EYP 1027 Modelos Probabilísticos

Clase 14

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2020

Contenido I

1. Esperanza y Varianza de Vectores Aleatorios
   - Vector esperado y matriz de varianza-covarianza
   - Matriz de covarianza
   - Ejemplos
2. Matriz de correlación
   - Ejemplo
   - Propiedades básicas
3. Distribuciones Multivariadas Especiales
   - Distribución Multinomial
   - Distribución Normal Multivariada
   - Ejemplo: Normal Bivariada

A continuación, se generalizán los conceptos de esperanza y varianza de una variable aleatoria a un vector aleatorio. Para esto, el álgebra requerida se desarrollará en términos de vectores columna.

Definición 1.1

Esperanza de un vector aleatorio Sea $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ un vector aleatorio de dimensión n. La esperanza de $\boldsymbol{X}$, denotado como $\mathrm{E}(\boldsymbol{X})$, se define como,

$$\mathrm{E}(\boldsymbol{X}) = (\mathrm{E}(X_1), \mathrm{E}(X_2), \cdots, \mathrm{E}(X_n))^T.$$

Es decir, si $\boldsymbol{\mu} := \mathrm{E}(\boldsymbol{X})$ y $\mu_i = \mathrm{E}(X_i)$ para $i = 1, \ldots, n$, entonces $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ y también se llama **vector esperado**.

Esperanza y Varianza de Vectores Aleatorios

Definición 1.2

Matriz de varianza-covarianza Sea $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ un vector aleatorio de dimensión n. La matriz de varianza-covarianza de $\boldsymbol{X}$ se define como,

$$
\begin{aligned}
\mathrm{Var}(\boldsymbol{X}) &= \mathrm{E}\left([\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})][\boldsymbol{X} - \mathrm{E}(\boldsymbol{X})]^T\right) \\
&= \begin{pmatrix}
\mathrm{Var}\,(X_1) & \mathrm{Cov}\,(X_1, X_2) & \ldots & \mathrm{Cov}\,(X_1, X_n) \\
\mathrm{Cov}\,(X_2, X_1) & \mathrm{Var}\,(X_2) & \ldots & \mathrm{Cov}\,(X_2, X_n) \\
\vdots & \vdots & \ddots & \vdots \\
\mathrm{Cov}\,(X_n, X_1) & \mathrm{Cov}\,(X_n, X_2) & \cdots & \mathrm{Var}\,(X_n)
\end{pmatrix}.
\end{aligned}
$$

Es decir, si $\boldsymbol{\Sigma} := \mathrm{Var}(\boldsymbol{X})$ y $\sigma_{ij} = \mathrm{Cov}(X_j, X_j) = \mathrm{E}\{(X_i - \mu_i)(X_j - \mu_j)\}$, $i, j = 1, \ldots, n$, entonces $\boldsymbol{\Sigma} = ((\sigma_{ij}))_{i,j=1,\ldots,n}$.

Como una extensión de la matriz de varianza-covarianza, se define la matriz de covarianzas entre dos vectores aleatorios.

Definición 1.3

Matriz de covarianza Sean $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ e $\boldsymbol{Y} = (Y_1, \ldots, Y_m)^T$ dos vectores aleatorios definidos en un mismo espacio de probabilidad. La matriz de covarianza entre $\boldsymbol{X}$ e $\boldsymbol{Y}$ se define como,

$$
\begin{aligned}
\operatorname{Cov}(\boldsymbol{X}, \boldsymbol{Y}) &= \operatorname{E}\left([\boldsymbol{X} - \operatorname{E}(\boldsymbol{X})][\boldsymbol{Y} - \operatorname{E}(\boldsymbol{Y})]^T\right) \\
&= \begin{pmatrix}
\operatorname{Cov}(X_1, Y_1) & \operatorname{Cov}(X_1, Y_2) & \ldots & \operatorname{Cov}(X_1, Y_m) \\
\operatorname{Cov}(X_2, Y_1) & \operatorname{Cov}(X_2, Y_2) & \ldots & \operatorname{Cov}(X_2, Y_m) \\
\vdots & \vdots & \ddots & \vdots \\
\operatorname{Cov}(X_n, Y_1) & \operatorname{Cov}(X_n, Y_2) & \cdots & \operatorname{Cov}(X_n, Y_m)
\end{pmatrix}.
\end{aligned}
$$

