) = \frac{f(\mathbf{x}_1, \mathbf{x}_2^0)}{f_{\mathbf{X}_2}(\mathbf{x}_2^0)} = \frac{f(x_1, \ldots, x_{p_1}, x_{p_1+1}^0, \ldots, x_p^0)}{f_{\mathbf{X}_2}(x_{p_1+1}^0, \ldots, x_p^0)},$$

que da lugar a una función de $\mathbb{R}^{p_1}$ en $[0, +\infty)$.

Esperanza, covarianza y correlación

Sea $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$ un vector aleatorio con función de densidad conjunta $f(\mathbf{x})$.

La **esperanza** (o primer momento) de $\mathbf{X}$ se define como el vector de esperanzas de las variables aleatorias que lo forman, es decir,

$$\boldsymbol{\mu} = \mathsf{E}(\mathbf{X}) = (\mathsf{E}(X_1), \mathsf{E}(X_2), \ldots, \mathsf{E}(X_p))' \in \mathbb{R}^p,$$

donde $\mathsf{E}(X_j) = \int_{\mathbb{R}} t\, f_{X_j}(t) dt$, para $j = 1, \ldots, p$.

El segundo momento de $\mathbf{X}$ se define como la matriz que contiene los segundos momentos de las variables aleatorias que forman el vector

$$\boldsymbol{\mu}_2 = \mathsf{E}(\mathbf{X}\,\mathbf{X}') = \begin{pmatrix} \mathsf{E}(X_1^2) & \mathsf{E}(X_1X_2) & \ldots & \mathsf{E}(X_1X_p) \\ \mathsf{E}(X_2X_1) & \mathsf{E}(X_2^2) & \ldots & \mathsf{E}(X_2X_p) \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{E}(X_pX_1) & \mathsf{E}(X_pX_2) & \ldots & \mathsf{E}(X_p^2) \end{pmatrix}_{p\times p}$$

donde $\mathsf{E}(X_j^2) = \int_{\mathbb{R}} t^2 f_{X_j}(t) dt$ y $\mathsf{E}(X_jX_k) = \int_{\mathbb{R}^2} t\, s\, f_{(X_j,X_k)}(t,s)\, dt\, ds$.

La **covarianza** de $\mathbf{X}$ se define como

$$\boldsymbol{\Sigma} = \mathsf{Cov}(\mathbf{X}) = \mathsf{E}((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})') = \boldsymbol{\mu}_2 - \boldsymbol{\mu}\boldsymbol{\mu}',$$

que da lugar a la siguiente matriz de covarianzas

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \ldots & \sigma_{pp} \end{pmatrix}, \text{ con } \sigma_{jk} = \begin{cases} \mathsf{E}(X_j^2) - (\mathsf{E}(X_j))^2, & \text{si } j = k, \\ \mathsf{E}(X_jX_k) - \mathsf{E}(X_j)\,\mathsf{E}(X_k), & \text{si } j \neq k. \end{cases}$$

Observación 1 *La diagonal de $\boldsymbol{\Sigma}$ contiene las varianzas de las X_j's, mientras que fuera de la diagonal están las covarianzas de los pares de variables. Para las varianzas también se utiliza la siguiente notación alternativa $\sigma_{jj} = \sigma_j^2 = \mathsf{var}(X_j)$. Además hay que recordar que $\mathsf{var}(X_j) = \mathsf{cov}(X_j, X_j)$.*

La matriz $\boldsymbol{\Sigma}$ es simétrica y semidefinida positiva.

La **correlación** de $\mathbf{X}$ se define a partir de la matriz de covarianzas

$$\mathbf{P} = \mathsf{Corr}(\mathbf{X}) = \mathbf{D}_\sigma^{-1}\,\boldsymbol{\Sigma}\,\mathbf{D}_\sigma^{-1},$$

donde $\mathbf{D}_\sigma = \mathsf{diag}(\sqrt{\sigma_{11}}, \ldots, \sqrt{\sigma_{pp}}) = \mathsf{diag}(\sigma_1, \ldots, \sigma_p)$, dando lugar a la siguiente matriz de correlaciones

$$\mathbf{P} = \begin{pmatrix} 1 & \rho_{12} & \ldots & \rho_{1p} \\ \rho_{21} & 1 & \ldots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \ldots & 1 \end{pmatrix}, \text{ con } \rho_{jk} = \mathsf{corr}(X_j, X_k) = \frac{\sigma_{jk}}{\sigma_j\sigma_k}.$$

La matriz $\mathbf{P}$ es simétrica y semidefinida positiva.

Variables aleatorias (compuestas)

Sea $\mathbf{X}$ un vector aleatorio con esperanza $\boldsymbol{\mu}$ y covarianza $\boldsymbol{\Sigma}$.

