Tema 3 - Variables Aleatorias discretas multidimensionales

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Variables aleatorias bidimensionales discretas

Variables aleatorias bidimensionales discretas. Introducción

Definición de variable aleatoria bidimensional.

Sea Ω es espacio muestral de un experimento. Diremos que (X, Y) es una **variable aleatoria bidimensional** cuando tanto X como Y toman valores reales para cada elemento del espacio Ω.

Diremos que es **discreta** cuando su conjunto de valores en $\mathbb{R}^2$, $(X, Y)(\Omega)$ es un conjunto finito o numerable.

Diremos que es **continua** cuando su conjunto de valores en $\mathbb{R}^2$, $(X, Y)(\Omega)$ es un producto de intervalos.

Diremos que es **heterogénea** cuando X e Y no compartan ser continuas o discretas.

Función de probabilidad conjunta

Definición de función de probabilidad conjunta: Dada una **variable aleatoria bidimensional discreta** (X, Y)

definimos la función de **probabilidad discreta bidimensional** como

$$\begin{aligned} P_{XY} : \mathbb{R}^2 &\longrightarrow [0,1] \\ (x,y) &\longrightarrow P_{XY}(x,y) = P(X = x,\ Y = y). \end{aligned}$$

Llamaremos dominio de la variable conjunta a

$$D_{XY} = \{(x,y) \in \mathbb{R}^2 | P_{XY}(x,y) = P(X = x,\ Y = y) > 0\}.$$

Es decir es el conjunto de valores posibles que toma la v.a. (X, Y).

Función de probabilidad conjunta

Por tanto, de cara a calcular P_{XY} basta calcular $P_{XY}(x_i, y_j)$ para $(x_i, y_j) \in D_{XY}$:

X/Y	y_1	y_2	$\ldots$	y_N
x_1	$P_{XY}(x_1, y_1)$	$P_{XY}(x_1, y_2)$	$\ldots$	$P_{XY}(x_1, y_N)$
x_2	$P_{XY}(x_2, y_1)$	$P_{XY}(x_2, y_2)$	$\ldots$	$P_{XY}(x_2, y_N)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_M	$P_{XY}(x_M, y_1)$	$P_{XY}(x_M, y_2)$	$\ldots$	$P_{XY}(x_M, y_N)$

Propiedades de la función de probabilidad conjunta

Sea (X, Y) una **variable aleatoria bidimensional discreta** con dominio $D_{XY} = \{(x_i, y_j)\, i = 1, 2, \ldots,\ j = 1, 2, \ldots\}$.

Su **función de probabilidad conjunta** verifica las siguientes propiedades:

La suma de todos los valores de la **función de probabilidad conjunta** sobre el conjunto de valores siempre vale 1:

$$\sum_i \sum_j P_{XY}(x_i, y_j) = 1.$$

Propiedades de la función de probabilidad conjunta

Sea B un subconjunto cualquiera del dominio D_{XY}. El valor de la probabilidad $P((X, Y) \in B)$ se puede calcular de la forma siguiente:

$$P((X, Y) \in B) = \sum_{(x_i, y_j) \in B} P_{XY}(x_i, y_j).$$

Es decir, la probabilidad de que la variable bidimensional tome valores en B es igual a la suma de todos aquellos valores de la función de probabilidad conjunta que están en B.

Función de distribución acumulada

Definición función de distribución conjunta

La función de distribución acumulada conjunto o simplemente distribución conjunta se define como

$$F_{XY}(x, y) = P(X \leq x, Y \leq y).$$

Propiedad

La **función de distribución conjunta** se puede obtener conociendo la **función de probabilidad conjunta**

$$F_{XY}(x, y) = \sum_{x_i \leq x, y_j \leq y} P_{XY}(x_i, y_j).$$

Distribuciones marginales

Variables aleatorias marginales y su distribución

Consideremos una variable aleatoria **bidimensional discreta** (X, Y) con **función de probabilidad conjunta** $P_{XY}(x_i, y_j)$, para cada $(x_i, y_j) \in D_{XY}$.

