

# Probabilidad

FACULTAD  
DE  
CIENCIAS  
UNIVERSIDAD DE GRANADA



Los Del DGIIM, [losdeldgiim.github.io](https://losdeldgiim.github.io)

Doble Grado en Ingeniería Informática y Matemáticas  
Universidad de Granada



Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional (CC BY-NC-ND 4.0).

Eres libre de compartir y redistribuir el contenido de esta obra en cualquier medio o formato, siempre y cuando des el crédito adecuado a los autores originales y no persigas fines comerciales.

# Probabilidad

Los Del DGIIM, [losdeldgiim.github.io](https://losdeldgiim.github.io)

Arturo Olivares Martos

Granada, 2024-2025



# Índice general

<b>1. Algunos Modelos Multivariantes</b>	<b>5</b>
1.1. Distribución Multinomial . . . . .	5
1.2. Distribución Normal Bidimensional . . . . .	11
<b>2. Relaciones de problemas</b>	<b>19</b>
2.1. Algunos Modelos Multivariantes . . . . .	20



# 1. Algunos Modelos Multivariantes

En el presente Capítulo, y al igual que hemos hecho en el caso unidimensional, vamos a estudiar ahora ciertos modelos multivariantes que nos permitirán trabajar con más de una variable aleatoria. En concreto, vamos a generalizar la distribución binomial y la normal, ya estudiadas previamente.

## 1.1. Distribución Multinomial

Como ya hemos adeñantado, esta distribución es una generalización de la distribución binomial. Fijamos  $k \in \mathbb{N}$ , y consideramos un experimento aleatorio con  $k + 1$  posibles resultados asociados a otros tantos sucesos  $A_1, A_2, \dots, A_{k+1}$  exhaustivos y mutuamente excluyentes, por lo tanto constituyen una partición del espacio muestral, con probabilidad de ocurrencia  $p_i \in ]0, 1[$ ,  $i = 1, \dots, k + 1$ . Es decir:

$$\begin{aligned}\Omega &= \bigcup_{i=1}^{k+1} A_i, \\ A_i \cap A_j &= \emptyset, \quad i, j = 1, \dots, k + 1, \quad i \neq j, \\ P(A_i) &= p_i, \quad i = 1, \dots, k + 1\end{aligned}$$

Tenemos entonces que:

$$P(A_{k+1}) = 1 - P\left(\bigcup_{i=1}^k A_i\right) = 1 - \sum_{i=1}^k P(A_i) = 1 - \sum_{i=1}^k p_i$$

Supongamos que realizamos  $n$  repeticiones independientes del experimento en las mismas condiciones, por tanto las probabilidades  $p_i$ ,  $i = 1, 2, \dots, k + 1$  se mantienen constantes a lo largo de las repeticiones, y sean  $X_1, \dots, X_{k+1}$  variables aleatorias tales que  $X_i$  contabiliza el número de veces que ocurre el suceso  $A_i$  en las  $n$  repeticiones del experimento,  $i = 1, 2, \dots, k + 1$ . Tenemos que:

$$X_1 + X_2 + \dots + X_{k+1} = n$$

En estas condiciones, podemos definir la distribución multinomial.

**Definición 1.1** (Distribución Multinomial). En las condiciones anteriores, se define la distribución multinomial  $k$ -dimensional con parámetros  $n$  y  $p_1, \dots, p_k$  como la

distribución del vector aleatorio  $X = (X_1, \dots, X_k)$  en el cual cada componente  $X_i$  contabiliza el número de ocurrencias del suceso  $A_i$ ,  $i = 1, \dots, k$ . La notaremos por:

$$X \sim M_k(n, p_1, \dots, p_k)$$

*Observación.* Notemos que la distribución binomial es un caso particular de la distribución multinomial, en concreto, la distribución multinomial con  $k = 1$ .

Razonemos ahora la función de masa de probabilidad de la distribución multinomial. Consideramos el vector  $x = (x_1, \dots, x_k)$  tal que  $x_i \in \mathbb{N}_0$ ,  $i = 1, \dots, k$ . Como se realizan  $n$  repeticiones del experimento, tenemos que:

$$\sum_{i=1}^k x_i \leq n$$

donde la igualdad solo se dará si no se produce el suceso  $A_{k+1}$  en ninguna de las repeticiones. Una de las posibles ordenaciones para que  $x$  tome dicho valor es que primero ocurran  $x_1$  veces el suceso  $A_1$ , luego  $x_2$  veces el suceso  $A_2$ , y así sucesivamente, hasta que ocurran  $x_k$  veces el suceso  $A_k$  y finalmente ocurran  $n - \sum_{i=1}^k x_i$  veces el suceso  $A_{k+1}$ . Es decir, el suceso sería (notando  $A_i^{x_i}$  al suceso  $A_i$  ocurrido  $x_i$  veces):

$$A_1^{x_1} \cap A_2^{x_2} \cap \dots \cap A_k^{x_k} \cap A_{k+1}^{n - \sum_{i=1}^k x_i}$$

La probabilidad de que ocurra este suceso, usando la independencia de las repeticiones del experimento, es:

$$P \left[ \bigcap_{i=1}^k \left( \bigcap_{j=1}^{x_i} A_i \right) \cap \left( \bigcap_{j=1}^{n - \sum_{i=1}^k x_i} A_{k+1} \right) \right] = \prod_{i=1}^k p_i^{x_i} \cdot p_{k+1}^{n - \sum_{i=1}^k x_i} = \prod_{i=1}^k p_i^{x_i} \cdot \left( 1 - \sum_{i=1}^k p_i \right)^{n - \sum_{i=1}^k x_i}$$

Veamos ahora cuántas ordenaciones distintas nos pueden dar dicho vector  $x$ . Estas ordenaciones son permutaciones con repetición de  $n$  elementos clasificados en  $k + 1$  grupos, habiendo  $x_i$  del tipo  $i$ -ésimo,  $i = 1, \dots, k$  y  $n - \sum_{i=1}^k x_i$  del grupo  $k + 1$ -ésimo. Por tanto, el número de ordenaciones posibles es<sup>1</sup>:

$$PR_n^{x_1, \dots, x_k, n - \sum_{i=1}^k x_i} = \frac{n!}{\left( \prod_{i=1}^k x_i! \right) \cdot \left( n - \sum_{i=1}^k x_i \right)!}$$

Por tanto, tenemos que:

$$P[X = (x_1, \dots, x_k)] = \frac{n!}{\left( \prod_{i=1}^k x_i! \right) \cdot \left( n - \sum_{i=1}^k x_i \right)!} \cdot \prod_{i=1}^k p_i^{x_i} \cdot \left( 1 - \sum_{i=1}^k p_i \right)^{n - \sum_{i=1}^k x_i}$$

Comprobemos que, efectivamente, esta función de masa de probabilidad es tal. Para ello, es necesario introducir el siguiente lema técnico, descubierto por Leibniz.

<sup>1</sup>Ver sección de Combinatoria en los apuntes de EDIP.



