

# VECTORES ALEATORIOS

Julián de la Horra

Departamento de Matemáticas U.A.M.

## 1 Introducción

Desde un punto de vista formal, los vectores aleatorios son la herramienta matemática adecuada para transportar un modelo de probabilidad de un espacio muestral cualquiera a  $\mathbb{R}^n$ . En este sentido, los vectores aleatorios constituyen un caso particular de las funciones medibles que se estudian en Teoría de la Medida.

Por otro lado, desde un punto de vista más aplicado, los vectores aleatorios son la base matemática adecuada para:

¶ La modelización y el estudio de dos o más características numéricas en un experimento aleatorio (por ejemplo, para estudiar conjuntamente el peso y la estatura en una población, para estudiar la dosis de un principio activo para tratar la hipercolesterolemia junto con la reducción de colesterol que se ha conseguido, ...).

¶ La formalización del concepto de una muestra aleatoria que estará formada por observaciones independientes e idénticamente distribuidas de la característica que se desee estudiar.

## 2 Vectores aleatorios

Comenzamos con la definición formal de vector aleatorio:

**Definición.-** Consideramos un espacio de probabilidad  $(\Omega, S, P)$ . Un *vector aleatorio* de  $n$  dimensiones es una función  $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$  que verifica:

$$X^{-1}(A) \in S, \text{ para todo suceso } A \text{ de la } \sigma\text{-álgebra de Borel en } \mathbb{R}^n. \quad \bullet$$

A continuación, veremos que con esta definición se cumple el objetivo formal de trasladar perfectamente un modelo de probabilidad de un espacio muestral cualquiera,  $\Omega$ , a  $\mathbb{R}^n$ :

**Teorema.-** Consideramos un espacio de probabilidad  $(\Omega, S, P)$  y un vector aleatorio,  $X = (X_1, \dots, X_n)$ . Definimos:

$P_X(A) = P(X^{-1}(A))$ , para todo suceso  $A$  de la  $\sigma$ -álgebra de Borel en  $\mathfrak{R}^n$ .

Entonces, la función  $P_X$  es un modelo de probabilidad sobre  $\mathfrak{R}^n$ .

*Demostración.*- Similar a la demostración para variables aleatorias. •

Con la construcción anterior, hemos definido un modelo de probabilidad sobre  $\mathfrak{R}^n$ , a partir de un vector aleatorio. Además, recordemos que siempre hay una función de distribución asociada a un modelo de probabilidad sobre  $\mathfrak{R}^n$ . Todo esto, motiva las siguientes definiciones.

**Definiciones.-** Consideramos un espacio de probabilidad  $(\Omega, S, P)$  y un vector aleatorio,  $X = (X_1, \dots, X_n)$ .

El modelo de probabilidad  $P_X$ , definido sobre  $\mathfrak{R}^n$  por

$$P_X(A) = P(X^{-1}(A)),$$

recibe el nombre de *modelo de probabilidad generado o inducido por el vector aleatorio  $X$* .

La función de distribución  $F_X$  definida como

$$\begin{aligned} F_X(x_1, \dots, x_n) &= P_X\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\} \\ &= P(X^{-1}\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\}), \end{aligned}$$

recibe el nombre de *función de distribución generada o inducida por el vector aleatorio  $X$* . •

Lo único que realmente nos va a interesar de un vector aleatorio es el modelo de probabilidad que genera sobre  $\mathfrak{R}^n$ . Es decir, el modelo de probabilidad del que partimos,  $P$ , no nos va a interesar en absoluto, una vez que dispongamos del modelo de probabilidad,  $P_X$ , que ha generado el vector aleatorio sobre  $\mathfrak{R}^n$ .

Sólo un par de cosas más para finalizar la sección. En primer lugar, recordando que la  $\sigma$ -álgebra de Borel en  $\mathfrak{R}^n$  es la  $\sigma$ -álgebra generada por la clase  $C = \{(-\infty, x_1] \times \dots \times (-\infty, x_n]\}$ , podemos ver el siguiente resultado teórico:

**Teorema.-** Consideramos un espacio de probabilidad  $(\Omega, S, P)$  y una función  $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathfrak{R}^n$ . Entonces:

$X$  es vector aleatorio si y sólo si  $X^{-1}\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\} \in S$ , para todos los conjuntos de la forma  $(-\infty, x_1] \times \dots \times (-\infty, x_n]$ .