Ejemplo 1.1

Sean X e Y variables aleatorias con fmp conjunta dada por,

$x \backslash y$	0	1	P(X=x)
-1	1/7	1/7	2/7
0	2/7	1/7	3/7
1	1/7	1/7	2/7
P(Y=y)	4/7	3/7	1

Es claro que,

$$\mathrm{E}(X) = -1 \times \frac{2}{7} + 1 \times \frac{2}{7} = 0,$$
$$\mathrm{Var}(X) = \mathrm{E}(X^2) = 2 \times \frac{2}{7} = \frac{4}{7},$$

$$\begin{aligned}
\mathsf{E}(Y) &= 1 \times \frac{3}{7} = \frac{3}{7}, \\
\mathsf{Var}(Y) &= \left(0 - \frac{3}{7}\right)^2 \times \frac{4}{7} + \left(1 - \frac{3}{7}\right)^2 \times \frac{3}{7} = \frac{12}{49}, \\
\mathsf{Cov}(X, Y) &= \mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y) \\
&= -1 \times 1 \times \frac{1}{7} + 1 \times 1 \times \frac{1}{7} - 0 \times \frac{3}{7} = 0.
\end{aligned}$$

Entonces, el vector de medias y la matriz de varianza-covarianza de $\boldsymbol{X} = (X, Y)^\top$ están dadas por,

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 3/7 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 4/7 & 0 \\ 0 & 12/49 \end{pmatrix}.$$

Ejemplo 1.2

Sea (X, Y) un vector aleatorio con fdp conjunta dada por,

$$f_{X,Y}(x, y) = \begin{cases} \dfrac{1}{x} & \text{si } 0 < x \leq 1,\ 0 < y \leq x \\ 0 & \text{eoc} \end{cases}$$

En esta caso tenemos,

$$E(X) = \int_0^1 \int_y^1 x\frac{1}{x}dxdy = \frac{1}{2},$$

$$E(Y) = \int_0^1 \int_0^x y\frac{1}{x}dydx = \frac{1}{4},$$

$$E\left(X^2\right) = \int_0^1 \int_0^x x^2\frac{1}{x}dydx = \frac{1}{3} \Longrightarrow \text{Var}(X) = \frac{1}{3} - \frac{1}{2^2},$$

$$E\left(Y^2\right)=\int_0^1\int_0^x y^2\frac{1}{x}dydx=\frac{1}{9}\Longrightarrow \mathrm{Var}(Y)=\frac{1}{9}-\frac{1}{4^2},$$
$$E(XY)=\int_0^1\int_0^x xy\frac{1}{x}dydx=\frac{1}{6}\Longrightarrow \mathrm{Cov}(X,Y)=\frac{1}{6}-\frac{1}{2}\frac{1}{4}.$$

Por lo tanto, el valor esperado $\boldsymbol{\mu}=\mathrm{E}(\boldsymbol{X})$ está dado por,

$$\mathrm{E}(\boldsymbol{X})=\boldsymbol{\mu}=\begin{pmatrix}1/2\\1/4\end{pmatrix},$$

y la matriz de varianza-covarianza $\boldsymbol{\Sigma}=\mathrm{Var}(\boldsymbol{X})$ es,

$$\boldsymbol{\Sigma}=\begin{pmatrix}1/12 & 1/24\\1/24 & 7/144\end{pmatrix}.$$

Definición 2.1

Matriz de Correlación Sea $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ un vector aleatorio n dimensional. La matriz de correlación de $\boldsymbol{X}$, denotada como R, se define como,

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_{X_1,X_2} & \cdots & \rho_{X_1,X_n} \\ \rho_{X_2,X_1} & 1 & \cdots & \rho_{X_2,X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{X_n,X_1} & \rho_{X_n,X_2} & \cdots & 1 \end{pmatrix}.$$

De forma similar se define la matriz de correlación $\mathbf{R}_{XY} = ((\rho_{X_i,Y_j}))_{n\times m}$ entre dos vectores aleatorios $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ e $\boldsymbol{Y} = (Y_1, \ldots, Y_m)^\top$, de modo que la matriz $\mathbf{R}$ se tiene para $\boldsymbol{Y} = \boldsymbol{X}$.