Consideremos la variable aleatoria (compuesta) $Y = \mathbf{a}'\mathbf{X}$, donde $\mathbf{a} \in \mathbb{R}^p$, es decir, Y es una combinación lineal de las componentes del vector aleatorio $\mathbf{X}$.

La esperanza y varianza de Y son

$$\begin{aligned} \mathsf{E}(Y) &= \mathsf{E}(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\,\mathsf{E}(\mathbf{X}) = \mathbf{a}'\boldsymbol{\mu} \in \mathbb{R}, \\ \mathsf{var}(Y) &= \mathsf{var}(\mathbf{a}'\mathbf{X}) = \mathsf{cov}(\mathbf{a}'\mathbf{X}, \mathbf{a}'\mathbf{X}) = \mathsf{E}((\mathbf{a}'\mathbf{X} - \mathsf{E}(\mathbf{a}'\mathbf{X}))(\mathbf{a}'\mathbf{X} - \mathsf{E}(\mathbf{a}'\mathbf{X}))') \\ &= \mathsf{E}((\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu})(\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu})') = \mathbf{a}'\,\mathsf{E}((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})')\mathbf{a} = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}. \end{aligned}$$

Si $Z = \mathbf{b}'\mathbf{X}$ es otra variable compuesta, con $\mathbf{b} \in \mathbb{R}^p$, entonces la covarianza entre Y y Z es

$$\begin{aligned} \mathsf{cov}(Y, Z) &= \mathsf{cov}(\mathbf{a}'\mathbf{X}, \mathbf{b}'\mathbf{X}) = \mathsf{E}((\mathbf{a}'\mathbf{X} - \mathsf{E}(\mathbf{a}'\mathbf{X}))(\mathbf{b}'\mathbf{X} - \mathsf{E}(\mathbf{b}'\mathbf{X}))') \\ &= \mathsf{E}((\mathbf{a}'\mathbf{X} - \mathbf{a}'\boldsymbol{\mu})(\mathbf{b}'\mathbf{X} - \mathbf{b}'\boldsymbol{\mu})') \\ &= \mathbf{a}'\,\mathsf{E}((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})')\mathbf{b} = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{b} = \mathbf{b}'\boldsymbol{\Sigma}\mathbf{a}. \end{aligned}$$

3.2 La distribución normal multivariante

Toda variable aleatoria $X \sim N(\mu, \sigma)$ puede expresarse como $X = \mu + \sigma Z$, donde $Z \sim N(0, 1)$. La generalización de esta propiedad en dimensión $p > 1$ permite definir la distribución normal multivariante.

Se dice que $\mathbf{X} = (X_1, X_2)'$ tiene ley **normal bivariante** con parámetros $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ y $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}$, y se denotará como $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, si

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}$$

donde Z_1, Z_2 son variables aleatorias independientes con ley $N(0, 1)$. Los coeficientes a_{ij}(para $j = 1, 2$) vienen determinados por los parámetros $\sigma_1^2 = \mathsf{var}(X_1)$, $\sigma_2^2 = \mathsf{var}(X_2)$, $\rho = \mathsf{corr}(X_1, X_2)$.

La función de densidad de probabilidad de $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ en el punto $(x_1, x_2) \in \mathbb{R}^2$ es

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}Q(x_1, x_2)\right\},$$

donde $Q(x_1, x_2)$ es la forma cuadrática siguiente:

$$Q(x_1, x_2) = \frac{1}{1-\rho^2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)$$

Observación 2 *Si $\rho = 0$ (X_1 y X_2 están incorreladas), entonces*

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right\} \\ &= \frac{1}{2\pi\sigma_1} \exp\left\{-\frac{1}{2}\frac{(x_1-\mu_1)^2}{\sigma_1^2}\right\} \frac{1}{2\pi\sigma_2} \exp\left\{-\frac{1}{2}\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right\} \\ &= f_{X_1}(x_1)\, f_{X_2}(x_2), \end{aligned}$$

donde $f_{X_1}(x_1)$ y $f_{X_2}(x_2)$ son las funciones de densidad marginales de dos v.a. con leyes $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$, respectivamente.

En la Observación anterior hemos demostrado que si $\mathbf{X} = (X_1, X_2)' \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $\rho = \mathsf{corr}(X_1, X_2) = 0$, entonces las variables aleatorias X_1 y X_2 son **normales** e **independientes**. (Recordad que esto último se debe a que la función de densidad conjunta puede expresarse como el producto de las funciones de densidad marginales).

La implicación contraria siempre es cierta. Es decir, si X_1 y X_2 son variables aleatorias independientes, entonces $\mathsf{corr}(X_1, X_2) = 0$. (Recordad que $\mathsf{corr}(X_1, X_2) = \sigma_{12}/(\sigma_1\sigma_2)$ donde $\sigma_{12} = \mathsf{E}(X_1X_2) - \mathsf{E}(X_1)\,\mathsf{E}(X_2)$. Cuando X_1 y X_2 son v.a. independientes se cumple que $\mathsf{E}(X_1X_2) = \mathsf{E}(X_1)\,\mathsf{E}(X_2)$).