*Demostración.*- La demostración es similar a la dada para variables aleatorias. •

En segundo lugar, cuando estamos trabajando con un vector aleatorio,  $X = (X_1, \dots, X_n)$ , uno tiene inmediatamente la tentación de pensar que las componentes,  $X_i$ , de ese vector son variables aleatorias. A continuación, comprobamos que esa tentación está plenamente justificada:

**Teorema.-** Consideramos un espacio de probabilidad  $(\Omega, S, P)$  y una función  $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ . Entonces:

$X$  es un vector aleatorio si y sólo si todas sus componentes  $X_i$  son variables aleatorias.

*Demostración.*-

Implicación hacia la derecha.- Tenemos que probar que  $X_i^{-1}(-\infty, x_i] \in S$ . En efecto:

$$X_i^{-1}(-\infty, x_i] = X^{-1}\{\mathbb{R} \times \dots \times (-\infty, x_i] \times \dots \times \mathbb{R}\} \in S$$

Implicación hacia la izquierda.- Tenemos que probar que  $X^{-1}\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\} \in S$ . En efecto:

$$X^{-1}\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\} = \cap_{i=1}^n X_i^{-1}(-\infty, x_i] \in S \quad \bullet$$

A continuación, nos centramos en los dos tipos más habituales de vectores aleatorios: el tipo discreto y el tipo continuo. Para llevar a cabo su exposición, nos vamos a centrar en el plano,  $\mathbb{R}^2$ , por razones de sencillez de notación. Así que, en las siguientes secciones, vamos a considerar vectores aleatorios de dos dimensiones,  $(X, Y)$ . Por supuesto, las ideas se trasladan a  $\mathbb{R}^n$  sin mayores problemas.

### 3 Vectores aleatorios discretos

**Definición.-** Un vector aleatorio,  $(X, Y)$ , es de *tipo discreto*, cuando el modelo de probabilidad que genera sobre  $\mathbb{R}^2$  reparte la probabilidad sobre un conjunto finito o numerable de elementos.

Un vector aleatorio discreto queda caracterizado por su *función de masa conjunta*:

$$P_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j)$$

la cual, obviamente, cumple las dos características que tiene que verificar una función de masa:

- (a)  $P(X = x_i, Y = y_j) \in [0, 1]$ , para todo  $(x_i, y_j)$ .
- (b)  $\sum_i \sum_j P(X = x_i, Y = y_j) = 1$ . •

La función de masa conjunta define la distribución conjunta del vector aleatorio. A partir de la función de masa conjunta, podemos definir las distribuciones marginales y condicionadas:

**Definiciones.-** La *distribución marginal de  $X$*  es, simplemente, la distribución de la variable aleatoria  $X$ . La *función de masa marginal de  $X$*  se puede obtener a partir de la función de masa conjunta, mediante la relación:

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j)$$

Análogamente, la *distribución marginal de  $Y$*  es, simplemente, la distribución de la variable aleatoria  $Y$ . La *función de masa marginal de  $Y$*  se puede obtener a partir de la función de masa conjunta, mediante la relación:

$$P(Y = y_j) = \sum_i P(X = x_i, Y = y_j) \quad \bullet$$

Por lo tanto, a partir del vector aleatorio  $(X, Y)$  tenemos dos distribuciones marginales. En realidad, la única novedad es la relación de las funciones de masa de las variables aleatorias  $X$  e  $Y$  con la función de masa conjunta del vector aleatorio  $(X, Y)$ .

**Definiciones.-** La *distribución de  $X$  condicionada por un valor fijo de  $Y = y_j$*  viene definida por la siguiente función de masa:

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$$

Obsérvese que, aunque están involucradas la  $X$  y la  $Y$ , la única variable es la  $X$ , ya que la  $Y$  está fija.