**Lema 1.1** (Fórmula de Leibniz). Sean  $p_1, \dots, p_j$  números reales y  $n \in \mathbb{N}$ . Entonces, se cumple que:

$$\left( \sum_{i=1}^j p_i \right)^n = \sum_{\sum_{i=1}^j x_i = n} \frac{n!}{\prod_{i=1}^j x_i!} \cdot \prod_{i=1}^j p_i^{x_i}$$

**Proposición 1.2.** Sea  $X \sim M_k(n, p_1, \dots, p_k)$ . Sean  $x_1, \dots, x_k$  tales que  $x_i \in \{0, \dots, n\}$ ,  $i = 1, \dots, k$  y  $\sum_{i=1}^k x_i \leq n$ . Entonces, la función de masa de probabilidad de  $X$  es:

$$P[X = (x_1, \dots, x_k)] = \frac{n!}{\left( \prod_{i=1}^k x_i! \right) \cdot \left( n - \sum_{i=1}^k x_i \right)!} \cdot \prod_{i=1}^k p_i^{x_i} \cdot \left( 1 - \sum_{i=1}^k p_i \right)^{n - \sum_{i=1}^k x_i}$$

*Demostración.* Usaremos para ello la Fórmula de Leibniz identificando  $k+1 = j$  y definiendo  $p_{k+1} = 1 - \sum_{i=1}^k p_i$  y  $x_{k+1} = n - \sum_{i=1}^k x_i$ :

$$\begin{aligned} \sum_{\sum_{i=1}^k x_i \leq n} \frac{n!}{\left( \prod_{i=1}^k x_i! \right) \cdot \left( n - \sum_{i=1}^k x_i \right)!} \cdot \prod_{i=1}^k p_i^{x_i} \cdot \left( 1 - \sum_{i=1}^k p_i \right)^{n - \sum_{i=1}^k x_i} &= \\ &= \sum_{\substack{\sum_{i=1}^{k+1} x_i = n \\ x_{k+1} = n - \sum_{i=1}^k x_i}} \frac{n!}{\prod_{i=1}^{k+1} x_i!} \cdot \prod_{i=1}^{k+1} p_i^{x_i} = \left( \sum_{i=1}^{k+1} p_i \right)^n = \left( \sum_{i=1}^k p_i + 1 - \sum_{i=1}^k p_i \right)^n = 1^n = 1 \end{aligned}$$

Además, como todos los términos de la suma son no negativos, tenemos efectivamente que se trata de una función de masa de probabilidad.  $\square$

Calculemos ahora la función generatriz de momentos de la distribución multinomial.

**Proposición 1.3.** Sea  $X \sim M_k(n, p_1, \dots, p_k)$ . Entonces, la función generatriz de momentos de  $X$  es:

$$M_X(t_1, \dots, t_k) = \left[ \left( \sum_{i=1}^k p_i e^{t_i} \right) + \left( 1 - \sum_{i=1}^k p_i \right) \right]^n \quad \forall t_1, \dots, t_k \in \mathbb{R}$$

*Demostración.* Usaremos la definición de función generatriz de momentos:

$$\begin{aligned}
M_X(t_1, \dots, t_k) &= E[e^{t_1 X_1 + \dots + t_k X_k}] = \\
&= \sum_{\sum_{i=1}^k x_i \leq n} e^{t_1 x_1 + \dots + t_k x_k} \cdot \frac{n!}{\left(\prod_{i=1}^k x_i!\right) \cdot \left(n - \sum_{i=1}^k x_i\right)!} \cdot \prod_{i=1}^k p_i^{x_i} \cdot \left(1 - \sum_{i=1}^k p_i\right)^{n - \sum_{i=1}^k x_i} = \\
&= \sum_{\sum_{i=1}^k x_i \leq n} \prod_{i=1}^k (e^{t_i})^{x_i} \cdot \frac{n!}{\left(\prod_{i=1}^k x_i!\right) \cdot \left(n - \sum_{i=1}^k x_i\right)!} \cdot \prod_{i=1}^k p_i^{x_i} \cdot \left(1 - \sum_{i=1}^k p_i\right)^{n - \sum_{i=1}^k x_i} = \\
&= \sum_{\sum_{i=1}^k x_i \leq n} \frac{n!}{\left(\prod_{i=1}^k x_i!\right) \cdot \left(n - \sum_{i=1}^k x_i\right)!} \cdot \prod_{i=1}^k (p_i e^{t_i})^{x_i} \cdot \left(1 - \sum_{i=1}^k p_i\right)^{n - \sum_{i=1}^k x_i} \stackrel{(*)}{=} \\
&\stackrel{(*)}{=} \left[ \left( \sum_{i=1}^k p_i e^{t_i} \right) + \left( 1 - \sum_{i=1}^k p_i \right) \right]^n
\end{aligned}$$

donde en (\*) hemos usado la Fórmula de Leibniz.  $\square$

Calculamos ahora las marginales de la distribución multinomial.

**Proposición 1.4.** *Dado el vector aleatorio  $X \sim M_k(n, p_1, \dots, p_k)$ , consideramos un subvector de  $X$   $(X_{i_1}, \dots, X_{i_l})$ , con  $i_1, \dots, i_l \in \{1, \dots, k\}$ , siendo  $i_p \neq i_q$  si  $p \neq q$ . Entonces,*

$$(X_{i_1}, \dots, X_{i_l}) \sim M_l(n, p_{i_1}, \dots, p_{i_l}).$$

*Observación.* Notemos que esta notación no es del todo precisa, pero se emplea para no complicar la notación. La distribución marginal de un subvector aleatorio de  $l$  componentes sigue una distribución multinomial con los parámetros  $n$  y las probabilidades asociadas a las componentes del subvector.

*Demostración.* Esto se deduce directamente de la función generatriz de momentos de un subvector de  $X$ , que es:

$$\begin{aligned}
M_{(X_{i_1}, \dots, X_{i_l})}(t_{i_1}, \dots, t_{i_l}) &= M_X(0, \dots, 0, t_{i_1}, 0, \dots, 0, t_{i_l}, 0, \dots, 0) = \\
&= \left[ \left( \sum_{i=1}^k p_i e^{t_i} \right) + \left( 1 - \sum_{i=1}^k p_i \right) \right]^n = \\
&= \left[ \left( \sum_{j=1}^l p_{i_j} e^{t_{i_j}} \right) + \left( 1 - \sum_{j=1}^l p_{i_j} \right) \right]^n
\end{aligned}$$

Como esta función generatriz de momentos coincide con la de una distribución multinomial con los parámetros  $n$  y las probabilidades asociadas a las componentes del subvector, se tiene que la distribución marginal de un subvector de  $l$  componentes de  $X$  sigue una distribución multinomial con los parámetros  $n$  y las probabilidades asociadas a las componentes del subvector.  $\square$

Tenemos que, en particular, para  $j \in \{1, \dots, k\}$ , tenemos que:

$$X_j \sim M_1(n, p_j)$$

Su función generatriz de momentos entonces es:

$$M_{X_j}(t) = [p_j e^t + (1 - p_j)]^n$$

Por la unicidad de la función generatriz de momentos, tenemos que  $X_j \sim B(n, p_j)$ , como introducimos ya anteriormente.

Calculamos ahora las distribuciones condicionadas de la distribución multinomial.