Acabamos de demostrar (para el caso $p = 2$) que **bajo el modelo de normalidad multivariante, incorrelación equivale a independencia**.

Generalización al caso multivariante

Se dice que el vector $\mathbf{X} = (X_1, \dots, X_p)'$ tiene ley **normal p-variante** si existen p variables aleatorias independientes con ley $N(0,1)$, $Z_1, Z_2, \dots, Z_p$, tales que

$$\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\,\mathbf{Z},$$

donde $\mathbf{Z} = (Z_1, \dots, Z_p)'$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p)'$ y $\mathbf{A}$ es una matrix $p \times p$.

Propiedades:

1. La esperanza y covarianza del vector $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\,\mathbf{Z}$ son

$$\mathsf{E}(\mathbf{X}) = \boldsymbol{\mu}, \quad \mathsf{Var}(\mathbf{X}) = \mathbf{A}'\,\mathbf{A} = \boldsymbol{\Sigma}.$$

Utilizaremos la notación $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Observad que $\boldsymbol{\Sigma} \geq 0$. Además, $\boldsymbol{\Sigma} > 0$ si $\mathsf{rang}(\mathbf{A}) = p$.

Observación 3 *Esta propiedad se utiliza para generar muestras de un vector aleatorio $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} = \mathbf{A}'\mathbf{A}$.*

2. Si $\mathbf{X}$ es no singular, su función de densidad de probabilidad en el punto $\mathbf{x} \in \mathbb{R}^p$ es

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

donde $(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ es el cuadrado de la distancia de Mahalanobis del punto $\mathbf{x}$ al centro de la distribución $\boldsymbol{\mu}$.

3. Todo vector $\mathbf{Y}$ obtenido como cominación lineal de $\mathbf{X}$ tiene ley normal multivariante. En particular, todo subconjunto de $\mathbf{X}$ tiene ley normal multivariante y las marginales $X_1, \dots, X_p$ son normales univariantes.

4. Para variables aleatorias con distribución conjunta normal multivariante, incorrelación equivale a independencia.

5. Si $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ es no singular, la variable aleatoria

$$U_p = (\mathbf{X}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2_p.$$

6. Consideremos la siguiente partición del vector $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$, donde $\mathbf{X}_1 = (X_1, \dots, X_{p_1})'$, $\mathbf{X}_2 = (X_{p_1+1}, \dots, X_p)'$ y $p_2 = p - p_1$. Sean $\boldsymbol{\mu}_1 = \mathsf{E}(\mathbf{X}_1)$ y $\boldsymbol{\mu}_2 = \mathsf{E}(\mathbf{X}_2)$ los vectores de esperanzas. La matriz de covarianzas de $\mathbf{X}$ puede particionarse en bloques como:

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array}\right),$$

donde $\boldsymbol{\Sigma}_{11} = \mathsf{Var}(\mathbf{X}_1)$, $p_1 \times p_1$, $\boldsymbol{\Sigma}_{22} = \mathsf{Var}(\mathbf{X}_2)$, $p_2 \times p_2$, $\boldsymbol{\Sigma}_{12} = \mathsf{Cov}(\mathbf{X}_1, \mathbf{X}_2)$, $p_1 \times p_2$, y $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}'$.

La **distribución** de $\mathbf{X}_1$ **condicionada** a $\mathbf{X}_2 = \mathbf{x}_2^0$ es normal con vector de esperanzas $\mathsf{E}(\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2^0) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2^0 - \boldsymbol{\mu}_2)$ y matriz de covarianzas $Var(\mathbf{X}_1|\mathbf{X}_2 = \mathbf{x}_2^0) = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$.

Ejemplo 3.1 *Sea $\mathbf{X} = (X_1, X_2, X_3)'$ un vector aleatorio con ley normal de media $\boldsymbol{\mu} = (-1, 1, 0)'$ y matriz de covarianzas*

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

a) Encontrar la ley de la v.a. $Y = X_1 + 2\,X_2 - 3\,X_3$

b) Encontrar un vector $\mathbf{a}$, 2×1, tal que las v.a. X_1 y $X_1 - \mathbf{a}'\binom{X_2}{X_3}$ sean independientes.

c) Calcular la distribución de X_3 condicionada a $(X_1, X_2)' = (x_1^0, x_2^0)'$.

a) La v.a. $Y = X_1 + 2\,X_2 - 3\,X_3$ puede escribirse como $Y = \mathbf{b}'\mathbf{X}$, donde $\mathbf{b} = (1, 2, -3)'$. Puesto que $\mathbf{X} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, la v.a. Y tendrá ley normal. Su esperanza y varianza son:

$$\begin{aligned} \mathsf{E}(Y) &= \mathsf{E}(\mathbf{b}'\mathbf{X}) = \mathbf{b}'\boldsymbol{\mu} = (1,2,-3)\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 1. \\ \mathsf{var}(Y) &= \mathsf{var}(\mathbf{b}'\mathbf{X}) = \mathbf{b}'\boldsymbol{\Sigma}\mathbf{b} = (1,2,-3)\begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = 13. \end{aligned}$$

b) $\mathbf{a} = (a_1, a_2)'$ tal que X_1 y $X_1 - a_1X_2 - a_2X_3$ sean independientes.

