Parte 1. Tema 4: Breve Introducción a las Variables Aleatorias multidimensionales

Estadística

2022-09-13

1. Variables aleatorias bidimensionales
2. Distribuciones marginales
3. Esperanzas de funciones de v.a. discretas bidimensionales. Covarianza y correlación
4. Covarianza y correlación
5. Distribuciones multidimensionales
6. Ejemplo caso discreto
7. Ejemplo bivariante continua normal bivariante

Lección 1

Variables aleatorias bidimensionales

Variables aleatorias bidimensionales. Introducción

Definición de variable aleatoria bidimensional.

En este caso tendremos un experimento con dos resultados.

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

Variables aleatorias bidimensionales. Introducción

Por ejemplo

- Lanzamos un dado rojo y una azul veces (X, Y) =("resultado dado rojo", "resultado dado azul"). Dominio $D_{X,Y} = \{(i,j) | \quad i,j = 1,2,3,4,5,4\}$.
- (X, Y) =("tamaño en memoria del proceso", "tiempo de CPU usado") de un proceso de un servidor escogido al azar. Dominio $D_{X,Y} = \{(x,y) \in \mathbb{R}^2 | \quad x \geq 0, \quad y \geq 0\}$.

Variables aleatorias bidimensionales. Introducción

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 distribución acumulada

Definición función de distribución conjunta

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

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

Esta función existe para variables aleatorias discretas y continuas.

Función de probabilidad conjunta para variables aleatorias discretas.

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

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

$$P_{XY}(x, y) = P(X = x,\ Y = y),\ \text{para cada } (x, y) \in D_{XY}.$$

Así el dominio de la variable conjunta es

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

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.

Propiedades de la función de probabilidad conjunta

Propiedad

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

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

Función de distribución acumulada, función de densidad

Definición función de densidad conjunta

Sea $f_{XY}:\mathbb{R}\times\mathbb{R}\mapsto[0,+\infty)$ que cumple que:

- $f_{XY}(x,y)>0$ para todo $(x,y)\in D_{XY}$.
- $\displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{XY}(t_x,t_y)dt_x dt_y=1.$
- $\displaystyle F_{XY}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(t_x,t_y)dt_x dt_y.$

El dominio (valores posibles) de la variable conjunta es

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

Lección 2

Distribuciones marginales

Variables aleatorias marginales y su distribución

Consideremos una variable aleatoria **bidimensional** (X, Y) llamaremos distribuciones marginales de las variables X e Y a las distribuciones individuales de cada variable obtenidas desde la distribución conjunta.

Dichas variables X e Y se denominan **variables marginales** y sus correspondientes **funciones de probabilidad o de densidad**, y se denominan **funciones de probabilidad o densidad marginales**

Funciones de probabilidad marginales caso discreto

Cálculo de las funciones de probabilidad marginales caso discreto.

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:

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

Funciones de probabilidad marginales caso discreto

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

Funciones de probabilidad marginales continuas

Proposición. Cálculo de las funciones de densidad marginales.

Sea (X, Y) una variable aleatoria **bidimensional continua** con **función de densidad conjunta** $f_{XY}(x, y)$, con $(x, y) \in D_{XY}$.

Las **funciones de densidad marginales** $f_X(x)$ y $f_Y(y)$ se calculan usando las expresiones siguientes:

- $f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y) \quad dy.$
- $f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x, y) \quad dx.$

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 se cumple alguna de estas condiciones:

- $P_{XY}(x_i, y_j) = P_X(x_i) \cdot P_Y(y_j),\ i = 1, 2, \ldots, j = 1, 2, \ldots$
- $P(X = x_i,\ Y = y_j) = P(X = x_i) \cdot P(Y = y_j),\ i = 1, 2, \ldots, j = 1, 2, \ldots$
- $F_{XY}(x, y) = F_X(x) \cdot F_Y(y)$.

Independencia de variables aleatorias continuas

Condiciones para independencia de variables aleatorias bidimensionales continuas

Dada (X, Y) una **variable aleatoria bidimensional continua** con **función de densidad** f_{XY} y **funciones de probabilidad marginales** f_X y f_Y.

Diremos que X e Y son independientes si se cumple al menos una de las siguientes condiciones:

- $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$ para todo $(x, y) \in D_{XY}$
- $F_{XY}(x, y) = F_X(x) \cdot F_Y(y)$ para todo $(x, y) \in D_{XY}$

Esperanza y varianza de las distribuciones marginales dicretas.