La tabla de la **función de probabilidad conjunta** contiene suficiente información para obtener las **funciones de probabilidad** de las variables X e Y.

Dichas variables X e Y se denominan **variables marginales** y sus correspondientes **funciones de probabilidad**, **funciones de probabilidad marginales** P_X de la variable X y P_Y de la variable Y.

Veamos cómo obtener P_X y P_Y a partir de la tabla P_{XY}.

Funciones de probabilidad marginales

Proposición. Cálculo de las funciones de probabilidad marginales.

Sea (X, Y) una variable aleatoria **bidimensional discreta** con **función de probabilidad conjunta** $P_{XY}(x_i, y_j)$, con $(x_i, y_j) \in D_{XY}$.

Las **funciones de probabilidad marginales** $P_X(x_i)$ y $P_Y(y_j)$ se calculan usando las expresiones siguientes:

$$\begin{aligned} P_X(x_i) &= \sum_j P_{XY}(x_i, y_j),\ i = 1, 2, \ldots, \\ P_Y(y_j) &= \sum_i P_{XY}(x_i, y_j),\ \ j = 1, 2, \ldots \end{aligned}$$

Variables aleatorias marginales

- Podemos representar P_{XY} como una tabla bidimensional en la primera fila están los valores de la variable Y $(y_1, y_2, \ldots)$ y en la primera columna están los valores de la variable X $(x_1, x_2, \ldots)$
- Para obtener la **función de probabilidad marginal** de la variable X en el valor x_i, $P_X(x_i)$, hay que sumar todos los valores de $P_{XY}(x_i, y_j)$ correspondientes a la fila i-ésima
- De forma análoga para obtener la **función de probabilidad marginal** de la variable Y en el valor y_j, $P_Y(y_j)$, hay que sumar todos los valores de $P_{XY}(x_i, y_j)$ correspondientes a la columna j-ésima.

Variables aleatorias marginales

$X\backslash Y$	y_1	y_2	$\ldots$	y_N	$P_X(x_i) = \sum_j P_{XY}(x_i, y_j)$
x_1	$P_{XY}(x_1, y_1)$	$P_{XY}(x_1, y_2)$	$\ldots$	$P_{XY}(x_1, y_N)$	$P_X(x_1)$
x_2	$P_{XY}(x_2, y_1)$	$P_{XY}(x_2, y_2)$	$\ldots$	$P_{XY}(x_2, y_N)$	$P_X(x_2)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
x_M	$P_{XY}(x_M, y_1)$	$P_{XY}(x_M, y_2)$	$\ldots$	$P_{XY}(x_M, y_N)$	$P_X(x_M)$
$P_Y(y_j) = \sum_i P_{XY}(x_i, y_j)$	$P_Y(y_1)$	$P_Y(y_2)$	$\ldots\ldots$	$P_Y(y_N)$	1

Independencia de variables aleatorias discretas

Recordemos que dos sucesos A y B son independientes si

$$P(A \cap B) = P(A) \cdot P(B).$$

¿Cómo trasladar dicho concepto al caso de variables aleatorias?

Dada una variable aleatoria bidimensional discreta (X, Y) con $D_{XY} = \{(x_i, y_j),\ i = 1, 2, \ldots, j = 1, 2, \ldots\}$

Así que al menos todos los sucesos de la forma $\{X = x_i,\ Y = y_j\}$ deberán ser independientes.

Independencia de variables aleatorias discretas

Definición de independencia para variables aleatorias bidimensionales discretas.

Dada (X, Y) una **variable aleatoria bidimensional discreta** con **función de probabilidad** P_{XY} y **funciones de probabilidad marginales** P_X y P_Y.

Diremos que X e Y son independientes si:

$$P_{XY}(x_i, y_j) = P_X(x_i) \cdot P_Y(y_j),\ i = 1, 2, \ldots, j = 1, 2, \ldots$$

o dicho de otra forma:

$$P(X = x_i,\ Y = y_j) = P(X = x_i) \cdot P(Y = y_j),\ i = 1, 2, \ldots, j = 1, 2, \ldots$$

Propiedad

Las v.a. X e Y son independientes si y solo si $F_{XY}(x, y) = F_X(x) \cdot F_Y(y)$.