La *distribución de  $Y$  condicionada por un valor fijo de  $X = x_i$*  viene definida por la siguiente función de masa:

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}$$

Obsérvese que, aunque están involucradas la  $X$  y la  $Y$ , la única variable es la  $Y$ , ya que la  $X$  está fija. •

### Observaciones

(a) Es muy interesante destacar que los tres tipos de funciones de masa (conjunta, marginales y condicionadas) están muy relacionadas entre ellas:

$$P(X = x_i, Y = y_j) = P(X = x_i | Y = y_j)P(Y = y_j) = P(Y = y_j | X = x_i)P(X = x_i)$$

(b) También es interesante resaltar que las distribuciones condicionadas utilizan la misma idea que las probabilidades condicionadas.

Los conceptos de esperanza y varianza eran importantes para las variables aleatorias. Para los vectores aleatorios tenemos el concepto adicional de covarianza y el concepto relacionado de coeficiente de correlación lineal, que en el caso discreto se definen de la siguiente forma:

**Definiciones.-** La *covarianza entre  $X$  e  $Y$*  se define como

$$\begin{aligned} \text{Cov}(X, Y) &= \sum_i \sum_j (x_i - E[X])(y_j - E[Y])P(X = x_i, Y = y_j) = \dots \\ &= \sum_i \sum_j x_i y_j P(X = x_i, Y = y_j) - E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

El *coeficiente de correlación lineal entre  $X$  e  $Y$*  se define como

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}$$

Cuando  $\text{Cov}(X, Y) = 0$ ,  $\rho(X, Y) = 0$ , y decimos que  $X$  e  $Y$  están *incorreladas*. •

## 4 Vectores aleatorios continuos

**Definición.-** Una variable aleatoria,  $X$ , es de *tipo continuo*, cuando el modelo de probabilidad que genera sobre  $\mathbb{R}^2$  reparte la probabilidad mediante una función de densidad, que es una función  $f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$  que verifica las dos siguientes propiedades:

- (a)  $f_{X,Y}(x, y) \geq 0$ , para todo  $(x, y) \in \mathbb{R}^2$ .
- (b)  $\int_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = 1$ .

La probabilidad de cualquier suceso  $A$ , se obtiene de la siguiente forma:

$$P_{X,Y}(A) = P\{(X, Y) \in A\} = \int_A f_{X,Y}(x, y) dx dy \quad \bullet$$

La función de densidad conjunta define la distribución conjunta del vector aleatorio. A partir de la función de densidad conjunta, podemos definir las distribuciones marginales y condicionadas:

**Definiciones.-** La *distribución marginal de  $X$*  es, simplemente, la distribución de la variable aleatoria  $X$ . La *función de densidad marginal de  $X$*  se puede obtener a partir de la función de densidad conjunta, mediante la relación:

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x, y) dy$$

Análogamente, la *distribución marginal de  $Y$*  es, simplemente, la distribución de la variable aleatoria  $Y$ . La *función de densidad marginal de  $Y$*  se puede obtener a partir de la función de densidad conjunta, mediante la relación:

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) dx \quad \bullet$$

Por lo tanto, a partir del vector aleatorio  $(X, Y)$  tenemos dos distribuciones marginales. En realidad, la única novedad es la relación de las funciones de densidad de las variables aleatorias  $X$  e  $Y$  con la función de densidad conjunta del vector aleatorio  $(X, Y)$ .

**Definiciones.-** La *distribución de  $X$  condicionada por un valor fijo de  $Y = y$*  viene definida por la siguiente función de densidad:

$$f(x|Y = y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Obsérvese que, aunque están involucradas la  $X$  y la  $Y$ , la única variable es la  $X$ , ya que la  $Y$  está fija.