Ejemplo 2.1

Sea $\boldsymbol{X} = (X, Y)^T$ un vector aleatorio bi-dimensional con fdp conjunta dada por,

$$f(x,y) = \begin{cases} 2 & \text{si } 0 < x < y, 0 < y < 1 \\ 0 & \text{eoc} \end{cases}$$

Tenemos que,

$$E(X) = \int_0^1 \int_0^y 2x dx dy = \frac{1}{3}$$

$$E(Y) = \int_0^1 \int_x^1 2y dy dx = \frac{2}{3},$$

$$\begin{aligned}
\mathrm{Var}(X) &= \int_0^1 \int_0^y 2\left(x - \frac{1}{3}\right)^2 dxdy = \frac{1}{18} \\
\mathrm{Var}(Y) &= \int_0^1 \int_x^1 2\left(y - \frac{2}{3}\right)^2 dydx = \frac{1}{18} \\
\mathrm{Cov}(X,Y) &= \int_0^1 \int_0^y 2\left(x - \frac{1}{3}\right)\left(y - \frac{2}{3}\right) dxdy = \frac{1}{36}.
\end{aligned}$$

Por lo tanto la matriz de correlación está dada por,

$$\mathbf{R} = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}.$$

1) $\mathsf{E}(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b})=\boldsymbol{A}\mathsf{E}(\boldsymbol{X})+\boldsymbol{b}$ y $\mathsf{Var}(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{b})=\boldsymbol{A}\mathsf{Var}(\boldsymbol{X})\boldsymbol{A}^\top$

2a) $\boldsymbol{\Sigma}=\mathsf{Var}(\boldsymbol{X})$ es una matriz simétrica, $\boldsymbol{\Sigma}=((\sigma_{ij}))=((\sigma_{ji}))=\boldsymbol{\Sigma}^\top$

2b) $\boldsymbol{\Sigma}=\mathsf{Var}(\boldsymbol{X})$ es una matriz semidefinida positiva (s.d.p.), es decir, $\boldsymbol{a}^\top\boldsymbol{\Sigma}\boldsymbol{a}\geq 0\ \forall\,\boldsymbol{a}\in\mathbb{R}^n$

2c) $\boldsymbol{\Sigma}=\mathsf{Var}(\boldsymbol{X})$ es una matriz definida positiva (d.p.), es decir, $\boldsymbol{a}^\top\boldsymbol{\Sigma}\boldsymbol{a}\geq 0\quad\forall\,\boldsymbol{a}\neq\boldsymbol{0}\iff X_1,\ldots,X_n$ son linealmente independientes.

3) $\mathsf{Cov}(\boldsymbol{X},\boldsymbol{Y})=\mathsf{E}(\boldsymbol{X}\boldsymbol{Y}^\top)-\mathsf{E}(\boldsymbol{X})\mathsf{E}(\boldsymbol{Y})^\top$, $\mathsf{Cov}(\boldsymbol{X},\boldsymbol{X})=\mathsf{Var}(\boldsymbol{X})$ y $\mathsf{Cov}(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{C},\boldsymbol{B}\boldsymbol{Y}+D)=\boldsymbol{A}\mathsf{Cov}(\boldsymbol{X},\boldsymbol{Y})\boldsymbol{B}^\top$.

Note que $\mathsf{Cov}(\boldsymbol{X},\boldsymbol{Y})=\mathsf{Cov}(\boldsymbol{Y},\boldsymbol{X})^\top$, de modo que
$\mathsf{Var}(\boldsymbol{X}\pm\boldsymbol{Y})=\mathsf{Var}(\boldsymbol{X})+\mathsf{Var}(\boldsymbol{Y})\pm\mathsf{Cov}(\boldsymbol{X},\boldsymbol{Y})\pm\mathsf{Cov}(\boldsymbol{Y},\boldsymbol{X})$.