Esperanza y varianza de las distribuciones marginales

- $E(X) = \sum_{x \in D_X} x \cdot P_X(x) = \sum_{x \in D_X} x \cdot P(X = x).$
- $E(Y) = \sum_{y \in D_Y} y \cdot P_Y(y) = \sum_{y \in D_Y} y \cdot P(Y = y).$
- $\sigma_X^2 = Var(X) = E(X - E(X)) = E(X) - E(X)^2.$
- $\sigma_Y^2 = Var(Y) = E(Y - E(Y)) = E(Y) - E(Y)^2.$

Distibuciones condicionales

- Dado un valor fijo $y \in D_Y$ definimos la distribución condicional de la v.a. X condicionada a que $Y = y$ como

$$P(X = x | Y = y) = \frac{P_{XY}(x, y)}{P_Y(y)} = \frac{P(X = x, Y = y)}{P(Y = y)},\ \text{para todo } x \in D_X.$$

- Dado un valor fijo $y \in D_Y$ definimos la distribución condicional de la v.a. Y condicionada a que $X = x$ como

$$P(Y = y | X = x) = \frac{P_{XY}(x, y)}{P_X(x)} = \frac{P(X = x, Y = y)}{P(X = x)},\ \text{para todo } y \in D_Y.$$

Distibuciones condicionales e independencia

Propiedad

Si las variables X e Y son independientes se cumple que

1. $P(X = x | Y = y) = P(X = x)$
2. $P(Y = y | X = x) = P(Y = y).$

Esperanzas condicionales

$$E(X|Y=y)=\sum_{x\in D_X} x\cdot P(X=x|Y=y)$$

$$E(Y|X=x)=\sum_{y\in D_Y} y\cdot P(Y=y|X=x)$$

Propiedad

Si las variables X e Y son independientes se cumple que

1. $E(X|Y=y)=E(X)$
2. $E(Y|X=x)=E(Y)$

Esperanzas de funciones de v.a. discretas bidimensionales. Covarianza y correlación

Esperanzas de funciones de v.a. discretas bidimensionales

Definición:

Sea (X,Y) una variable aleatoria bidimensional discreta y $g(X,Y)$ una función de esa variable bidimensional entonces $E(g(X,Y))=\sum_i\sum_j g(x_i,y_j)\cdot P(X=x_i,Y=y_j)$.

En particular:

- $E(X+Y)=\sum_i\sum_j (x_i+y_j)\cdot P(X=x_i,Y=y_j)=\mu_X+\mu_Y$.
- $Var(X+Y)=E\left((X+Y-E(X+Y))^2\right)=\sum_i\sum_j (x_i+y_j-(\mu_X+\mu_Y))^2\cdot P(X=x_i,Y=y_j)$.

Esperanzas de funciones de v.a. discretas bidimensionales

Propiedad: Sea (X,Y) una variable aleatoria bidimensional entonces se cumple que:

- $E(X+Y)=E(X)+E(Y)=\mu_X+\mu_y$
- Si X e Y son independientes entonces $E(X\cdot Y)=E(X)\cdot E(Y)=\mu_X\cdot\mu_y$
- Si X e Y son independientes entonces $Var(X+Y)=Var(X)+Var(Y)=\sigma_X^2+\sigma_y^2$

Covarianza y correlación

Medida de la variación conjunta: covarianza

El **momento conjunto centrado en las medias para** $k=1$ **y** $l=1$ se denomina **covarianza** entre las variables X e Y:

$$\sigma_{XY}=Cov(X,Y)=E((X-\mu_X)(Y-\mu_Y)).$$

La covarianza puede calcularse también con:

$$Cov(X,Y)=E(X\cdot Y)-E(X)\cdot E(Y)=E(X\cdot Y)-\mu_X\cdot\mu_Y,$$

Propiedad. Si las variables X e Y son **independientes**, entonces $Cov(X,Y)=0$.