**Proposición 1.5.** *Consideramos el vector aleatorio  $X \sim M_k(n, p_1, \dots, p_k)$ . Sean los conjuntos disjuntos  $\{i_1, \dots, i_q\}, \{j_1, \dots, j_p\} \subset \{1, \dots, k\}$ , con  $i_p \neq i_q$  y  $j_p \neq j_q$  si  $p \neq q$ . Entonces:*

$$(X_{i_1}, \dots, X_{i_q}) \mid (X_{j_1} = x_{j_1}, \dots, X_{j_p} = x_{j_p}) \sim M_q \left( n - \sum_{l=1}^p x_{j_l}, \frac{p_{i_1}}{1 - \sum_{l=1}^p p_{j_l}}, \dots, \frac{p_{i_q}}{1 - \sum_{l=1}^p p_{j_l}} \right)$$

*Observación.* De nuevo, la notación no es del todo precisa, pero se emplea para no complicar la notación.

La distribución condicionada de un subvector de  $q$  componentes de  $X$  dado que otro subvector de  $p$  componentes de  $X$  toma unos valores concretos, sigue una distribución multinomial con los parámetros  $n$  menos la suma de los valores que toman las componentes del subvector condicionante, y las probabilidades asociadas a las componentes del subvector condicionado divididas por la probabilidad 1 menos la suma de las probabilidades asociadas a las componentes del subvector condicionante.

En el caso de que condicionemos una componente  $X_i$  de  $X$  a otra componente  $X_j$  de  $X$  ( $i \neq j$ ), tenemos que:

$$X_i \mid X_j = x_j \sim M_1 \left( n - x_j, \frac{p_i}{1 - p_j} \right) = B \left( n - x_j, \frac{p_i}{1 - p_j} \right)$$

Veamos ahora que se trata de una distribución reproductiva.

**Proposición 1.6** (Distribución Reproductiva - Multinomial). *Fijado  $p \in \mathbb{N}$ , consideramos  $p$  vectores aleatorios  $X_1, \dots, X_p$  independientes tales que  $X_i \sim M_k(n_i, p_1, \dots, p_k)$ ,  $i = 1, \dots, p$ . Se tiene entonces que:*

$$\sum_{i=1}^p X_i \sim M_k \left( \sum_{i=1}^p n_i, p_1, \dots, p_k \right)$$

*Demostración.* Por ser independientes, tenemos que:

$$\begin{aligned} M_{\sum_{i=1}^p X_i}(t_1, \dots, t_k) &= \prod_{i=1}^p M_{X_i}(t_1, \dots, t_k) = \prod_{i=1}^p \left[ \left( \sum_{j=1}^k p_j e^{t_j} \right) + \left( 1 - \sum_{j=1}^k p_j \right) \right]^{n_i} = \\ &= \sum_{j=1}^k \left[ \left( \sum_{j=1}^k p_j e^{t_j} \right) + \left( 1 - \sum_{j=1}^k p_j \right) \right]^{\sum_{i=1}^p n_i} \end{aligned}$$

Por tanto, la función generatriz de momentos de  $\sum_{i=1}^p X_i$  coincide con la de una distribución multinomial con los parámetros  $\sum_{i=1}^p n_i$  y las probabilidades asociadas a las componentes de  $X_i$ .  $\square$

Veamos ahora la esperanza de la distribución multinomial.

**Proposición 1.7.** *Sea  $X \sim M_k(n, p_1, \dots, p_k)$ . Entonces, la esperanza de  $X$  es:*

$$E[X] = n(p_1, \dots, p_k)$$

*Demostración.* De la definición de esperanza, si  $X = (X_1, \dots, X_k)$ , tenemos que:

$$E[X] = (E[X_1], \dots, E[X_k])$$

Como hemos visto anteriormente, para cada  $j \in \{1, \dots, k\}$ , tenemos que  $X_j \sim B(n, p_j)$ , por lo que  $E[X_j] = n \cdot p_j$ . Por tanto, tenemos que:

$$E[X] = (n \cdot p_1, \dots, n \cdot p_k) = n(p_1, \dots, p_k)$$

$\square$

Veamos ahora la covarianza de cada par de componentes de la distribución multinomial.

**Proposición 1.8.** *Sea  $X \sim M_k(n, p_1, \dots, p_k)$ . Entonces, si  $X = (X_1, \dots, X_k)$ , se tiene que:*

$$\text{Cov}(X_i, X_j) = -n \cdot p_i p_j \quad \forall i, j \in \{1, \dots, k\}, i \neq j$$

*Demostración.* De la definición de covarianza, tenemos que:

$$\text{Cov}(X_i, X_j) = E[X_i \cdot X_j] - E[X_i] \cdot E[X_j]$$

Para calcular  $E[X_i \cdot X_j]$ , usamos la función generatriz de momentos de  $X$ :

$$\begin{aligned} E[X_i \cdot X_j] &= \frac{\partial^2 M_X(t_1, \dots, t_k)}{\partial t_i \partial t_j} \Big|_{t_1=\dots=t_k=0} = \\ &= n(n-1)p_i p_j e^{t_i} e^{t_j} \left[ \left( \sum_{i=1}^k p_i e^{t_i} \right) + \left( 1 - \sum_{i=1}^k p_i \right) \right]^{n-2} \Big|_{t_1=\dots=t_k=0} = \\ &= n(n-1)p_i p_j \left[ \left( \sum_{i=1}^k p_i \right) + \left( 1 - \sum_{i=1}^k p_i \right) \right]^{n-2} = n(n-1)p_i p_j \end{aligned}$$

Por tanto, como  $E[X_i] = n \cdot p_i$ ,  $E[X_j] = n \cdot p_j$ , tenemos que:

$$\text{Cov}(X_i, X_j) = n(n-1)p_i p_j - n^2 p_i p_j = -n \cdot p_i \cdot p_j$$

$\square$

## 1.2. Distribución Normal Bidimensional

Generalizamos ahora la distribución normal a dos dimensiones.

**Definición 1.2** (Distribución Normal Bidimensional). Consideramos un vector aleatorio bidimensional continuo  $X = (X_1, X_2)$  y los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ . Consideramos además las siguientes matrices:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$\mu = (\mu_1 \quad \mu_2)$$

donde  $\sigma_i = \sqrt{\sigma_i^2}$ ,  $i = 1, 2$ . Diremos que  $X$  sigue una distribución normal bidimensional con parámetros  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  y  $\rho$  si su función de densidad es:

$$f_X(x) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{(x-\mu)\Sigma^{-1}(x-\mu)^t}{2}\right) \quad x \in \mathbb{R}^2$$

La notaremos por las siguientes dos formas:

$$X \sim \mathcal{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \quad \text{o} \quad X \sim \mathcal{N}_2(\mu, \Sigma)$$

donde  $\mu, \Sigma$  son las matrices definidas anteriormente.