Esperanza y Varianza de Vectores Aleatorios

Para demostrar las Propiedades 2b) y 2c) de la matriz de varianza-covarianza $\boldsymbol{\Sigma} = \mathsf{Var}(\boldsymbol{X})$, sea

$$Y = \sum_{i=1}^{n} a_i X_i = (a_1, \ldots, a_n) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \boldsymbol{a}^\top \boldsymbol{X}$$

$$\Longrightarrow \mathsf{Var}(Y) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \mathsf{Cov}(X_i, X_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \sigma_{ij} = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a}.$$

Pero, para toda variable aleatoria Y,

$$\mathsf{Var}(Y) \geq 0 \Longrightarrow \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a} \geq 0 \quad \forall\, \boldsymbol{a}$$

$$\overset{\text{def.}}{\Longleftrightarrow} \boldsymbol{\Sigma} \text{ es s.d.p. } (\boldsymbol{\Sigma} \geq 0)$$

Esperanza y Varianza de Vectores Aleatorios

Para $\boldsymbol{a} \neq \mathbf{0}$,

$$
\begin{aligned}
\mathsf{Var}(Y) = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a} > 0 &\iff Y = \sum_{i=1}^{n} a_i X_i \neq \text{constante} \\
&\iff X_1, \ldots, X_n \text{ son linealmente independientes} \\
&\overset{\text{def.}}{\iff} \boldsymbol{\Sigma} \text{ es d.p. } (\boldsymbol{\Sigma} > 0).
\end{aligned}
$$

Considere un experimento aleatorio con k resultados posibles $R_1, \ldots, R_k$ exhaustivos y mutuamente excluyentes, es decir,

$$\cup_{i=1}^{k} R_i = \Omega \quad \text{y} \quad R_i \cap R_j = \emptyset \quad \forall\ i \neq j$$

Sea $p_i = P(R_i)$, $i = 1, \ldots, k$,

$$\Longrightarrow p_i \geq 0 \quad \forall\ i \quad \text{y} \quad \sum_{i=1}^{k} p_i = 1. \tag{3.1}$$

Para n repeticiones independientes de este experimento, defina las variables aleatorias,

$$X_i = \text{ número de veces que ocurre } R_i, \quad i = 1, \ldots, k,$$

$$\Longrightarrow X_i(\omega) \in \mathcal{X}_i = \{0, 1, \ldots, n\} \quad \forall\ i \quad \text{y} \quad \sum_{i=1}^{k} X_i(\omega) = n \quad \forall\ \omega \in \Omega.$$

Luego, en n repeticiones independientes, la probabilidad del evento

$$\begin{aligned}\{X_1 = x_1, \ldots, X_k = x_k\} &:= \cap_{i=1}^{k}\{X_i = x_i\} \\ &= \cap_{i=1}^{k}\{R_i \text{ ocurre exactamente } x_i \text{ veces}\},\end{aligned}$$

esta dada por,

$$P(X_1 = x_1, \ldots, X_k = x_k) = \begin{cases} \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} & \text{si } (x_1, \ldots, x_k) \in \mathcal{X}, \\ \\ 0 & \text{eoc.} \end{cases} \tag{3.2}$$

donde $\mathcal{X} = \{(x_1, \ldots, x_k) : x_i = 1, \ldots, n, \ \sum_{i=1}^{k} x_i = n\}$ es el recorrido conjunto de $X_1, \ldots, X_k$.

Definición 3.1

Se dice que un vector aleatorio discreto $(X_1, \ldots, X_k)$ tiene distribución multinomial con parámetros n y $(p_1, \ldots, p_k)$, si las variables aleatorias $X_1, \ldots, X_k$ tienen fmp conjunta dada por (3.2); en tal caso, se escribe,

$$(X_1, \ldots, X_k) \sim Mult(n, p_1, \ldots, p_k),$$

donde $n \in \{1, 2, \ldots\}$ y $p_1, \ldots, p_k$ satisfacen (3.1).

Nota: Ya que $\sum_{i=1}^{n} X_i = n$ ($X_1, \ldots, X_k$ son linealmente dependientes), basta con especificar la distribución conjunta de $X_1, \ldots, X_{k-1}$.