La *distribución de  $Y$  condicionada por un valor fijo de  $X = x_i$*  viene definida por la siguiente función de densidad:

$$f(y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

Obsérvese que, aunque están involucradas la  $X$  y la  $Y$ , la única variable es la  $Y$ , ya que la  $X$  está fija. •

Sigue siendo interesante resaltar que los tres tipos de funciones de densidad (conjunta, marginales y condicionadas) están muy relacionadas entre ellas:

$$f_{X,Y}(x, y) = f(x|Y = y)f_Y(y) = f(y|X = x)f_X(x)$$

Los conceptos de esperanza y varianza eran importantes para las variables aleatorias. Para los vectores aleatorios tenemos el concepto adicional de covarianza y el concepto relacionado de coeficiente de correlación lineal, que en el caso continuo se definen de la siguiente forma:

**Definiciones.-** La *covarianza entre  $X$  e  $Y$*  se define como

$$\begin{aligned} \text{Cov}(X, Y) &= \int_{\mathbb{R}^2} (x - E[X])(y - E[Y])f_{X,Y}(x, y)dxdy = \dots \\ &= \int_{\mathbb{R}^2} xyf_{X,Y}(x, y)dxdy - E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

El *coeficiente de correlación lineal entre  $X$  e  $Y$*  se define como

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}$$

Cuando  $\text{Cov}(X, Y) = 0$ ,  $\rho(X, Y) = 0$ , y decimos que  $X$  e  $Y$  están *independientes*. •

## 5 Cambio de variables

Es frecuente e interesante, tanto en los desarrollos teóricos como en las aplicaciones, considerar transformaciones o cambio de variables, a partir de un vector aleatorio  $(X, Y)$ . La situación general es que estaremos interesados en conocer la distribución conjunta de un nuevo vector aleatorio  $(U, V)$ , donde  $U$  y  $V$  se definen a partir de las funciones:

$$\begin{aligned} u &= h_1(x, y) \\ v &= h_2(x, y) \end{aligned}$$

**(a) Caso discreto.-** Si el vector aleatorio  $(X, Y)$  es discreto, con una función de masa  $P_{X,Y}(x_i, y_j) = P(X = x_i, Y = y_j)$ , todo se reducirá a calcular la función de masa del nuevo vector aleatorio  $(U, V)$ , teniendo en cuenta

cómo se transforman los puntos  $(x, y)$  que tienen probabilidad positiva, es decir, teniendo en cuenta cómo se transforma el soporte de la distribución conjunta de  $(X, Y)$ :

$$P_{U,V}(u_i, v_j) = P(U = u_i, V = v_j) = P\{(x, y) : h_1(x, y) = u_i, h_2(x, y) = v_j\}$$

Este proceso no tiene ninguna complicación teórica, aunque puede ser tedioso en la práctica.

**(b) Caso continuo.-** Si el vector aleatorio  $(X, Y)$  es continuo, con una función de densidad  $f_{X,Y}(x, y)$ , tendremos que calcular la función de densidad  $f_{U,V}(u, v)$  del nuevo vector aleatorio  $(U, V)$ .

Para llevar a cabo este proceso, supongamos que la función bidimensional  $h = (h_1, h_2)$  es inyectiva en el soporte de la distribución conjunta de  $(X, Y)$  (es decir, en el conjunto donde se reparte la probabilidad del vector aleatorio  $(X, Y)$ ) y que, por tanto, en esa región (que es la única que nos interesa a efectos de la probabilidad) existe la función inversa:

$$\begin{aligned} x &= g_1(u, v) \\ y &= g_2(u, v) \end{aligned}$$

Si las funciones  $g_1$  y  $g_2$  tienen derivadas parciales continuas y el jacobiano

$$J = \begin{vmatrix} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{vmatrix}$$

es distinto de cero, entonces, para cualquier conjunto de Borel  $A$  contenido en el soporte de  $(X, Y)$ , tendremos:

$$\begin{aligned} P_{U,V}[h(A)] &= P_{X,Y}(A) = \int_A f_{X,Y}(x, y) dx dy \\ &= \int_{h(A)} f_{X,Y}(g_1(u, v), g_2(u, v)) |J| du dv \end{aligned}$$

Por lo tanto, la función de densidad del vector aleatorio  $(U, V)$  es:

$$f_{U,V}(u, v) = f_{X,Y}(g_1(u, v), g_2(u, v)) |J|$$

Si la función  $h = (h_1, h_2)$  no es inyectiva, se descompone en “componentes inyectivas”, y se suman las densidades que resultan al tratar cada una de las “componentes inyectivas” como se acaba de indicar.