Veamos que está bien definida. En primer lugar, hemos de calcular  $|\Sigma|$ :

$$|\Sigma| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1 - \rho^2) > 0$$

Por tanto,  $\Sigma$  es singular y, por tanto,  $\exists \Sigma^{-1}$ . Calculémosla:

$$\Sigma^{-1} = \frac{1}{|\Sigma|} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$

Calculamos ahora el producto de matrices descrito, considerando  $x = (x_1, x_2)$ :

$$\begin{aligned} (x-\mu)\Sigma^{-1}(x-\mu)^t &= \frac{1}{|\Sigma|} \cdot (x_1 - \mu_1 \quad x_2 - \mu_2) \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} = \\ &= \frac{1}{|\Sigma|} \cdot (x_1 - \mu_1 \quad x_2 - \mu_2) \begin{pmatrix} \sigma_2^2(x_1 - \mu_1) - \rho\sigma_1\sigma_2(x_2 - \mu_2) \\ -\rho\sigma_1\sigma_2(x_1 - \mu_1) + \sigma_1^2(x_2 - \mu_2) \end{pmatrix} = \\ &= \frac{1}{\sigma_1^2\sigma_2^2 \cdot (1 - \rho^2)} \cdot [\sigma_2^2(x_1 - \mu_1)^2 - 2\rho\sigma_1\sigma_2(x_1 - \mu_1)(x_2 - \mu_2) + \sigma_1^2(x_2 - \mu_2)^2] = \\ &= \frac{1}{1 - \rho^2} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x_1 - \mu_1}{\sigma_1} \right) \left( \frac{x_2 - \mu_2}{\sigma_2} \right) + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right] \end{aligned}$$

Por tanto, la función de densidad de la distribución normal bidimensional es:

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left( \frac{x_1 - \mu_1}{\sigma_1} \right) \left( \frac{x_2 - \mu_2}{\sigma_2} \right) + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right]\right)$$

Veamos ahora distintas expresiones alternativas para la función de densidad de la distribución normal bidimensional. Multiplicando el término de  $\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2$  por  $(1 - \rho^2 + \rho^2)$ , el exponente queda:

$$\begin{aligned} & -\frac{1}{2(1 - \rho^2)} \left[ \left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 (1 - \rho^2 + \rho^2) - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1}\right) \left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 \right] = \\ & = -\frac{(x_1 - \mu_1)^2}{2\sigma_1^2} - \frac{1}{2(1 - \rho^2)} \left[ \rho^2 \left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1}\right) \left(\frac{x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 \right] = \\ & = -\frac{(x_1 - \mu_1)^2}{2\sigma_1^2} - \frac{1}{2(1 - \rho^2)} \left[ -\rho \left(\frac{x_1 - \mu_1}{\sigma_1}\right) + \left(\frac{x_2 - \mu_2}{\sigma_2}\right) \right]^2 = \\ & = -\frac{(x_1 - \mu_1)^2}{2\sigma_1^2} - \frac{1}{2\sigma_2^2(1 - \rho^2)} \left[ -\rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1 - \mu_1) + (x_2 - \mu_2) \right]^2 \end{aligned}$$

Por tanto, podemos descomponer la función de densidad de la distribución normal bidimensional en dos funciones:

$$f_X(x_1, x_2) = \overbrace{\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1 - \rho^2}} \exp \left( -\frac{1}{2\sigma_2^2(1 - \rho^2)} \left[ -\rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1 - \mu_1) + (x_2 - \mu_2) \right]^2 \right)}^{h_{x_1}(x_2)} \cdot \underbrace{\frac{1}{\sqrt{2\pi}\sigma_1} \exp \left( -\frac{(x_1 - \mu_1)^2}{2\sigma_1^2} \right)}_{g(x_1)} \quad (1.1)$$

Tenemos que  $g(x_1)$  es la función de densidad de  $\mathcal{N}(\mu_1, \sigma_1^2)$  y, para cada  $x_1$ ,  $h_{x_1}(x_2)$  es la función de densidad de  $\mathcal{N}(\mu_2 + \rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1 - \mu_1), \sigma_2^2(1 - \rho^2))$ . Esta descomposición es útil, y la emplearemos en diversas demostraciones en el futuro. Se podrá considerar también la descomposición análoga.

Para terminar de aceptar dicha definición, hemos de comprobar que, efectivamente, se trata de una función de densidad. Como  $f_X(x) \geq 0 \forall x \in \mathbb{R}^2$ , tan solo nos falta por comprobar que:

$$\begin{aligned} \int_{\mathbb{R}^2} f_X(x) dx &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_X(x_1, x_2) dx_2 dx_1 \stackrel{(*)}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x_1) h_{x_1}(x_2) dx_2 dx_1 = \\ &= \int_{-\infty}^{+\infty} g(x_1) dx_1 \int_{-\infty}^{+\infty} h_{x_1}(x_2) dx_2 \stackrel{(**)}{=} \int_{-\infty}^{+\infty} g(x_1) dx_1 \stackrel{(**)}{=} 1 \end{aligned}$$

donde en  $(*)$  hemos empleado la descomposición de la Ecuación 1.1 y en  $(**)$  hemos usado que  $g(x_1)$  y  $h_{x_1}(x_2)$  son funciones de densidad.

Veamos ahora las marginales de la distribución normal bidimensional.

**Proposición 1.9.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $(X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces:

$$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad y \quad X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

*Demostración.* Calculamos la marginal de  $X_1$ :

$$f_{X_1}(x_1) = \int_{-\infty}^{+\infty} f_X(x_1, x_2) dx_2 \stackrel{(*)}{=} \int_{-\infty}^{+\infty} g(x_1) h_{x_1}(x_2) dx_2 = g(x_1) \int_{-\infty}^{+\infty} h_{x_1}(x_2) dx_2 \stackrel{(**)}{=} g(x_1)$$

donde en  $(*)$  hemos empleado la descomposición de la Ecuación 1.1 y en  $(**)$  hemos usado que  $h_{x_1}(x_2)$  es una función de densidad. Por tanto, como vimos que  $g(x_1)$  es la función de densidad de  $\mathcal{N}(\mu_1, \sigma_1^2)$ , tenemos que  $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ . Usando la descomposición análoga, se demuestra que  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ .  $\square$

Veamos ahora las condicionadas de la distribución normal bidimensional.

**Proposición 1.10.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $(X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces, dado  $x_1^*, x_2^* \in \mathbb{R}$ , se tiene que:

$$\begin{aligned} X_1 \mid X_2 = x_2^* &\sim \mathcal{N}\left(\mu_1 + \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot (x_2^* - \mu_2), \sigma_1^2(1 - \rho^2)\right) \\ X_2 \mid X_1 = x_1^* &\sim \mathcal{N}\left(\mu_2 + \rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1^* - \mu_1), \sigma_2^2(1 - \rho^2)\right) \end{aligned}$$

*Demostración.* Dado  $x_1^* \in \mathbb{R}$ , calculamos la condicionada de  $X_2$ :

$$f_{X_2 \mid X_1 = x_1^*}(x_2) = \frac{f_X(x_1^*, x_2)}{f_{X_1}(x_1^*)} \stackrel{(*)}{=} \frac{g(x_1^*) h_{x_1^*}(x_2)}{g(x_1^*)} = h_{x_1^*}(x_2)$$

donde en  $(*)$  hemos empleado la descomposición de la Ecuación 1.1. Como  $h_{x_1^*}(x_2)$  es la función de densidad de  $\mathcal{N}\left(\mu_2 + \rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1^* - \mu_1), \sigma_2^2(1 - \rho^2)\right)$ , tenemos que  $X_2 \mid X_1 = x_1^* \sim \mathcal{N}\left(\mu_2 + \rho \cdot \frac{\sigma_2}{\sigma_1} \cdot (x_1^* - \mu_1), \sigma_2^2(1 - \rho^2)\right)$ .  $\square$

*Observación.* Una vez llegados aquí, podemos olvidarnos de la descomposición de la función de densidad de la distribución normal bidimensional descrita en la Ecuación 1.1. Esta descomposición en verdad era:

$$f_X(x_1, x_2) = f_{X_1}(x_1) f_{X_2 \mid X_1 = x_1}(x_2)$$

Esta descomposición ya es conocida por el lector.

Veamos ahora la función generatriz de momentos de la distribución normal bidimensional.