Ejemplo 3.1

En 12 lanzamientos de un dado honesto, sea $X_i =$ número de veces que aparece el i-ésimo resultado, $i = 1, \ldots, 6$. Entonces,

$$(X_1, \ldots, X_6) \sim Mult(n = 10, p_1 = 1/6, \ldots, p_6 = 1/6).$$

Así,

$$P(X_1 = 2, X_2 = 3, X_3 = 2, X_4 = 2, X_5 = 2, X_6 = 3) = \frac{12!}{(2!)^3(3!)^3}(1/6)^{12}.$$

Propiedades: Si $(X_1, \dots, X_k) \sim Mult(n, p_1, \dots, p_k)$, entonces,

1) La fmg conjunta de $X_1, \dots, X_k$ es,

$$\begin{aligned}
M_{X_1,\dots,X_k}(t_1, \dots, t_k) &= \mathsf{E}\left(e^{\sum_{i=1}^k t_i X_i}\right) \\
&= \sum_{(x_1,\dots,x_k)\in\mathcal{X}} e^{\sum_{i=1}^k t_i x_i} \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \\
&= \sum_{(x_1,\dots,x_k)\in\mathcal{X}} \frac{n!}{x_1! \cdots x_k!} (p_1 e^{t_1})^{x_1} \cdots (p_k e^{t_k})^{x_k} \\
&= (p_1 e^{t_1} + \cdots + p_k e^{t_k})^n \quad \forall\ (t_1, \dots, t_k) \in \mathbb{R}^k
\end{aligned}$$

(teorema del multinomio).

2) Para todo $m = 1, \ldots, k$,

$$(X_1, \ldots, X_m; n - \sum_{i=1}^{m} X_i) \sim Mult(n; p_1, \ldots, p_m; 1 - \sum_{i=1}^{m} p_i).$$

En particular,

$$(X_i; n - \sum_{l \neq i} X_i) \sim Bin(n, p_i; 1 - \sum_{l \neq i} p_l)$$

$$\Longrightarrow \mathsf{E}(X_i) = np_i \ \text{ y } \ \mathsf{Var}(X_i) = np_i(1 - p_i), \quad i = 1, \ldots, k,$$

$$(X_i, X_j; n - \sum_{l \neq i,j} X_l) \sim Trin(n; p_i, p_j; 1 - \sum_{l \neq i,j} p_l)$$

$$\Longrightarrow \mathsf{Cov}(X_i, X_j) = -np_ip_j \quad \forall\ i \neq j.$$

Distribuciones Multivariadas Especiales

3) Suponga que se reagrupan los resultados $R_1, \ldots, R_k$ en m grupos exhaustivos y excluyentes $S_1, \ldots, S_m$, es decir,

$$\cup_{j=1}^{m} S_j = \Omega \quad \text{y} \quad S_i \cap S_j = \emptyset \quad \forall\, i \neq j,$$

de modo que

$$P(S_j) = \pi_j \geq 0, \quad j = 1, \ldots, m, \;\; \text{y} \;\; \sum_{j=1}^{m} \pi_j = 1.$$

Sea $Y_j =$ número de veces que ocurre S_j en las n repeticiones independientes, $j = 1, \ldots, m$. Entonces,

$$(Y_1, \ldots, Y_m) \sim Mult(n, \pi_1, \ldots, \pi_m).$$

Sea $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^\top$, donde $Z_1, \ldots, Z_m$ son variables aleatorias iid $N(0,1)$. Entonces,

$$
\begin{aligned}
f_{\boldsymbol{Z}}(\boldsymbol{z}) &= f_{Z_1,\ldots,Z_m}(z_1, \ldots, z_m) \\
&= \prod_{i=1}^{m} f_{Z_i}(z_i) \quad \text{(hip. de independencia)} \\
&= \prod_{i=1}^{m} (2\pi)^{-1/2} e^{-\frac{1}{2} z_i^2} \quad (Z_i \sim N(0,1) \ \ \forall\ i) \\
&= (2\pi)^{-m/2} e^{-\frac{1}{2}\sum_{i=1}^{m} z_i^2} \quad z_i \in \mathbb{R} \ \ \forall\ i, \\
&= (2\pi)^{-m/2} e^{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}} \quad \boldsymbol{z} = (z_1, \ldots, z_m)^\top \in \mathbb{R}^m.
\end{aligned}
$$