Puesto que bajo el supuesto de normalidad, independencia equivale a incorrelación, debemos encontar aquel vector $\mathbf{a}$ tal que $\mathsf{cov}(X_1, X_1 - a_1X_2 - a_2X_3) = 0$.

Observemos que $X_1 - a_1X_2 - a_2X_3 = (1, -a_1, -a_2)\mathbf{X}$ y $X_1 = (1,0,0)\mathbf{X}$, por tanto:

$$0 = \mathsf{cov}(X_1, X_1 - a_1X_2 - a_2X_3) = (1,0,0)\boldsymbol{\Sigma}\begin{pmatrix} 1 \\ -a_1 \\ -a_2 \end{pmatrix} = 1 - a_2 \Rightarrow a_2 = 1.$$

Por tanto, cualquier vector $\mathbf{a} = (k, 1)$, $k \in \mathbb{R}$ hace que las variables X_1 y $X_1 - a_1X_2 - a_2X_3$ sean independientes.

c) Distribución de X_3 condicionada a $(X_1, X_2)' = (x_1^0, x_2^0)'$.

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \text{ donde } \boldsymbol{\mu} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \boldsymbol{\mu}_3 \end{pmatrix} \text{ y}$$

$$\boldsymbol{\Sigma} = \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 3 & 1 \\ \hline 1 & 1 & 2 \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{(12)(12)} & \boldsymbol{\Sigma}_{(12)3} \\ \hline \boldsymbol{\Sigma}_{3(12)} & \boldsymbol{\Sigma}_{33} \end{array}\right)$$

La distribución de X_3 condicionada a $(X_1, X_2)' = (x_1^0, x_2^0)'$ será normal univariante, de esperanza

$$\begin{aligned} & \mathsf{E}(X_3|(X_1,X_2)' = (x_1^0, x_2^0)') = \mu_3 + \boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}\begin{pmatrix} x_1^0 - \mu_1 \\ x_2^0 - \mu_2 \end{pmatrix} \\ = \; & 0 + (1,1)\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^{-1}\begin{pmatrix} x_1^0 + 1 \\ x_2^0 - 1 \end{pmatrix} = (1,1)\begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} x_1^0 + 1 \\ x_2^0 - 1 \end{pmatrix} \\ = \; & (1, 1/3)\begin{pmatrix} x_1^0 + 1 \\ x_2^0 - 1 \end{pmatrix} = x_1^0 + 1 + \frac{1}{3}x_2^0 - \frac{1}{3} = x_1^0 + \frac{1}{3}x_2^0 + \frac{2}{3}, \end{aligned}$$

y varianza:

$$\begin{aligned} & \mathsf{var}(X_3|(X_1,X_2)' = (x_1^0, x_2^0)') = \boldsymbol{\Sigma}_{33} - \boldsymbol{\Sigma}_{3(12)}\boldsymbol{\Sigma}_{(12)(12)}^{-1}\boldsymbol{\Sigma}_{(12)3} \\ = \; & 2 - (1,1)\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 - (1,1)\begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ = \; & 2 - (1, 1/3)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 - \left(1 + \frac{1}{3}\right) = \frac{2}{3}. \end{aligned}$$

Resumiendo, $X_3|(X_1, X_2)' = (x_1^0, x_2^0)' \sim N\left(x_1^0 + \frac{1}{3}x_2^0 + \frac{2}{3}, \frac{2}{3}\right)$.

Ejemplo 3.2 *Sean* $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4$ *vectores aleatorios independientes con ley* $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, *donde* $\boldsymbol{\mu} = (1, 2)'$ *y* $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.1 \\ 0.1 & 2 \end{pmatrix}$.

a) Encontrar la ley del vector aleatorio $\mathbf{Y} = \frac{1}{4}\mathbf{X}_1 - \frac{1}{4}\mathbf{X}_2 + \frac{1}{4}\mathbf{X}_3 - \frac{1}{4}\mathbf{X}_4$.

b) Escribir la funbción de densidad del vector $\mathbf{Y}$.

c) Calcular la correlación ρ *correspondiente a la matriz de covarianzas* $\boldsymbol{\Sigma}$. *Observar qué cambios se producen en la densidad de* $\mathbf{Y}$ *al variar el valor de* $\rho = \mathsf{corr}(Y_1, Y_2)$.