**Proposición 1.11.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $(X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces, la función generatriz de momentos de  $X_1$  y  $X_2$  es:

$$\begin{aligned} M_X(t) &= \exp\left(t\mu + \frac{t\Sigma t^t}{2}\right) = \\ &= \exp\left(t_1\mu_1 + t_2\mu_2 + \frac{t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2t_1t_2\rho\sigma_1\sigma_2}{2}\right) \quad \forall t = (t_1, t_2) \in \mathbb{R}^2 \end{aligned}$$

*Demostración.* Sea  $t = (t_1, t_2) \in \mathbb{R}^2$ . Calculamos la función generatriz de momentos de  $X$ :

$$\begin{aligned} M_X(t) &= E[e^{t_1 X_1 + t_2 X_2}] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{t_1 x_1 + t_2 x_2} f_X(x_1, x_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{t_1 x_1} e^{t_2 x_2} f_{X_2}(x_2) f_{X_1|X_2=x_2}(x_1) dx_1 dx_2 = \\ &= \int_{-\infty}^{+\infty} e^{t_2 x_2} f_{X_2}(x_2) \left[ \int_{-\infty}^{+\infty} e^{t_1 x_1} f_{X_1|X_2=x_2}(x_1) dx_1 \right] dx_2 \end{aligned}$$

Como  $X_1 | X_2 = x_2 \sim \mathcal{N}\left(\mu_1 + \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot (x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\right)$ , usando la función generatriz de momentos de la distribución normal evaluada en  $t_1$ , tenemos que:

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{+\infty} e^{t_2 x_2} f_{X_2}(x_2) M_{X_1|X_2=x_2}(t_1) dx_2 = \\ &= \int_{-\infty}^{+\infty} e^{t_2 x_2} f_{X_2}(x_2) \exp \left[ t_1 \left( \mu_1 + \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot (x_2 - \mu_2) \right) + \frac{t_1^2 \sigma_1^2 (1 - \rho^2)}{2} \right] dx_2 = \\ &= \exp \left[ t_1 \left( \mu_1 - \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 \right) + \frac{t_1^2 \sigma_1^2 (1 - \rho^2)}{2} \right] \int_{-\infty}^{+\infty} e^{t_2 x_2} f_{X_2}(x_2) \exp \left[ t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot x_2 \right] dx_2 = \\ &= \exp \left[ t_1 \left( \mu_1 - \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 \right) + \frac{t_1^2 \sigma_1^2 (1 - \rho^2)}{2} \right] \int_{-\infty}^{+\infty} \exp \left[ \left( t_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \right) x_2 \right] f_{X_2}(x_2) dx_2 \end{aligned}$$

Como  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ , usando la función generatriz de momentos de la distribución normal evaluada en  $t_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2}$ , tenemos que:

$$\begin{aligned} M_X(t) &= \exp \left[ t_1 \left( \mu_1 - \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 \right) + \frac{t_1^2 \sigma_1^2 (1 - \rho^2)}{2} \right] M_{X_2} \left( t_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \right) \\ &= \exp \left[ t_1 \left( \mu_1 - \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 \right) + \frac{t_1^2 \sigma_1^2 (1 - \rho^2)}{2} \right] \exp \left[ \left( t_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \right) \mu_2 + \frac{\left( t_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \right)^2 \sigma_2^2}{2} \right] \\ &= \exp \left[ t_1 \mu_1 - t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 + \frac{t_1^2 \sigma_1^2}{2} - \rho^2 \cdot \frac{t_1^2 \sigma_1^2}{2} + t_2 \mu_2 + t_1 \rho \cdot \frac{\sigma_1}{\sigma_2} \cdot \mu_2 + \frac{t_2^2 \sigma_2^2 + \cancel{t_1^2 \rho^2 \sigma_1^2} + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right] \\ &= \exp \left[ t_1 \mu_1 + t_2 \mu_2 + \frac{t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right] \end{aligned}$$

Calculemos ahora el producto de matrices  $t \Sigma t^t$ :

$$\begin{aligned} t \Sigma t^t &= (t_1 \ t_2) \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = (t_1 \ t_2) \begin{pmatrix} \sigma_1^2 t_1 + \rho \sigma_1 \sigma_2 t_2 \\ \rho \sigma_1 \sigma_2 t_1 + \sigma_2^2 t_2 \end{pmatrix} = \\ &= t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2 \end{aligned}$$

Llegando así a las dos expresiones de la función generatriz de momentos de la distribución normal bidimensional.  $\square$

Una vez estudiadas las marginales y la función generatriz de momentos, estamos en condiciones de ver por qué los parámetros de la distribución normal bidimensional



se denominan  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  y  $\rho$ , y por qué se definen las matrices  $\mu$  y  $\Sigma$  asociadas a estos parámetros. Veamos en primer lugar que, como era de esperar,  $\mu$  es el vector esperanza de la distribución normal bidimensional.

**Proposición 1.12.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $X = (X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces:

$$E[X] = \mu = \begin{pmatrix} \mu_1 & \mu_2 \end{pmatrix}$$

*Demostración.* De la definición de esperanza, tenemos que:

$$E[X] = \begin{pmatrix} E[X_1] & E[X_2] \end{pmatrix}$$

Considerando las marginales, como  $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$  y  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ , tenemos que  $E[X_1] = \mu_1$  y  $E[X_2] = \mu_2$ . Por tanto, tenemos que:

$$E[X] = \begin{pmatrix} \mu_1 & \mu_2 \end{pmatrix} = \mu$$

□

Veamos ahora que la matriz  $\Sigma$  es la matriz de covarianzas de la distribución normal bidimensional.

**Proposición 1.13.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $X = (X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces:

$$\text{Cov}_X = \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

*Demostración.* De la definición de covarianza, tenemos que:

$$\text{Cov}_X = \begin{pmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) \end{pmatrix}$$

Como la covarianza de una variable consigo misma es su varianza y la covarianza es simétrica, tenemos que:

$$\text{Cov}_X = \begin{pmatrix} \text{Var}[X_1] & \text{Cov}(X_1, X_2) \\ \text{Cov}(X_1, X_2) & \text{Var}[X_2] \end{pmatrix}$$

Considerando las marginales, como  $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$  y  $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ , tenemos que  $\text{Var}[X_1] = \sigma_1^2$  y  $\text{Var}[X_2] = \sigma_2^2$ . Para calcular la covarianza de  $X_1$  y  $X_2$ , consideramos la función generatriz de momentos de  $X$ :

$$\begin{aligned} E[XY] &= \frac{\partial^2 M_X(t)}{\partial t_1 \partial t_2} \Big|_{t_1=t_2=0} = \frac{\partial^2}{\partial t_1 \partial t_2} \exp \left( t_1 \mu_1 + t_2 \mu_2 + \frac{t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right) \Big|_{t_1=t_2=0} = \\ &= \frac{\partial}{\partial t_2} \exp \left( t_1 \mu_1 + t_2 \mu_2 + \frac{t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right) (\mu_1 + t_1 \sigma_1^2 + \rho \sigma_1 \sigma_2 t_2) \Big|_{t_1=t_2=0} = \\ &= \exp \left( t_1 \mu_1 + t_2 \mu_2 + \frac{t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right) (\mu_2 + t_2 \sigma_2^2 + \rho \sigma_1 \sigma_2 t_1) (\mu_1 + t_1 \sigma_1^2 + \rho \sigma_1 \sigma_2 t_2) + \\ &\quad + \exp \left( t_1 \mu_1 + t_2 \mu_2 + \frac{t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 + 2t_1 t_2 \rho \sigma_1 \sigma_2}{2} \right) (\rho \sigma_1 \sigma_2) \Big|_{t_1=t_2=0} = \\ &= \exp(0) (\mu_2) (\mu_1) + \exp(0) (\rho \sigma_1 \sigma_2) = \mu_1 \mu_2 + \rho \sigma_1 \sigma_2 \end{aligned}$$

Por tanto, tenemos que:

$$\text{Cov}(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2] = \mu_1 \mu_2 + \rho \sigma_1 \sigma_2 - \mu_1 \mu_2 = \rho \sigma_1 \sigma_2$$

Por tanto, hemos demostrado que  $\text{Cov}_X = \Sigma$ .  $\square$

Por último, interpretemos el parámetro  $\rho$  de la distribución normal bidimensional.