Además,

$$\begin{aligned}
M_{\boldsymbol{Z}}(\boldsymbol{t}) &= M_{Z_1,\ldots,Z_m}(t_1,\ldots,t_m) \\
&= \prod_{i=1}^{m} M_{Z_i}(t_i) \quad \text{(hip. de independencia)} \\
&= \prod_{i=1}^{m} e^{\frac{1}{2}t_i^2} \quad (Z_i \sim N(0,1) \ \ \forall\ i) \\
&= e^{-\frac{1}{2}\sum_{i=1}^{m} t_i^2} \quad t_i \in \mathbb{R} \ \ \forall\ i, \\
&= e^{-\frac{1}{2}\boldsymbol{t}^\top \boldsymbol{t}} \quad \boldsymbol{t} = (t_1,\ldots,t_m)^\top \in \mathbb{R}^m.
\end{aligned}$$

Note también que:

$$\mathsf{E}(Z_i) = 0, \quad i = 1, \ldots, m, \quad \text{y}$$

$$\mathsf{Cov}(Z_i, Z_j) = \begin{cases} 1 & \text{si } i = j,\ i, j = 1, \ldots, m, \\ 0 & \text{si } i \neq j,\ i, j = 1, \ldots, m. \end{cases}$$

$$\Longrightarrow \begin{cases} \mathsf{E}(\boldsymbol{Z}) = (\mathsf{E}(Z_1), \ldots, \mathsf{E}(Z_m))^\top = \boldsymbol{0} \in \mathbb{R}^m, \quad \text{y} \\ \mathsf{Var}(\boldsymbol{Z}) = ((\mathsf{Cov}(Z_i, Z_j))) = \boldsymbol{I}_m. \end{cases}$$

Sean $X_i = \sum_{j=1}^n a_{ij} Z_j + \mu_i$ para $i = 1, \ldots, n$, donde a_{ij} y μ_i, $i, j = 1, \ldots, n$, son constantes reales. Defina $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$, $\boldsymbol{A} = ((a_{ij}))$ y $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$. Matricialmente, se tiene que

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{Z} + \boldsymbol{\mu}, \quad \boldsymbol{A} \in \mathbb{R}^{n \times m}, \quad \boldsymbol{\mu} \in \mathbb{R}^{n \times 1}.$$

Entonces,

$$\mathsf{E}(\boldsymbol{X}) = \mathsf{E}(\boldsymbol{A}\boldsymbol{Z} + \boldsymbol{\mu}) = \boldsymbol{A}\mathsf{E}(\boldsymbol{Z}) + \boldsymbol{\mu} = \boldsymbol{\mu} \quad (\mathsf{E}(\boldsymbol{Z}) = \mathbf{0}),$$

$$\mathsf{Var}(\boldsymbol{X}) = \mathsf{Var}(\boldsymbol{A}\boldsymbol{Z} + \boldsymbol{\mu}) = \boldsymbol{A}\mathsf{Var}(\boldsymbol{Z})\boldsymbol{A}^\top = \boldsymbol{A}\boldsymbol{A}^\top := \boldsymbol{\Sigma} \quad (\mathsf{Var}(\boldsymbol{Z}) = \boldsymbol{I}_n).$$

Además,

$$\begin{aligned}
\mathsf{M}_{\boldsymbol{X}}(\boldsymbol{t}) &= \mathsf{E}\left(e^{\boldsymbol{t}^\top \boldsymbol{X}}\right) \\
&= \mathsf{E}\left(e^{\boldsymbol{t}^\top (\boldsymbol{A}\boldsymbol{Z} + \boldsymbol{\mu})}\right) \\
&= \mathsf{E}\left(e^{\boldsymbol{t}^\top \boldsymbol{A}\boldsymbol{Z} + \boldsymbol{t}^\top \boldsymbol{\mu}}\right) \\
&= e^{\boldsymbol{t}^\top \boldsymbol{\mu}} \mathsf{E}\left(e^{(\boldsymbol{A}^\top \boldsymbol{t})^\top \boldsymbol{Z}}\right) \\
&= e^{\boldsymbol{t}^\top \boldsymbol{\mu}} M_{\boldsymbol{Z}}\left(\boldsymbol{A}^\top \boldsymbol{t}\right) \\
&= e^{\boldsymbol{t}^\top \boldsymbol{\mu}} e^{\frac{1}{2}(\boldsymbol{A}^\top \boldsymbol{t})^\top (\boldsymbol{A}^\top \boldsymbol{t})} \\
&= e^{\boldsymbol{t}^\top \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top \boldsymbol{A}\boldsymbol{A}^\top \boldsymbol{t}} \\
&= e^{\boldsymbol{t}^\top \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}}, \quad \boldsymbol{t} \in \mathbb{R}^n.
\end{aligned}$$