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

Esperanza y varianza de las distribuciones marginales continuas.

- $E(X) = \int_{-\infty}^{+\infty} x \cdot f_X(x) \quad dx.$
- $E(Y) = \int_{-\infty}^{+\infty} y \cdot f_Y(y) \quad dy.$
- $\sigma_X^2 = Var(X) = E\left((X - E(X))^2\right) = E(X^2) - E(X)^2.$
- $\sigma_Y^2 = Var(Y) = E\left((Y - E(Y))^2\right) = E(Y^2) - E(Y)^2.$

Distibuciones condicionales dicretas

- 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 continuas

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

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

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

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

Distibuciones condicionales e independencia

Propiedad

Las variables discretas X e Y son independientes si y solo sí se cumple que

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

Si las variables X e Y son independientes si y solo sí se cumple que

1. $f_{X|Y=y}(x) = f_X(x)$
2. $f_{Y|X=x}(y) = f_Y(y)$

Esperanzas condicionales

- **Caso discreto** $E(X|Y=y) = \sum_{x\in D_X} x \cdot P(X=x|Y=y)$
- **Caso continuo** $E(X|Y=y) = \int_{-\infty}^{+\infty} x \cdot f_{X|Y=y}(x) \quad dx.$

Las definiciones para $E(Y|X=x)$ son similares.

Propiedad

Si las variables X e Y son independientes se cumple que

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

Lección 3

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

$$E(g(X,Y)) = \sum_i \sum_j g(x_i, y_j) \cdot P(X = x_i, Y = y_j).$$

- Sea (X,Y) una variable aleatoria bidimensional continua y $g(X,Y): \mathbb{R}^2 \mapsto \mathbb{R}$ una función de esa variable bidimensional entonces

$$E(g(X,Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x,y) \cdot f_{XY}(x,y) \quad dxdy.$$

Esperanzas de funciones de v.a. bidimensionales

Caso discreto:

$$E(X+Y)=\sum_i\sum_j (x_i+y_j)\cdot P(X=x_i,Y=y_j)=\mu_X+\mu_Y.$$

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

Caso continuo:

$$E(X+Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)\cdot f_{XY}(x,y)\quad dxdy=\mu_X+\mu_Y.$$

$$Var(X+Y)=E\left((X+Y-E(X+Y))^2\right)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y-(\mu_X+\mu_Y))^2\cdot f_{XY}(x,y)\quad dxdy.$$

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 $Var(X+Y) = Var(X) + Var(Y) = \sigma_X^2 + \sigma_y^2$

Lección 4

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:

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

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

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

Lección 5

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

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

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

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

Lección 6

Ejemplo caso discreto

Lanzamiento de dos dados

Lanzamos dos dados numerados de 1 a 4 caras dos veces de forma independiente (X, Y) resultado del lanzamiento de la primera y la segunda vez respectivamente

$$P_{XY}(x,y) = \begin{cases} \frac{1}{16} & \text{si } x, y = 1, 2, 3, 4 \\ 0 & \text{en otro caso.} \end{cases}$$

Dados Marginales

$P_X(x) = \sum_{y=1}^{4} P_{XY}(x,y) = P(x,1) + \cdots + P(x,4) = 4 \cdot \frac{1}{6} = \frac{1}{4}$ si $x = 1, 2, \ldots, 4$

$P_Y(y) = \sum_{x=1}^{4} P_{XY}(x,y) = P(1,y) + \cdots + P(4,y) = 4 \cdot \frac{1}{6} = \frac{1}{4}$ si $y = 1, 2, \ldots, 4$

$E(X) = \sum_{x=1}^{4} x \cdot P_X(x) = \sum_{x=1}^{4} x \cdot \frac{1}{4} = \frac{10}{4} = 2.5.$

$E(X^2) = \sum_{x=1}^{4} x^2 \cdot P_X(x) = \sum_{x=1}^{4} x^2 \cdot \frac{1}{4} = 7.5.$

$Var(X) = E(X^2) - E(X)^2 = 1.25$

Los mismo resultados se obtienen para $E(Y)$ y $Var(Y)$.