**Proposición 1.14.** *Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $X = (X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces  $\rho$  es el coeficiente de correlación de  $X_1$  y  $X_2$ .*

$$\rho_{X_1, X_2} = \rho$$

*Demostración.* De la definición de Coeficiente de Correlación, tenemos que:

$$\rho_{X_1, X_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}[X_1] \text{Var}[X_2]}}$$

Usando que  $\Sigma$  es la matriz de covarianzas de  $X$ , tenemos que:

$$\rho_{X_1, X_2} = \frac{\rho \sigma_1 \sigma_2}{\sqrt{\sigma_1^2 \sigma_2^2}} = \rho$$

$\square$

Por tanto, ya hemos interpretado todos los parámetros de la distribución normal bidimensional. Por tanto, de aquí en adelante no es necesario prefijar  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  y  $\rho$  para definir la distribución normal bidimensional, sino que podemos definirla directamente con  $\mu$  y  $\Sigma$ , siendo la primera un vector esperanza y la segunda una matriz de covarianzas. De ahí la notación de la distribución normal bidimensional  $\mathcal{N}_2(\mu, \Sigma)$ .

Estudiemos ahora las combinaciones lineales de variables aleatorias normales bidimensionales.

**Proposición 1.15.** *Fijado  $q \in \{1, 2\}$ , consideramos la matriz  $A \in \mathcal{M}_{2 \times q}(\mathbb{R})$  de rango  $q$ . Sea también  $X = (X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ . Entonces, tenemos que:*

$$Y := XA \sim \mathcal{N}_q(\mu A, A^t \Sigma A)$$

*Demostración.* Como  $A \in \mathcal{M}_{2 \times q}(\mathbb{R})$  de rango  $q$ , tenemos que  $XA \in \mathcal{M}_{1 \times q}(\mathbb{R})$ , por lo que  $Y$  es un vector aleatorio de dimensión  $q$ . Dado  $t \in \mathbb{R}^q$ , calculamos la función generatriz de momentos de  $Y$ :

$$\begin{aligned} M_Y(t) &= E[e^{\langle t, Y \rangle}] = E[e^{\langle t, XA \rangle}] = E[e^{\langle X, tA^t \rangle}] = M_X(tA^t) = \exp\left(\langle tA^t, \mu \rangle + \frac{tA^t \Sigma (tA^t)^t}{2}\right) = \\ &= \exp\left(\langle t, \mu A \rangle + \frac{tA^t \Sigma A t^t}{2}\right) \end{aligned}$$

Por tanto, tenemos que  $Y \sim \mathcal{N}_q(\mu A, A^t \Sigma A)$  (Nótese que, si  $q = 1$ , también se trata de la función generatriz de momentos de la distribución normal unidimensional, donde usamos el producto escalar en  $\mathbb{R}$ ).  $\square$

Por último, caractericemos la independencia en el caso de la distribución normal bidimensional.

**Proposición 1.16.** Sean los parámetros  $\mu_1, \mu_2 \in \mathbb{R}$ ,  $\sigma_1^2, \sigma_2^2 \in \mathbb{R}^+$  y  $\rho \in ]-1, 1[$ , y sean sus matrices  $\mu$  y  $\Sigma$  asociadas. Si  $X = (X_1, X_2) \sim \mathcal{N}_2(\mu, \Sigma)$ , entonces:

$$X_1 \text{ y } X_2 \text{ son independientes} \iff \rho = 0$$

*Demostración.* Demostramos mediante doble implicación:

$\implies$ ) Supongamos que  $X_1$  y  $X_2$  son independientes. Entonces, vimos en su momento que  $\text{Cov}(X_1, X_2) = 0$ , por lo que  $\rho_{X_1, X_2} = 0$ . Además, como  $\rho_{X_1, X_2} = \rho$ , tenemos que  $\rho = 0$ .

$\impliedby$ ) Supongamos que  $\rho = 0$ . Entonces, la función de densidad de  $X$  queda:

$$\begin{aligned} f_X(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2} \left[ \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right]\right) = \\ &= \underbrace{\frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2}\right)}_{g_1(x_1)} \underbrace{\exp\left(-\frac{(x_2 - \mu_2)^2}{2\sigma_2^2}\right)}_{g_2(x_2)} \end{aligned}$$

De esta forma, tenemos que  $f_X(x_1, x_2) = g_1(x_1)g_2(x_2)$  para determinadas funciones  $g_1$  y  $g_2$ . Por tanto,  $X_1$  y  $X_2$  son independientes.

□

Notemos que, como ya hemos visto en otros casos, esta equivalencia no es cierta en general para variables aleatorias normales bidimensionales.



## 2. Relaciones de problemas

## 2.1. Algunos Modelos Multivariantes

**Ejercicio 2.1.1.** El 60 % de los clientes de un almacén paga con dinero, el 30 % con tarjeta y el 10 % con cheques. Calcular la probabilidad de que de 10 clientes, 5 paguen con dinero, 2 con tarjeta y 3 con cheques.

Dado  $i \in \{\text{dinero, tarjeta, cheques}\}$ , definimos  $X_i$  como la variable aleatoria que cuenta el número de clientes que pagan con el método  $i$ . Entonces,  $(X_{\text{dinero}}, X_{\text{tarjeta}}, X_{\text{cheques}})$  sigue una distribución multinomial con parámetros  $n = 10$  y probabilidades  $(p_{\text{dinero}}, p_{\text{tarjeta}}, p_{\text{cheques}})$ , donde:

$$p_{\text{dinero}} = 0,6, \quad p_{\text{tarjeta}} = 0,3, \quad p_{\text{cheques}} = 0,1.$$

Tenemos entonces que:

$$\begin{aligned} P[X_{\text{dinero}} = 5, X_{\text{tarjeta}} = 2, X_{\text{cheques}} = 3] &= \frac{10!}{5!2!3!} \cdot p_{\text{dinero}}^5 \cdot p_{\text{tarjeta}}^2 \cdot p_{\text{cheques}}^3 = \\ &= \frac{10!}{5!2!3!} \cdot 0,6^5 \cdot 0,3^2 \cdot 0,1^3 \approx 0,0176 \end{aligned}$$

**Ejercicio 2.1.2.** En un hotel hay tres salas de televisión. En un determinado instante, cada televisor puede sintonizar uno entre 6 canales distintos, A, B, C, D, E y F. Cada canal tiene probabilidad  $1/36, 3/36, 5/36, 7/36, 9/36$  y  $11/36$ , respectivamente, de ser sintonizado, con independencia unas televisiones de otras. Calcular:

1. Probabilidad de que en un instante dado se sintonicen los canales B, D y E.

Dado  $i \in \{A, B, C, D, E, F\}$ , definimos  $X_i$  como la variable aleatoria que cuenta el número de televisores sintonizando el canal  $i$ . Entonces,  $(X_A, X_B, X_C, X_D, X_E)$  sigue una distribución multinomial con parámetros  $n = 3$  y probabilidades  $(p_A, p_B, p_C, p_D, p_E)$ , donde

$$\begin{aligned} p_A &= \frac{1}{36}, & p_B &= \frac{3}{36}, \\ p_C &= \frac{5}{36}, & p_D &= \frac{7}{36}, & p_E &= \frac{9}{36}. \end{aligned}$$

Consideramos ahora el vector aleatorio  $(X_B, X_D, X_E)$ , que sabemos que:

$$(X_B, X_D, X_E) \sim M_3(3, p_B, p_D, p_E).$$

Tenemos que:

$$P[X_B = 1, X_D = 1, X_E = 1] = \frac{3!}{1!1!1!0!} \cdot p_B \cdot p_D \cdot p_E = 6 \cdot \frac{3}{36} \cdot \frac{7}{36} \cdot \frac{9}{36} = \frac{7}{288} \approx 0,0243$$

2. Probabilidad de que en un instante dado haya un televisor sintonizando el canal B y otro el E.

Consideramos ahora el vector aleatorio  $(X_B, X_E)$ , que sabemos que:

$$(X_B, X_E) \sim M_2(3, p_B, p_E).$$

Tenemos que:

$$P[X_B = 1, X_E = 1] = \frac{3!}{1!1!1!} \cdot p_B \cdot p_E \cdot (1 - p_B - p_E) = 6 \cdot \frac{3}{36} \cdot \frac{9}{36} \cdot \frac{24}{36} = \frac{1}{12} \approx 0,0833$$

3. Probabilidad de que en un instante dado los tres televisores sintonicen el canal F.

Consideramos ahora el vector aleatorio  $(X_F)$ , que sabemos que:

$$(X_F) \sim M_1(3, p_F).$$

Tenemos que:

$$P[X_F = 3] = \frac{3!}{3!0!} \cdot p_F^3 = 6 \cdot \left(\frac{11}{36}\right)^3 = \frac{1331}{7776} \approx 0,1711$$

4. Probabilidad de que en un instante dado no estén sintonizados A, B, C, D.

Consideramos ahora el vector aleatorio  $(X_A, X_B, X_C, X_D)$ , que sabemos que:

$$(X_A, X_B, X_C, X_D) \sim M_4(3, p_A, p_B, p_C, p_D).$$

Tenemos que:

$$\begin{aligned} P[X_A = 0, X_B = 0, X_C = 0, X_D = 0] &= \frac{3!}{0!0!0!3!} \cdot (1 - p_A - p_B - p_C - p_D)^3 = \\ &= 1 \cdot \left(1 - \frac{1}{36} - \frac{3}{36} - \frac{5}{36} - \frac{7}{36}\right)^3 = \left(\frac{5}{9}\right)^3 = \frac{125}{729} \approx 0,1715 \end{aligned}$$

**Ejercicio 2.1.3.** Un ordenador genera números aleatorios del 0 al 9 con independencia e igual probabilidad para cada dígito. Si se generan 12 números aleatorios, calcular:

1. La probabilidad de que aparezca 6 veces el dígito 0, 4 veces el 1 y 2 veces el 2.

Dado  $i \in \{0, 1, 2, \dots, 8\}$ , definimos  $X_i$  como la variable aleatoria que cuenta el número de veces que aparece el dígito  $i$ . Entonces,  $(X_0, X_1, X_2, \dots, X_8)$  sigue una distribución multinomial con parámetros  $n = 12$  y probabilidades  $(p_0, p_1, p_2, \dots, p_8)$ , donde:

$$p := p_0 = p_1 = p_2 = \dots = p_8 = \frac{1}{10}.$$

Consideramos ahora el vector aleatorio  $(X_0, X_1, X_2)$ , que sabemos que:

$$(X_0, X_1, X_2) \sim M_3(12, p_0, p_1, p_2).$$

Tenemos que:

$$P[X_0 = 6, X_1 = 4, X_2 = 2] = \frac{12!}{6!4!2!0!} \cdot p_0^6 \cdot p_1^4 \cdot p_2^2 = \frac{12!}{6!4!2!0!} p^{12} \approx 13,86 \cdot 10^{-9}$$

2. El número esperado de veces que aparece el 0.

Tenemos que  $X_0 \sim N_1(12, p_0)$ , por lo que:

$$E[X_0] = n \cdot p_0 = 12 \cdot \frac{1}{10} = \frac{6}{5} = 1,2$$

3. La probabilidad de que aparezca 4 veces el 1 y 3 veces el 6.

Consideramos ahora el vector aleatorio  $(X_1, X_6)$ , que sabemos que:

$$(X_1, X_6) \sim M_2(12, p_1, p_6).$$

Tenemos que:

$$P[X_1 = 4, X_6 = 3] = \frac{12!}{4!3!5!} \cdot p_1^4 \cdot p_6^3 \cdot (1-p_1-p_6)^5 = \frac{12!}{4!3!5!} p^7 (1-2p)^5 \approx 9,08 \cdot 10^{-4}$$

**Ejercicio 2.1.4.** Sea  $F : \mathbb{R} \rightarrow \mathbb{R}$  una función de distribución continua, y sean  $\alpha_1$  y  $\alpha_2$  números reales tales que  $F(\alpha_1) = 0,3$  y  $F(\alpha_2) = 0,8$ . Si se seleccionan al azar 25 observaciones independientes de la distribución cuya función de distribución es  $F$ . Calcular la probabilidad de que seis de los valores observados sean menores que  $\alpha_1$ , diez de los valores observados estén entre  $\alpha_1$  y  $\alpha_2$  y 9 sean mayores que  $\alpha_2$ .

Como  $F$  es una función de distribución, es creciente. Como  $F(\alpha_1) = 0,3$  y  $F(\alpha_2) = 0,8$ , tenemos que  $\alpha_1 \leq \alpha_2$ . Definimos las siguientes variables aleatorias:

$X_1 :=$  Número de observaciones menores que  $\alpha_1$ ,

$X_2 :=$  Número de observaciones entre  $\alpha_1$  y  $\alpha_2$ ,

$X_3 :=$  Número de observaciones mayores que  $\alpha_2$ .

Tenemos las siguientes probabilidades:

$p_1 =$  Prob. de que una observación sea menor que  $\alpha_1 = F(\alpha_1) = 0,3$ ,

$p_2 =$  Prob. de que una observación esté entre  $\alpha_1$  y  $\alpha_2 = F(\alpha_2) - F(\alpha_1) = 0,8 - 0,3 = 0,5$ ,

$p_3 =$  Prob. de que una observación sea mayor que  $\alpha_2 = 1 - F(\alpha_2) = 0,2 = 1 - p_1 - p_2$ .