Definición 3.2

Se dice que un vector aleatorio $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ tiene distribución normal n-variada con vector de medias $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ y matriz de varianza-covarianza $\boldsymbol{\Sigma} = ((\sigma_{ij}))_{i,j=1,\ldots,n}$, lo cual se escribe como $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, ssi

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

donde $\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{\Sigma}$ y $\boldsymbol{Z} = (Z_1, \ldots, Z_m)^\top \sim N_m(\boldsymbol{0}, \boldsymbol{I}_m)$, es decir, $Z_1, \ldots, Z_m \overset{iid}{\sim} N(0,1)$. En otras palabras,

$$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \overset{\text{def.}}{\Longleftrightarrow} M_{\boldsymbol{X}}(\boldsymbol{t}) = e^{\boldsymbol{t}^\top \boldsymbol{\mu} + \frac{1}{2}\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}}, \quad \boldsymbol{t} \in \mathbb{R}^n.$$

Equivalentemente, si $\boldsymbol{\Sigma} > 0$ (matriz definida positiva), entonces,

$$\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \overset{\text{def.}}{\Longleftrightarrow} f_{\boldsymbol{X}}(\boldsymbol{x}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}-\boldsymbol{\mu})},$$
$$\boldsymbol{x} \in \mathbb{R}^n.$$

Propiedades básicas: Sea $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Entonces:

1) $\boldsymbol{B}\boldsymbol{X} + \boldsymbol{b} \sim N_m(\boldsymbol{B}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top)$ para cualquier matriz $\boldsymbol{B} \in \mathbb{R}^{m \times n}$ y vector $\boldsymbol{b} \in \mathbb{R}^{m \times 1}$.

Demostración: Use fgm.

En particular, si $\boldsymbol{\Sigma} > 0$, entonces $\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \sim N_n(\boldsymbol{0}, \boldsymbol{I}_n)$, donde $\boldsymbol{B} = \boldsymbol{\Sigma}^{-1/2} = (\boldsymbol{\Sigma}^{1/2})^{-1}$ y $\boldsymbol{\Sigma}^{1/2}$ es la única raíz cuadrada simétrica de $\boldsymbol{\Sigma}$.

2) Considere la partición,

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

donde los $\boldsymbol{X}_j$'s y $\boldsymbol{\mu}_j$´s son vectores $n_j \times 1$, los $\boldsymbol{\Sigma}_{ij}$'s son matrices $n_i \times n_j$ y $n_1 + n_2 = n$. Entonces:

a) $\boldsymbol{X}_j \sim N_{n_j}(\boldsymbol{\mu}_j, \Sigma_{jj})$, $j = 1, 2$, donde $n_1 + n_2 = n$.

Demostración: Use la Propiedad 1), colocando $\boldsymbol{b} = \boldsymbol{0}$ y $\boldsymbol{B} = (\boldsymbol{I}_{n_1}, \boldsymbol{0})$ para $j = 1$, y $\boldsymbol{B} = (\boldsymbol{0}, \boldsymbol{I}_{n_1})$ para $j = 2$.

b) $\boldsymbol{X}_1$ y $\boldsymbol{X}_2 \Longleftrightarrow \text{Cov}(\boldsymbol{X}_1, \boldsymbol{X}_2) = \boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21}^\top = \boldsymbol{0}$.

Demostración: Use fgm.

3) Si $\boldsymbol{\Sigma} > 0$, entonces $(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2_n$.