Es una consecuencia de que si X e Y son independientes entonces que vimos que $E(X\cdot Y)=E(X)\cdot E(Y)=\mu_X\cdot\mu_y$.

Covarianza entre las variables

La **covarianza** es una medida de lo relacionadas están las variables X e Y:

- Si cuando $X\geq\mu_X$, también ocurre que $Y\geq\mu_Y$ o viceversa, cuando $X\leq\mu_X$, también ocurre que $Y\leq\mu_Y$, el valor $(X-\mu_X)(Y-\mu_Y)$ será positivo y la **covarianza** será positiva.
- Si por el contrario, cuando $X\geq\mu_X$, también ocurre que $Y\leq\mu_Y$ o viceversa, cuando $X\leq\mu_X$, también ocurre que $Y\geq\mu_Y$, el valor $(X-\mu_X)(Y-\mu_Y)$ será negativo y la **covarianza** será negativa.
- En cambio, si a veces ocurre una cosa y a veces ocurre otra, la **covarianza** va cambiando de signo y puede tener un valor cercano a 0.

Propiedades de la covarianza

- Sea (X,Y) una variable aleatoria bidimensional. Entonces la **varianza de la suma/resta** se calcula usando la expresión siguiente:

$$Var(X\pm Y)=Var(X)+Var(Y)\pm 2\cdot Cov(X,Y).$$

- Sea (X,Y) una variable aleatoria bidimensional donde las variables X e Y son **independientes**. Entonces:

$$Var(X+Y)=Var(X)+Var(Y).$$

Coeficiente de correlación

La **covarianza** depende de las unidades en las que se midan las variables X e Y ya que si $a>0$ y $b>0$, entonces:

$$Cov(a\cdot X,b\cdot Y)=a\cdot b\cdot Cov(X,Y).$$

Por tanto, si queremos "medir" la relación que existe entre las variables X e Y tendremos que "normalizar" la **covarianza** definiendo el **coeficiente de correlación** entre las variables X e Y:

Coeficiente de correlación entre las variables

Definición del coeficiente de correlación. Sea (X,Y) una variable aleatoria bidimensional. Se define el **coeficiente de correlación** entre las variables X e Y como:

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\cdot\sqrt{Var(Y)}} = \frac{E(X\cdot Y) - \mu_X\cdot\mu_Y}{\sqrt{E(X^2)-\mu_X^2}\cdot\sqrt{E(Y^2)-\mu_Y^2}}.$$

Coeficiente de correlación entre las variables

Observación. Si las variables X e Y son **independientes**, su **coeficiente de correlación** $\rho_{XY}=0$ es nulo ya que su **covarianza** lo es.

Notemos también que la **correlación** no tiene unidades y es invariante a cambios de escala.

Además, la **covarianza** de las **variables tipificadas** $\frac{X-\mu_X}{\sigma_X}$ y $\frac{Y-\mu_Y}{\sigma_Y}$ coincide con la **correlación** de X e Y.

El **coeficiente de correlación** es un valor normalizado ya que siempre está entre -1 y 1: $-1\leq\rho_{XY}\leq 1$.

Coeficiente de correlación entre las variables

Observación. Si las variables X e Y tiene dependencia lineal, por ejemplo si $Y=a\cdot X+b$ para algunas constantes $a,b\in\mathbb{R}$, entonces su **coeficiente de correlación** $\rho_{XY}=\pm 1$, es decir toma el valor 1 si la pendiente $a>0$ y -1 si $a<0$.

De forma similar:

- si $Cor(X,Y)=+1$ X e Y tienen relación lineal con pendiente positiva.
- si $Cor(X,Y)=-1$ X e Y tienen relación lineal con pendiente negativa.

Matriz de varianzas-covarianzas y matriz de correlaciones

Sea (X,Y) una variable bidimensional Notemos que

- $Cov(X,X)=\sigma_{XX}=\sigma_X^2$.
- $Cov(Y,Y)=\sigma_{YY}=\sigma_Y^2$.
- $\sigma_{XY}=Cov(X,Y)=Cov(Y,X)=\sigma_{YX}$.