Tenemos que:

$$(X_1, X_2) \sim M_2(25, p_1, p_2).$$

Por lo que la probabilidad pedida es:

$$P[X_1 = 6, X_2 = 10] = \frac{25!}{6!10!9!} \cdot p_1^6 \cdot p_2^{10} \cdot p_3^9 = \frac{25!}{6!10!9!} \cdot 0,3^6 \cdot 0,5^{10} \cdot 0,2^9 \approx 0,0059$$

**Ejercicio 2.1.5.** Si se lanzan 5 dados equilibrados de forma independiente, calcular la probabilidad de que los números 1 y 4 aparezcan el mismo número de veces.

Dado  $i \in \{1, 2, 3, 4, 5, 6\}$ , definimos  $X_i$  como la variable aleatoria que cuenta el número de veces que aparece el número  $i$ . Entonces,  $(X_1, X_2, X_3, X_4, X_5)$  sigue una distribución multinomial con parámetros  $n = 5$  y probabilidades  $(p_1, p_2, p_3, p_4, p_5)$ , donde:

$$p := p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{6}.$$

Consideramos ahora el vector aleatorio  $(X_1, X_4)$ , que sabemos que:

$$(X_1, X_4) \sim M_2(5, p_1, p_4).$$



Tenemos que la probabilidad pedida es:

$$P[X_1 = X_2] = \sum_{i=0}^2 P[X_1 = i, X_4 = i]$$

Calculamos cada uno de los sumandos:

$$P[X_1 = 0, X_4 = 0] = \frac{5!}{0!0!5!} \cdot p_1^0 \cdot p_4^0 \cdot (1 - p_1 - p_4)^5 = \frac{5!}{5!} \cdot \left(\frac{4}{6}\right)^5 = \frac{32}{243},$$

$$P[X_1 = 1, X_4 = 1] = \frac{5!}{1!1!3!} \cdot p_1^1 \cdot p_4^1 \cdot (1 - p_1 - p_4)^3 = 5 \cdot 4 \cdot \left(\frac{1}{6}\right) \cdot \left(\frac{1}{6}\right) \cdot \left(\frac{4}{6}\right)^3 = \frac{40}{243},$$

$$P[X_1 = 2, X_4 = 2] = \frac{5!}{2!2!1!} \cdot p_1^2 \cdot p_4^2 \cdot (1 - p_1 - p_4)^1 = \frac{5!}{2!2!} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{4}{6} = \frac{5}{324}.$$

Por lo que la probabilidad pedida es:

$$P[X_1 = X_2] = \frac{32}{243} + \frac{40}{243} + \frac{5}{324} = \frac{101}{324} \approx 0,312$$

**Ejercicio 2.1.6.** En un determinado juego de azar existen tres posibles resultados A, B y C, que se dan con probabilidad 0,8, 0,15 y 0,05, respectivamente. Una persona realiza 5 veces el juego de forma independiente, calcular la probabilidad de que no obtenga ninguna vez el resultado C ni más de una vez el resultado B.

Dado  $i \in \{A, B, C\}$ , sea  $p_i$  la probabilidad de obtener el resultado  $i$ . Definimos la variable aleatoria  $X_i$  como el número de veces que se obtiene el resultado  $i$ . Se pide:

$$P[X_B \leq 1, X_C = 0] = P[X_B = 0, X_C = 0] + P[X_B = 1, X_C = 0]$$

Para calcular cada una de las probabilidades, consideramos el vector aleatorio  $(X_B, X_C)$ , que sabemos que:

$$(X_B, X_C) \sim M_2(5, p_B, p_C).$$

Tenemos que:

$$P[X_B = 0, X_C = 0] = \frac{5!}{0!0!5!} \cdot p_B^0 \cdot p_C^0 \cdot (1 - p_B - p_C)^5 = \frac{5!}{5!} \cdot 0,8^5 = \frac{1024}{3125},$$

$$P[X_B = 1, X_C = 0] = \frac{5!}{1!0!4!} \cdot p_B^1 \cdot p_C^0 \cdot (1 - p_B - p_C)^4 = 5 \cdot 0,15 \cdot 0,8^4 = \frac{192}{625}.$$

Por lo que la probabilidad pedida es:

$$P[X_B \leq 1, X_C = 0] = \frac{1024}{3125} + \frac{192}{625} = \frac{1984}{3125} \approx 0,63488$$

**Ejercicio 2.1.7.** Sea  $(X, Y)$  un vector aleatorio con distribución normal con parámetros  $\mu_1 = 5$ ,  $\mu_2 = 8$ ,  $\sigma_1^2 = 16$ ,  $\sigma_2^2 = 9$ ,  $\rho = 0,6$ . Calcular

$$P[5 < Y < 11 \mid X = 2].$$

Calculemos en primer lugar la distribución condicional de  $Y$  dado  $X = 2$ . Sabemos que:

$$Y | X = 2 \sim \mathcal{N}\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(2 - \mu_1), \sigma_2^2(1 - \rho^2)\right) = \mathcal{N}(6,65, 5,76).$$

Por tanto, y notando por  $Z$  a la tipificada de  $Y | X = 2$ , tenemos que:

$$\begin{aligned} P[5 < Y < 11 | X = 2] &= P[-0,6875 < Z < 1,8125] = P[Z < 1,8125] - P[Z > -0,6875] = \\ &= P[Z < 1,8125] - P[Z < 0,6875] \approx 0,96485 - 0,75490 = 0,20995. \end{aligned}$$

**Ejercicio 2.1.8.** Calcular la distribución de probabilidad de la suma  $X + Y$  para  $(X, Y) \sim \mathcal{N}((\mu_1, \mu_2), \Sigma)$ , siendo

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix}.$$

**Ejercicio 2.1.9.** Si  $(X, Y)$  es un vector aleatorio con distribución normal, encontrar una condición necesaria y suficiente para que  $X + Y$  y  $X - Y$  sean independientes.

**Ejercicio 2.1.10.** Para cada una de las siguientes densidades normales bidimensionales, hallar  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  y  $\rho$ :

1.  $f_{(X,Y)}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}((x-1)^2 + (y-2)^2)\right)$
2.  $f_{(X,Y)}(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 + y^2 + 4x - 6y + 13)\right)$

**Ejercicio 2.1.11.** Sea  $(X, Y)$  un vector aleatorio con función de densidad en  $\mathbb{R}^2$  dada por:

$$\begin{aligned} f_{(X,Y)}(x, y) &= f_1(x, y) + f_2(x, y), \\ f_1(x, y) &= \frac{1}{4\pi} \sqrt{1 - \rho^2} \exp\left(-\frac{1}{2}(1 - \rho^2)(x^2 + y^2 - 2\rho xy)\right), \quad |\rho| < 1, \\ f_2(x, y) &= \frac{1}{4\pi} \sqrt{1 - \tau^2} \exp\left(-\frac{1}{2}(1 - \tau^2)(x^2 + y^2 - 2\tau xy)\right), \quad |\tau| < 1. \end{aligned}$$

1. Demostrar que las distribuciones marginales son normales de media cero y varianza uno.
2. Obtener la covarianza de  $X$  e  $Y$ .
3. ¿Son  $X$  e  $Y$  independientes? ¿Son  $X$  e  $Y$  incorreladas?

**Ejercicio 2.1.12.** Sea  $(X, Y)$  un vector aleatorio con distribución normal con parámetros  $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$  y  $\rho$ . Demostrar que  $X$  e  $Y$  son independientes  $\Leftrightarrow \rho = 0$ .