## 6 Independencia, esperanzas, varianzas, ...

En esta sección, vamos a abordar diferentes cuestiones referentes a las variables aleatorias que forman parte de un vector aleatorio. Para hacer esto, volvemos a considerar un vector aleatorio con  $n$  dimensiones.

**Definiciones.-** Consideramos un vector aleatorio  $X = (X_1, \dots, X_n)$ .

Caso discreto: Decimos que las variables aleatorias  $X_1, \dots, X_n$  son *independientes* si:

$$P_X(x_1, \dots, x_n) = \prod_{i=1}^n P_{X_i}(x_i) \text{ para los posibles valores } x_1, \dots, x_n$$

Caso continuo: Decimos que las variables aleatorias  $X_1, \dots, X_n$  son *independientes* si:

$$f_X(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i) \text{ para los posibles valores } x_1, \dots, x_n \quad \bullet$$

### Observaciones

(a) En realidad, se puede dar una definición general de independencia de variables aleatorias, que vale tanto para el caso discreto como para el caso continuo, de la siguiente forma:

Decimos que las variables aleatorias  $X_1, \dots, X_n$  son *independientes* si para los conjuntos  $B_1, \dots, B_n$  de la  $\sigma$ -álgebra de Borel tenemos:

$$P \left[ \bigcap_{i=1}^n X_i^{-1}(B_i) \right] = \prod_{i=1}^n P \left[ X_i^{-1}(B_i) \right]$$

El interés de esta definición radica en que es del mismo estilo que la definición que se da para la independencia de sucesos.

(b) A continuación, hacemos notar que:

$$\begin{aligned} P \left[ \bigcap_{i=1}^n X_i^{-1}(B_i) \right] &= P \left[ X^{-1}(B_1 \times \dots \times B_n) \right] = P_X(B_1 \times \dots \times B_n) \\ \prod_{i=1}^n P \left[ X_i^{-1}(B_i) \right] &= \prod_{i=1}^n P_{X_i}(B_i) \end{aligned}$$

Por tanto, de forma alternativa, podemos decir que las variables aleatorias  $X_1, \dots, X_n$  son *independientes* si para los conjuntos  $B_1, \dots, B_n$  de la  $\sigma$ -álgebra de Borel:

$$P_X(B_1 \times \dots \times B_n) = \prod_{i=1}^n P_{X_i}(B_i)$$

El interés de esta definición radica en que nos relaciona el modelo de probabilidad,  $P_X$ , generado por el vector aleatorio con los modelos de probabilidad,  $P_{X_i}$ , generados por las variables aleatorias.

(c) Finalmente, señalemos que las definiciones de independencia que se han dado para el caso discreto y para el caso continuo surgen de la definición general, como era de suponer, pero, para poder verlo, es necesario disponer de resultados de Teoría de la Medida sobre el espacio producto. •

**Teorema.-** Consideramos un vector aleatorio  $X = (X_1, \dots, X_n)$  en el que las variables aleatorias  $X_1, \dots, X_n$  son independientes. Entonces:

$$E[X_1 \dots X_n] = \prod_{i=1}^n E[X_i] .$$

*Demostración.-* Lo vemos para el caso continuo (para el caso discreto es análogo):

$$\begin{aligned} E[X_1 \dots X_n] &= \int_{\mathbb{R}^n} x_1 \dots x_n f_X(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \int_{\mathbb{R}^n} x_1 \dots x_n f_{X_1}(x_1) \dots f_{X_n}(x_n) dx_1 \dots dx_n \\ &= \left[ \int_{\mathbb{R}} x_1 f_{X_1}(x_1) dx_1 \right] \dots \left[ \int_{\mathbb{R}} x_n f_{X_n}(x_n) dx_n \right] \\ &= \prod_{i=1}^n E[X_i] \quad \bullet \end{aligned}$$