Dados Marginales

```r
media_X=sum(1:4)/4
media_X
```

```
## [1] 2.5
```

```r
media_X_cuadrados=sum((1:4)^2)/4
media_X_cuadrados
```

```
## [1] 7.5
```

Dados Marginales

```
var_X=media_X_cuadrados-media_X^2
var_X
```

```
## [1] 1.25
```

```
sd_X=sqrt(var_X)
sd_X
```

```
## [1] 1.118034
```

Dados

$$
\begin{aligned}
E(X\cdot Y) &= \sum_{x=1}^{4}\sum_{y=1}^{4} x\cdot y\frac{1}{16} = \frac{1}{16}\sum_{x=1}^{4} x\cdot\left(\sum_{y=1}^{4} y\right)\\
&= \frac{1}{16}\sum_{x=1}^{4} x\cdot\left(\frac{4\cdot 5}{2}\right)\frac{1}{16}10\sum_{x=1}^{4} x\\
&= \frac{1}{16}10\cdot 10 = \frac{1}{16}(10)^2 = \frac{10^2}{4^2}.
\end{aligned}
$$

$$
\begin{aligned}
Cov(X,Y) &= E(X,Y) - E(X)\cdot E(Y)\\
&= \frac{10^2}{4^2} - \left(\frac{10}{4}\right)^2 = 0.
\end{aligned}
$$

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\cdot Var(Y)}} = 0.$$

Dados

$$
\begin{aligned}
\Sigma &= \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(Y,X) & Cov(Y,Y) \end{pmatrix} = \begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{pmatrix} \\
&= \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} 1.25 & 0 \\ 0 & 1.25 \end{pmatrix}
\end{aligned}
$$

$$
R = \begin{pmatrix} 1 & \rho_{XY} \\ \rho_{XY} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
$$

Dados

Marginales Si $x, y = 1, 2, 3, 4$

$P_{X|Y=y}(X = x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{\frac{1}{16}}{\frac{1}{4}} = \frac{1}{4}$

Notemos que entonces $P_{x|Y=y}(X = x) = P_X(x)$ y por lo tanto son independientes

Efectivamente, son independientes

$$P_{XY}(x, y) = \frac{1}{16} = P_X(x) \cdot P_Y(y) = \frac{1}{4} \cdot \frac{1}{4}.$$

Ejemplo dados máximo y suma

Ahora tiramos dos veces un dado con valores de 1 a 4 de forma que (X, Y) son las variables X máximo de las dos tiradas e Y suma de las dos tiradas

```
dados=data.frame(d1=rep(1:4,times=4),d2=rep(1:4,each=4))
dados$X=pmax(dados$d1,dados$d2)
dados$Y=dados$d1+dados$d2
dados[1:6,]
```

```
##   d1 d2 X Y
## 1  1  1 1 2
## 2  2  1 2 3
## 3  3  1 3 4
## 4  4  1 4 5
## 5  1  2 2 3
## 6  2  2 2 4
```

Ejemplo dados máximo y suma

```
dados[1:16,]
```

```
##    d1 d2 X Y
## 1   1  1 1 2
## 2   2  1 2 3
## 3   3  1 3 4
## 4   4  1 4 5
## 5   1  2 2 3
## 6   2  2 2 4
## 7   3  2 3 5
## 8   4  2 4 6
## 9   1  3 3 4
## 10  2  3 3 5
## 11  3  3 3 6
## 12  4  3 4 7
```

Ejemplo dados máximo y suma

```
P_XY=prop.table(table(dados$X,dados$Y))
P_XY
```

```
##
##          2      3      4      5      6      7      8
##   1 0.0625 0.0000 0.0000 0.0000 0.0000 0.0000 0.0000
##   2 0.0000 0.1250 0.0625 0.0000 0.0000 0.0000 0.0000
##   3 0.0000 0.0000 0.1250 0.1250 0.0625 0.0000 0.0000
##   4 0.0000 0.0000 0.0000 0.1250 0.1250 0.1250 0.0625
```

```
str(P_XY)
```

```
##  'table' num [1:4, 1:7] 0.0625 0 0 0 0 0.125 0 0 0 0.0625 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : chr [1:4] "1" "2" "3" "4"
##   ..$ : chr [1:7] "2" "3" "4" "5"
```

Ejemplo dados máximo y suma

```r
P_X=margin.table(P_XY,1)
P_X
```

```
##
##      1      2      3      4
## 0.0625 0.1875 0.3125 0.4375
```

```r
P_Y=margin.table(P_XY,2)
P_Y
```

```
##
##      2      3      4      5      6      7      8
## 0.0625 0.1250 0.1875 0.2500 0.1875 0.1250 0.0625
```