a) El vector $\mathbf{Y}$ tiene ley normal bivariante con vector de esperanzas

$$\begin{aligned} \mathsf{E}(\mathbf{Y}) &= \mathsf{E}\left(\frac{1}{4}\mathbf{X}_1 - \frac{1}{4}\mathbf{X}_2 + \frac{1}{4}\mathbf{X}_3 - \frac{1}{4}\mathbf{X}_4\right) \\ &= \frac{1}{4}\mathsf{E}(\mathbf{X}_1) - \frac{1}{4}\mathsf{E}(\mathbf{X}_2) + \frac{1}{4}\mathsf{E}(\mathbf{X}_3) - \frac{1}{4}\mathsf{E}(\mathbf{X}_4) \\ &= \left(\frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \frac{1}{4}\right)\boldsymbol{\mu} = 0\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}, \end{aligned}$$

y matriz de covarianzas

$$\begin{aligned} \mathsf{Var}(\mathbf{Y}) &= \mathsf{Var}\left(\frac{1}{4}\mathbf{X}_1 - \frac{1}{4}\mathbf{X}_2 + \frac{1}{4}\mathbf{X}_3 - \frac{1}{4}\mathbf{X}_4\right) \\ &= \frac{1}{16}\mathsf{Var}(\mathbf{X}_1) + \frac{1}{16}\mathsf{Var}(\mathbf{X}_2) + \frac{1}{16}\mathsf{Var}(\mathbf{X}_3) + \frac{1}{16}\mathsf{E}(\mathbf{X}_4) \\ &= \frac{4}{16}\boldsymbol{\Sigma} = \frac{1}{4}\boldsymbol{\Sigma}. \end{aligned}$$

Resumiendo, $\mathbf{Y} \sim N_2\left(\vec{0}, \frac{1}{4}\boldsymbol{\Sigma}\right)$.

b) Puesto que $\mathsf{E}(\vec{0})$, la función de densidad de $\mathbf{Y} = (Y_1, Y_2)'$ en el punto $\mathbf{y} = (y_1, y_2)' \in \mathbb{R}^2$ es

$$\begin{aligned} f(\mathbf{y}) &= \frac{|\frac{1}{4}\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{2/2}} \exp\left\{-\frac{1}{2}\mathbf{y}'\left(\frac{1}{4}\boldsymbol{\Sigma}\right)^{-1}\mathbf{y}\right\} \\ &= \frac{1}{\pi|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{2\,\mathbf{y}'\boldsymbol{\Sigma}^{-1}\mathbf{y}\right\}, \quad \mathbf{y} \in \mathbb{R}^2. \end{aligned}$$

c) La matriz de covarianzas de $\mathbf{Y}$ es $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.1 \\ 0.1 & 2 \end{pmatrix}$. Por tanto, el coeficiente de correlación será $\rho = \mathsf{corr}(Y_1, Y_2) = \frac{\sigma_{12}}{\sigma_1\sigma_2} = \frac{0.1}{\sqrt{1\cdot 2}} \approx 0.071$.

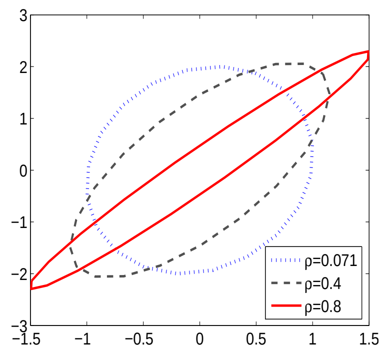
¿Cómo influye el valor de ρ en la densidad de $\mathbf{Y}$?

Puesto que $\sigma_{12} = \sigma_1\sigma_2\rho$, la matriz $\boldsymbol{\Sigma}$ puede escribirse en función de ρ como $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \sqrt{2}\rho \\ \sqrt{2}\rho & 2 \end{pmatrix}$.

En el Tema 2 vimos dos medidas de dispersión global:

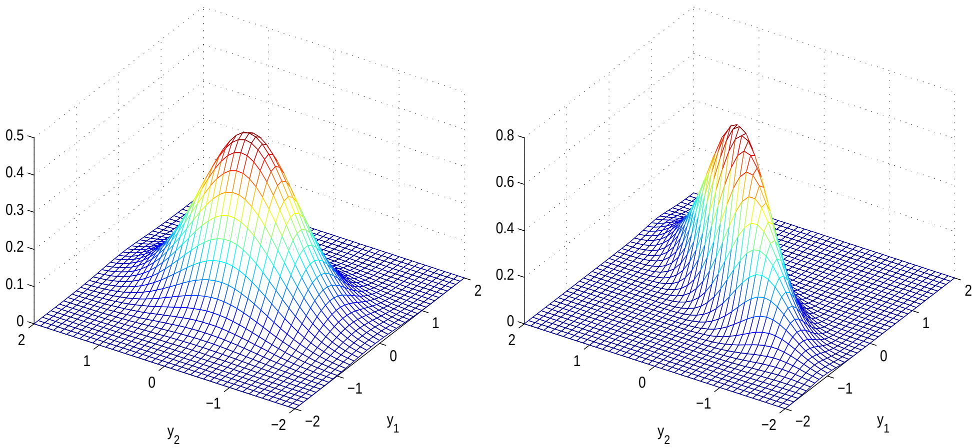
- La variación total $\mathsf{tr}(\boldsymbol{\Sigma}) = 3$, que no depende de ρ.
- La varianza generalizada $\mathsf{det}(\boldsymbol{\Sigma}) = 2(1 - \rho^2)$.
  $\rho^2 = 0 \Rightarrow \mathsf{det}(\boldsymbol{\Sigma}) = 2$
  $\rho^2 = 1 \Rightarrow \mathsf{det}(\boldsymbol{\Sigma}) = 0$.

A medida que ρ aumenta, la dispersión global disminuye.



Elipses de la forma cuadrática $\mathbf{y}\boldsymbol{\Sigma}^{-1}\mathbf{y}'$

Función de densidad de $\mathbf{Y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu} = (0,0)'$ y $\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \sqrt{2}\rho \\ \sqrt{2}\rho & 2 \end{pmatrix}$.



(a) $\rho = 0.071$ (b) $\rho = 0.8$

3.3 Distribuciones relacionadas con la normal multivariante

En esta sección consideraremos distribuciones asociadas a una muestra de tamaño n de una vector aleatorio normal p-variante.

Paralelismos entre el caso univariante y el multivariante:

Caso univariante	Caso multivariante
$N(\mu, \sigma^2)$	$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
χ^2_n	$W_p(\boldsymbol{\Sigma}, n)$
$F(m,n)$	$\Lambda(p,a,b)$
t	T^2

La ley Wishart

Dada una muestra aleatoria simple de tamaño n de una ley $N_p(\vec{0}, \boldsymbol{\Sigma})$ (es decir, $\mathbf{X}_1, \ldots, \mathbf{X}_n$ sucesión de vectores aleatorios iid$\sim N_p(\vec{0}, \boldsymbol{\Sigma})$), se define la distribución de Wishart como la ley de la matriz aleatoria, $p \times p$,

$$\mathbf{Q} = \sum_{i=1}^{n} \mathbf{X}_i \mathbf{X}_i' \sim W_p(\boldsymbol{\Sigma}, n)$$

Propiedades:

1. $Q_1 \sim W_p(\boldsymbol{\Sigma}, n_1)$, $Q_2 \sim W_p(\boldsymbol{\Sigma}, n_2)$ independientes, entonces $Q_1 + Q_2 \sim W_p(\boldsymbol{\Sigma}, n_1 + n_2)$.
2. (Teorema de Fisher) Si $\mathbf{S}$ es la matriz de covarianzas de una muestra de tamaño n obtenida a partir de un vector con ley $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces:
   (i) El vector $\overline{\mathbf{x}}$ y la matriz $\mathbf{S}$ son estocásticamente independientes,
   (ii) $\overline{\mathbf{x}} \sim N_p(\boldsymbol{\mu}, \frac{1}{n}\boldsymbol{\Sigma})$,
   (iii) $n\mathbf{S} \sim W_p(\boldsymbol{\Sigma}, n-1)$.

La ley Lambda de Wilks

La distribución Lambda de Wilks es la ley del cociente de determinantes

$$\Lambda = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|} \sim \Lambda(p,a,b),$$

donde $\mathbf{A} \sim W_p(\boldsymbol{\Sigma}, a)$, $\mathbf{B} \sim W_p(\boldsymbol{\Sigma}, b)$, independientes, siendo $\boldsymbol{\Sigma}$ no singular y $a \geq p$.

Propiedades:

1. $\Lambda(p,a,b) = \Lambda(b, a+b-p, p)$.
2. En general, una transformación de Λ equivale, exacta o asintóticamente, a una F de Fisher. Por ejemplo:
   si $p=1$, $\frac{1-\Lambda}{\Lambda}\frac{a}{b} = F(b,a)$,
   si $p=2$, $\frac{1-\Lambda^{1/2}}{\Lambda^{1/2}}\frac{a-1}{b} = F(2b, 2(a-1))$.

La ley T^2 de Hotelling

Se define la distribución T^2 de Hotelling como la ley de la v.a.

$$T^2 = n\,\mathbf{Z}'\mathbf{Q}^{-1}\mathbf{Z} \sim T^2(p,n),$$

donde $\mathbf{Z} \sim N_p(\vec{0}, \mathbf{I})$, $\mathbf{Q} \sim W_p(\mathbf{I}, n)$ estocásticamente independientes.