**Teorema.-** Si  $X$  e  $Y$  son variables aleatorias independientes, entonces también son incorreladas.

*Demostración.-* Si  $X$  e  $Y$  son independientes, tenemos:

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0 \quad \bullet$$

Sin embargo, el que  $X$  e  $Y$  sean incorreladas no implica que sean independientes:

**Ejemplo.-** Consideremos el vector aleatorio de tipo discreto cuya función de masa es:

$$P(X = 0, Y = 1) = P(X = 0, Y = -1) = P(X = 1, Y = 0) = P(X = -1, Y = 0) = 1/4$$

En este caso, las funciones de masa de las variables aleatorias  $X$  e  $Y$  son:

$$\begin{aligned} P(X = -1) &= 1/4 & P(X = 0) &= 1/2 & P(X = 1) &= 1/4 \\ P(Y = -1) &= 1/4 & P(Y = 0) &= 1/2 & P(Y = 1) &= 1/4 \end{aligned}$$

Por un lado, es fácil calcular su covarianza:

$$Cov(X, Y) = E[XY] - E[X]E[Y] = 0 - (0)(0) = 0$$

con lo cual,  $X$  e  $Y$  son incorreladas.

Por otro lado, tenemos, por ejemplo:

$$\begin{aligned} P(X = 1, Y = 1) &= 0 \\ P(X = 1)P(Y = 1) &= (1/4)(1/4) = 1/16 \end{aligned}$$

con lo cual,  $X$  e  $Y$  no son independientes. •

**Teorema.-** Consideramos un vector aleatorio  $X = (X_1, \dots, X_n)$  cualquiera. Entonces:

$$E \left[ \sum_{i=1}^n a_i X_i + b \right] = \sum_{i=1}^n a_i E[X_i] + b$$

*Demostración.-* Lo vemos para el caso continuo (para el caso discreto es análogo):

$$\begin{aligned} E \left[ \sum_{i=1}^n a_i X_i + b \right] &= \int_{\mathbb{R}^n} \left[ \sum_{i=1}^n a_i x_i + b \right] f_X(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \sum_{i=1}^n a_i \int_{\mathbb{R}^n} x_i f_X(x_1, \dots, x_n) dx_1 \dots dx_n + b \\ &= \sum_{i=1}^n a_i \int_{\mathbb{R}} x_i f_{X_i}(x_i) dx_i + b \\ &= \sum_{i=1}^n a_i E[X_i] + b \quad \bullet \end{aligned}$$

**Teorema.-**

(a) Consideramos un vector aleatorio  $X = (X_1, \dots, X_n)$  cualquiera. Entonces:

$$V \left( \sum_{i=1}^n a_i X_i + b \right) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{i < j} a_i a_j Cov(X_i, X_j)$$

(b) Consideramos un vector aleatorio  $X = (X_1, \dots, X_n)$  en el que las variables aleatorias  $X_1, \dots, X_n$  son independientes. Entonces:

$$V \left( \sum_{i=1}^n a_i X_i + b \right) = \sum_{i=1}^n a_i^2 V(X_i)$$

*Demostración.-* (a) Tenemos:

$$\begin{aligned}
 V\left(\sum_{i=1}^n a_i X_i + b\right) &= E\left[\left(\sum_{i=1}^n a_i X_i + b - \sum_{i=1}^n a_i E[X_i] - b\right)^2\right] \\
 &= E\left[\left(\sum_{i=1}^n a_i (X_i - E[X_i])\right)^2\right] \\
 &= E\left[\sum_{i=1}^n a_i^2 (X_i - E[X_i])^2 + 2 \sum_{i < j} a_i a_j (X_i - E[X_i]) (X_j - E[X_j])\right] \\
 &= \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{i < j} a_i a_j Cov(X_i, X_j)
 \end{aligned}$$

(b) Es inmediato porque, en este caso, las covarianzas son cero. •