Ejemplo dados máximo y suma

```r
df=as.data.frame(P_XY)
names(df)=c("X","Y","P_XY")
df$X=as.integer(df$X)
df$Y=as.integer(df$Y)
df[1:11,]
```

```
##     X Y   P_XY
## 1   1 1 0.0625
## 2   2 1 0.0000
## 3   3 1 0.0000
## 4   4 1 0.0000
## 5   1 2 0.0000
## 6   2 2 0.1250
## 7   3 2 0.0000
## 8   4 2 0.0000
```

Ejemplo dados máximo y suma

```
df[12:28,]
```

```
##    X Y   P_XY
## 12 4 3 0.0000
## 13 1 4 0.0000
## 14 2 4 0.0000
## 15 3 4 0.1250
## 16 4 4 0.1250
## 17 1 5 0.0000
## 18 2 5 0.0000
## 19 3 5 0.0625
## 20 4 5 0.1250
## 21 1 6 0.0000
## 22 2 6 0.0000
## 23 3 6 0.0000
```

Dados máximo y suma

```
df$xP_X=df$X*df$P_XY
df$x2P_X=(df$X)^2*df$P_XY
df$yP_Y=df$Y*df$P_XY
df$y2P_Y=(df$Y)^2*df$P_XY
df$xyP_XY=df$X*df$Y*df$P_XY
colSums(df[,-c(1:3)])
```

```
##    xP_X  x2P_X   yP_Y  y2P_Y xyP_XY
##   3.125 10.625  4.000 18.500 13.750
```

Dados máximo y suma

```
Esp_X=sum(df$xP_X)
Esp_X
```

```
## [1] 3.125
```

```
Esp_Y=sum(df$yP_Y)
Esp_Y
```

```
## [1] 4
```

Dados máximo y suma

```
Esp_X2=sum(df$x2P_X)
Esp_X2
```

```
## [1] 10.625
```

```
Esp_Y2=sum(df$y2P_Y)
Esp_Y2
```

```
## [1] 18.5
```

```
Var_X=Esp_X2-Esp_X^2
Var_X
```

```
## [1] 0.859375
```

Dados máximo y suma

```
Var_Y=Esp_Y2-Esp_Y^2
Var_Y
```

```
## [1] 2.5
```

```
Esp_XY=sum(df$xyP_XY)
Esp_XY
```

```
## [1] 13.75
```

```
Cov_XY=Esp_XY-Esp_X*Esp_Y
Cov_XY
```

```
## [1] 1.25
```

Dados máximo y suma

```
Cor_XY=Cov_XY/sqrt(Var_X*Var_Y)
Cor_XY
```

```
## [1] 0.8528029
```

Vector de medias

$$\begin{pmatrix}\mu_X\\ \mu_Y\end{pmatrix}=\begin{pmatrix}3.125\\ 4\end{pmatrix}.$$

Dados máximo y suma

Matriz de covarianzas

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} 0.859375 & 1.25 \\ 1.25 & 2.5 \end{pmatrix}.$$

Matriz de correlaciones

$$R = \begin{pmatrix} 1 & \rho_{XY} \\ \rho_{XY} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.8528029 \\ 0.8528029 & 1 \end{pmatrix}.$$

Dados ejercicio manual

Ejercicio

- Repetir los cálculos anteriores manualmente,
- Calcular manualmente las distribuciones condicionales
- . . . . TIEMPO

Lección 7

Ejemplo bivariante continua normal bivariante

Definición de distribción normal bivariante

Sea (X, Y) una variable continua bidimensional con $E(X) = \mu_X$, $E(Y) = \mu_X$
$\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$, $\sigma_{XY} = Cov(X, Y)$.

Y si denotamos por

$$\mu = \left(\begin{array}{c} \mu_X \\ \mu_Y \end{array} \right)$$

y por

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{pmatrix}.$$

Definición de distribución normal bivariante

Diremos que el vector $\begin{pmatrix} X \\ Y \end{pmatrix}$ sigue una ley **normal o gaussiana bidimensional**

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

si su densidad es

$$f_{XY}(x,y) = \frac{1}{\sqrt{(2\pi)^2 \cdot \det(\Sigma)}} \cdot e^{-\frac{1}{2}((x,y)-\mu)^t \cdot \Sigma^{-1} \cdot ((x,y)-\mu)}.$$

Gráfica de la distribución gaussiana (X, Y).