Propiedades:

1. Si $p = 1$, $T^2(1,n)$ es el cuadrado de una v.a. con ley t de Student con n grados de libertad.
2. En general, $T^2(p,n)$ es proporcional a una F de Fisher, es decir, $\frac{n-p+1}{n\,p} T^2(p,n) = F(p, n-p+1)$.
3. Transformaciones afines sobre $\mathbf{Z}$ no afectan a la ley T^2. Es decir, si $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{R} \sim W_p(\boldsymbol{\Sigma}, n)$ independientes, entonces $n(\mathbf{X}-\boldsymbol{\mu})'\mathbf{R}^{-1}(\mathbf{X}-\boldsymbol{\mu}) \sim T^2(p,n)$.

4. Si $\mathbf{X}_1, \ldots, \mathbf{X}_n$ es una muestra aleatoria simple de una ley $N_p(\vec{0}, \boldsymbol{\Sigma})$, entonces:

   $$(n-1)(\overline{\mathbf{x}}-\boldsymbol{\mu})'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu}) = n(\overline{\mathbf{x}}-\boldsymbol{\mu})'\tilde{\mathbf{S}}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu}) \sim T^2(p, n-1).$$

5. Dadas dos muestras aleatorias simples e independientes, $\mathbf{X}_1, \ldots, \mathbf{X}_{n_1}$ iid $\sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$, $\mathbf{Y}_1, \ldots, \mathbf{Y}_{n_2}$ iid $\sim N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$, si $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, entonces:

   $$\frac{n_1 n_2}{n_1+n_2}(\overline{\mathbf{x}}-\overline{\mathbf{y}})'\mathbf{S}_P^{-1}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \sim T^2(p, n_1+n_2-2),$$

   donde $\mathbf{S}_P = \frac{1}{n_1+n_2}(n_1\mathbf{S}_1 + n_2\mathbf{S}_2)$ es la media aritmética ponderada de las matrices de covarianzas muestrales (*pooled covariance matrix*).

Observación 4 *Las propiedades 4 y 5 nos dicen que la ley T^2 de Hotelling se utlizará en los contrastes multivariantes sobre medias, de forma análoga a la ley t de Student en el caso univariante.*

3.4 Aplicaciones a la inferencia multivariante

[1]. Sea $\mathcal{X}$ una matriz de datos $n \times p$, cuyas filas representan una muestra aleatoria simple de una ley $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Consideremos el siguiente contraste de hipótesis:

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0, \quad \boldsymbol{\mu}_0 \in \mathbb{R}^p.$$

Cuando H_0 sea cierta, el estadístico

$$(n-1)(\overline{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu}_0) \sim T^2(p, n-1).$$

Utilizando la propiedad 2 de la ley T^2 de Hotelling, tendremos que

$$\frac{n-p}{p}(\overline{\mathbf{x}}-\boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{x}}-\boldsymbol{\mu}_0) \sim F(p, n-p)$$

y podremos obtener una región crítica para el nivel de significación deseado.

[2]. Consideremos dos matrices de datos $\mathcal{X}$ e $\mathcal{Y}$:
$\mathcal{X}$, $n_1 \times p$, provinenete de una ley $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$,
$\mathcal{Y}$, $n_2 \times p$, provinenete de una ley $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$ independientes.

Consideremos el siguiente contraste de hipótesis:

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2.$$

Cuando H_0 sea cierta, el estadístico

$$\frac{n_1 n_2}{n_1+n_2}(\overline{\mathbf{x}}-\overline{\mathbf{y}})'\mathbf{S}_P^{-1}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \sim T^2(p, n_1+n_2-2),$$

donde $\mathbf{S}_P = \frac{1}{n_1+n_2}(n_1\mathbf{S}_1 + n_2\mathbf{S}_2)$. De forma equivalente (utilizando la propiedad 2 de la ley T^2 de Hotelling), el estadístico

$$\frac{n_1+n_2-p-1}{(n_1+n_2-2)p}\,\frac{n_1 n_2}{n_1+n_2}(\overline{\mathbf{x}}-\overline{\mathbf{y}})'\mathbf{S}_P^{-1}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \sim F(p, n_1+n_2-p-1)$$

y podremos obtener una región crítica para el nivel de significación deseado.