Se denomina matriz de varianzas-covarianzas y se suele denotar como Σ a

$$\Sigma=\begin{pmatrix} Cov(X,X) & Cov(X,Y)\\ Cov(Y,X) & Cov(Y,Y)\end{pmatrix}=\begin{pmatrix}\sigma_{XX} & \sigma_{XY}\\ \sigma_{YX} & \sigma_{YY}\end{pmatrix}=\begin{pmatrix}\sigma_X^2 & \sigma_{XY}\\ \sigma_{YX} & \sigma_Y^2\end{pmatrix}$$

Matriz de varianzas-covarianzas y matriz de correlaciones

Sea (X,Y) una variable bidimensional Notemos que

- $Cor(X,X)=\rho_{XX}=1$.
- $Cor(Y,Y)=\rho_{YY}=1$.
- $\rho_{XY}=Cor(X,Y)=Cor(Y,X)=\rho_{YX}$.

Se denomina matriz de correlaciones a

$$R=\begin{pmatrix} Cor(X,X) & Cor(X,Y)\\ Cor(Y,X) & Cor(Y,Y)\end{pmatrix}=\begin{pmatrix}1 & \rho_{XY}\\ \rho_{YX} & 1\end{pmatrix}=\begin{pmatrix}1 & \rho_{XY}\\ \rho_{XY} & 1\end{pmatrix}.$$

Distribuciones multidimensionales

Conceptos básicos. Función de probabilidad y de distribución.

Consideremos un vector compuesto de n variables aleatorias discretas $(X_1,X_2,\ldots,X_n)$

Su **función de probabilidad** es

$$\begin{aligned}P_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) &= P\Big((X_1,X_2,\ldots,X_n)=(x_1,x_2,\ldots,x_n)\Big)\\ &= P(X_1=x_1,X_2=x_2,\ldots,X_n=x_n).\end{aligned}$$

Su función de **distribución de probabilidad** es

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=P(X_1\leq x_1,X_2\leq x_2,\ldots,X_n\leq x_n).$$

Independencia

Definición independencia

Diremos que la variables $X_1,X_2,\ldots,X_n$ son **INDEPENDIENTES** cuando

$$P_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=P_{X_1}(x_1)\cdot P_{X_2}(x_2)\cdot\ldots\cdot P_{X_n}(x_n).$$

Propiedad

Las variables $X_1,X_2,\ldots,X_n$ son **INDEPENDIENTES** si y solo si

$$F_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n)=F_{X_1}(x_1)\cdot F_{X_2}(x_2)\cdot\ldots\cdot F_{X_n}(x_n).$$

Conceptos básicos

Vector de medias

Si denotamos $E(X_i)=\mu_i$ para $i=1,2,\ldots,n$ el **vector de medias** es

$$E(X_1,X_2,\ldots,X_n)=(E(X_1),E(X_2),\ldots,E(X_n))=(\mu_1,\mu_2,\ldots,\mu_n).$$

Covarianza y varianzas

Si denotamos $\sigma_{ij}=Cov(X_i,X_j)$ para todo i,j en $1,2,\ldots n$ entonces tenemos que

- $\sigma_{ii}=Cov(X_i,X_i)=\sigma_{ii}=\sigma_i^2$.
- $\sigma_{ij}=Cov(X_i,X_j)=Cov(X_j,X_i)=\sigma_{ji}$.

Conceptos básicos

Si denotamos $\rho_{ij} = Cor(X_i, X_j)$ para todo i, j en $1, 2, \ldots n$ entonces tenemos que

- $\rho_{ii} = Cor(X_i, X_i) = 1.$
- $\rho_{ij} = Cor(X_i, X_j) = Cor(X_j, X_i) = \rho_{ji}.$

Matrices de varianzas-covarianzas y de correlaciones

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \ldots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_n^2 \end{pmatrix}, \qquad R = \begin{pmatrix} 1 & \rho_{12} & \ldots & \rho_{1n} \\ \rho_{21} & 1 & \ldots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \ldots & 1 \end{pmatrix}.$$