Ejemplo 3.2

Para $n = 2$ con $X_1 = X$ y $X_2 = Y$, se tiene la distribución normal bivariada dada por

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}\right),$$

donde $\mu_X = \mathrm{E}(X)$, $\mu_Y = \mathrm{E}(Y)$, $\sigma_X^2 = \mathrm{Var}(X)$, $\sigma_Y^2 = \mathrm{Var}(Y)$ y $\sigma_{XY} = \mathrm{Cov}(X, Y)$. Note que

$$\sigma_{XY} = \rho_{XY}\sigma_X\sigma_Y,$$

donde ρ_{XY} es el coeficiente de correlación enetre X e Y, y σ_X y σ_Y son las desviaciones estándar de X e Y, respectivamente.

Si $-1 < \rho_{XY} < 1$, entonces $\Sigma > 0$ y, por lo tanto, es invertible. En tal caso, (X, Y) tiene fdp normal bivariada dada por,

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left\{-\frac{1}{2\left(1-\rho_{XY}^2\right)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 \right.\right.$$
$$\left.\left. -2\rho_{XY}\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\},$$

para $-\infty < x < \infty$ y $-\infty < y < \infty$.

Distribución Normal Bivariada

Aunque la fdp anterior parece complicada, la distribución normal bivariada es una de las más utilizadas. Algunas de sus muchas propiedades incluyen,

i) La distribución marginal de X es $N\left(\mu_X, \sigma_X^2\right)$
ii) La distribución marginal de Y is $N\left(\mu_Y, \sigma_Y^2\right)$
iii) $\rho_{XY} = 0 \Longleftrightarrow X$ e Y son independientes
iv) Para constantes cualquiera a y b, la distribución de $aX + bY$ es $N\left(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\rho_{XY}\sigma_X\sigma_Y\right)$

Distribución Normal Bivariada

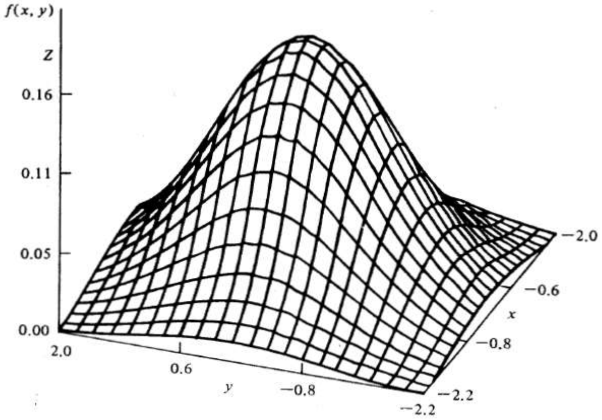

Figura 1: Densidad normal bivariada con $\mu_X = \mu_Y = 0, \sigma_X = \sigma_Y = 1$ y $\rho_{XY} = 0$.

Distribución Normal Bivariada

Ejemplo 3.3

Suponga que se selecciona al azar una pareja formada por un hombre y una mujer de una determinada población. Sea X la altura de la mujer e Y la altura del hombre, ambas medidas en pulgadas. Se sabe que la distribución conjunta de X e Y es normal bivariada con medias $\mu_X = 66.8$ y $\mu_Y = 70$, desviaciones estándar $\sigma_X = \sigma_Y = 2$, y correlación $\rho_{XY} = 0.68$. Calcule la probabilidad de que la mujer sea más alta que el hombre, es decir, $P(X - Y > 0)$.

Distribución Normal Bivariada

Dado que (X, Y) tienen una distribución normal bivariada, se tiene que la distribución de $X - Y$ también es normal, con una media

$$E(X - Y) = 66.8 - 70 = -3.2$$

y varianza

$$\begin{aligned} \mathrm{Var}(X - Y) &= \sigma_X^2 + \sigma_Y^2 - 2\mathrm{Cov}(X, Y) \\ &= 4 + 4 - 2(0.68)(2)(2) = 2.56. \end{aligned}$$

Luego, como la desviación estándar de $X - Y$ es $\sigma_{X-Y} = 1.6$, entonces la estandarización de $X - Y$ es $Z = (X - Y + 3.2)/1.6 \sim N(0, 1)$. Luego,

$$\begin{aligned} P(X - Y > 0) &= P(Z > 2) = 1 - \Phi(2) \\ &= 0.0227 \end{aligned}$$