[3]. Supongamos que tenemos g matrices de datos provinentes de g distribuciones normales multivariantes independientes:

muestra	tamaño matriz	media (muestral)	covarianza (muestral)	ley de probabilidad
$\mathcal{X}_1$	$n_1 \times p$	$\overline{\mathbf{x}}_1$	$\mathbf{S}_1$	$N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$
$\mathcal{X}_2$	$n_2 \times p$	$\overline{\mathbf{x}}_2$	$\mathbf{S}_2$	$N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathcal{X}_g$	$n_g \times p$	$\overline{\mathbf{x}}_g$	$\mathbf{S}_g$	$N_p(\boldsymbol{\mu}_g, \boldsymbol{\Sigma})$

La estimación muestral del vector de medias global y la matriz de covarianzas común son

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{g} n_i\,\overline{\mathbf{x}}_i, \quad \mathbf{S} = \frac{1}{n-g}\sum_{i=1}^{g} n_i\,\mathbf{S}_i, \quad \text{siendo } n = \sum_{i=1}^{g} n_i.$$

Queremos contraste de hipótesis $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_g$.

Para ello, introducimos las siguientes matrices:

$$\mathbf{B} = \sum_{i=1}^{g} n_i(\overline{\mathbf{x}}_i - \overline{\mathbf{x}}))(\overline{\mathbf{x}}_i - \overline{\mathbf{x}}))', \text{ (dispersión entre grupos)}$$

$$\mathbf{W} = \sum_{i=1}^{g}\sum_{k=1}(x_{ik} - \overline{\mathbf{x}}_i)(x_{ik} - \overline{\mathbf{x}}_i)' = \sum_{i=1}^{g} n_i\,\mathbf{S}_i, \text{ (dispersión dentro de los grupos)}$$

$$\mathbf{T} = \mathbf{B} + \mathbf{W}, \text{ (dispersión total)}.$$

Si H_0 es cierta, $\mathbf{B} \sim W_p(\boldsymbol{\Sigma}, g-1)$, $\mathbf{W} \sim W_p(\boldsymbol{\Sigma}, n-g)$ son independientes y $\mathbf{T} \sim W_p(\boldsymbol{\Sigma}, n-1)$. Para resolver este contraste utilizaremos el estadístico:

$$\Lambda = \frac{|\mathbf{W}|}{|\mathbf{W}+\mathbf{B}|} \sim \Lambda(p, n-g, g-1),$$

que se aproxima a una ley F de Fisher mediante la *aproximación asintótica de Rao*:

Si $\Lambda \sim \Lambda(p,a,b)$, entonces $\dfrac{1-\Lambda^{1/\beta}}{\Lambda^{1/\beta}}\,\dfrac{\alpha\,\beta - 2\gamma}{p\,b} \sim F(p\,b, \alpha\,\beta - 2\gamma)$,

donde $\alpha = a + b - (p+b+1)/2$, $\beta^2 = (p^2b^2 - 4)/(p^2 + b^2 - 5)$, $\gamma = (pb-2)/4$.

Ejemplo 3.3 *Un psicólogo realiza dos tests, uno de capacidad verbal (variable X_1) y otro de rendimiento intelectual general (variable X_2), a $n = 101$ individuos adultos, obteniendo los siguientes resultados:*

$$\overline{\mathbf{x}} = (55.24, 34.97)', \quad \mathbf{S} = \begin{pmatrix} 210.54 & 126.99 \\ 126.99 & 119.68 \end{pmatrix}$$

Contrastar la hipótesis de que la media poblacional es $\boldsymbol{\mu} = (60, 40)'$.

Se trata de un contraste para la media poblacional $H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0$, donde $\boldsymbol{\mu}_0 = (60, 40)'$, $n = 101$, $p = 2$. Por tanto, utilizaremos el estadístico T^2 de Hotelling

$$T^2 = (n-1)(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{x}} - \boldsymbol{\mu}_0) \sim T^2(p, n-1),$$

o, equivalentemente,

$$F = \frac{n-p}{p}(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)'\mathbf{S}^{-1}(\overline{\mathbf{x}} - \boldsymbol{\mu}_0) \sim F(p, n-p).$$

Mediante Matlab, obtenemos $F = 10.68 > F^{0.05}_{(2,99)} = 3.07$. Por tanto, se rechaza H_0, es decir, $\boldsymbol{\mu} \neq (60, 40)'$

Ejercicios computacionales

Ejercicio 3.1 *Generar dos muestras de tamaños $n_1 = 100$ y $n_2 = 150$ provinentes, respectivamente, de dos leyes normales multivariantes $N_3(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$ y $N_3(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$, con*

$$\boldsymbol{\mu}_1 = (1, 0, 1)',\ \boldsymbol{\mu}_2 = (0.9, 0.01, 1)' \text{ y } \boldsymbol{\Sigma} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

a) Obtener los vectores de medias y las matrices de covarianzas muestrales.

b) Contrastar la hipótesis $H_0 : \boldsymbol{\mu}_1 = (1, 0, 1)'$ vs. $H_1 : \boldsymbol{\mu}_1 \neq (1, 0, 1)'$.

c) Contrastar las hipótesis $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$ vs. $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, mediante el contraste de la T^2 de Hotelling y mediante el contraste de la Lambda de Wilks.