Introducción a la Probabilidad y la Estadística

Martes y Jueves Aula B17
Dra Ana Georgina Flesia

Distribución Conjunta: Ejemplo

Una gran tienda de departamentos vende camisas deportivas

1. en tres talles (pequeño, mediano y grande),
2. en tres modelos (a cuadros, estampados y de franjas)
3. con dos largos de mangas (corta y larga).

Si Ω es el conjunto de todas las camisas vendidas, sea

1. X la variable indicadora del talle,
2. Y la variable indicadora del modelo
3. W la variable indicadora del largo de manga.

Distribución Conjunta: Ejemplo

Las siguientes tablas presentan las proporciones de camisas vendidas que caen en varias combinaciones de categorías.

Manga corta			
		Modelo	
Talle	Cuadros	Estampada	Franjas
Pequeño	0,04	0,02	0,05
Mediano	0,08	0,07	0,12
Grande	0,03	0,07	0,08
Manga larga			
		Modelo	
Talle	Cuadros	Estampada	Franjas
Pequeño	0,03	0,02	0,03
Mediano	0,10	0,05	0,07
Grande	0,04	0,02	0,08

Distribución Conjunta: Ejemplo

1. ¿Que representa la tabla?
2. ¿Cuál es la función de densidad puntual de la variable X?
3. ¿Cuál es la función de densidad puntual de la variable Y ?
4. ¿Cuál es la función de densidad puntual de la variable W ?
5. ¿Puedo calcular funciones de las variables a partir de esta información?

$$X + Y + Z; \quad X^2 - Y; \quad [\sin(X) + \cos(Y)]W$$

Distribución Conjunta: Ejemplo

Tabla

La tabla muestra las probabilidades de las intersecciones de todos los valores de las indicadoras.

$$P((\text{pequeño,franjas,manga larga}))$$

Codificación

Codifiquemos las variables, $X = 1$ si el talle es pequeño, $X = 2$ si el talle es mediano y $X = 3$ si el talle es grande. $Y = 1$ si el modelo es a cuadros, $Y = 2$ si el modelo es estampado y $Y = 3$ si el modelo es a franjas; $W = 1$ si la camisa es manga corta y $W = 2$ si la camisa en manga larga.

$$P((\text{pequeño,franjas,manga larga})) = P(X = 1, Y = 3, W = 2)$$

Distribución Conjunta: Definición

Densidad de masa

Sean $(X_1, \cdots, X_n)$ un vector aleatorio, es decir, una n-upla donde cada coordenada es una variable aleatorias discreta sobre el espacio de probabilidad $(\Omega, \mathcal{A}, P)$, se define la función densidad de masa conjunta del vector $(X_1, \cdots, X_n)$ como

$$p(x_1, \cdots, x_n) = P[(X_1 = x_1, \cdots, X_n = x_n)] \quad \forall x_i \in \mathbb{R}$$

Distribución acumulada conjunta

Se define la función de distribución acumulada conjunta del vector $(X_1, \cdots, X_n)$ como

$$F(x_1, \cdots, x_n) = P[(X_1 \leq x_1, \cdots, X_n \leq x_n)] \quad \forall x_i \in \mathbb{R}$$

Distribución Conjunta: Propiedades

Propiedades

La función densidad de probabilidad (densidad de masa) conjunta del vector $(X_1, \cdots, X_n)$ cumple

1. $0 \leq p(x_1, \cdots, x_n) \leq 1$
2. $P[(X_1 = x_1, \cdots, X_n = x_n) \in A] = \sum_{(x_1, \cdots, x_n) \in A} p(x_1, \cdots, x_n)$, $\forall A \subset \mathbb{R}^n$
3. $\sum_{(x_1, \cdots, x_n) \in \mathbb{R}^n} p(x_1, \cdots, x_n) = 1$

Marginales

Observación

- Notemos que conocer la función de masa conjunta nos dice mucho mas que conocer las funciones de masa puntual individuales.
- Consideremos el experimento de tirar una moneda honesta. Las variables indicadoras del evento 'cara' y del evento 'contracara' toman los mismos valores con igual probabilidad. Sin embargo, de la distribución conjunta surge la información de que se relacionan linealmente $X = 1 - Y$.

Distribución Conjunta: Ejemplo

Variable X

1. ¿Cuál es la función de densidad puntual de la variable X?

$P(X = 1)$

$$\begin{aligned} P(X=1) &= P(X=1, Y=1, W=1) + P(X=1, Y=1, W=2) \\ &+ P(X=1, Y=2, W=1) + P(X=1, Y=2, W=2) \\ &+ P(X=1, Y=3, W=1) + P(X=1, Y=3, W=2) \\ &= 0.04 + 0.03 + 0.02 + 0.02 + 0.05 + 0.03 \\ &= 0.19 \end{aligned}$$

Distribución Conjunta: Ejemplo

$P(X = 2)$

$$
\begin{aligned}
P(X = 2) &= P(X = 2, Y = 1, W = 1) + P(X = 2, Y = 1, W = 2) \\
&+ P(X = 2, Y = 2, W = 1) + P(X = 2, Y = 2, W = 2) \\
&+ P(X = 2, Y = 3, W = 1) + P(X = 2, Y = 3, W = 2) \\
&= 0.08 + 0.1 + 0.07 + 0.05 + 0.12 + 0.07 \\
&= 0.49
\end{aligned}
$$

Distribución Conjunta: Ejemplo

$P(X = 3)$

$$
\begin{aligned}
P(X = 3) &= P(X = 3, Y = 1, W = 1) + P(X = 3, Y = 1, W = 2) \\
&+ P(X = 3, Y = 2, W = 1) + P(X = 3, Y = 2, W = 2) \\
&+ P(X = 3, Y = 3, W = 1) + P(X = 3, Y = 3, W = 2) \\
&= 0.03 + 0.04 + 0.07 + 0.02 + 0.08 + 0.08 \\
&= 0.32
\end{aligned}
$$

Distribución Conjunta:Ejemplo

Vector(X, Y)

1. ¿Cuál es la función densidad de masa conjunta del vector (X, Y)?

- La tabla es la función de masa conjunta del vector $(X; Y; W)$. Si unimos los eventos relacionados con los valores de W obtenemos la distribución conjunta de (X, Y)

$$P(X = x, Y = y) = P(X = x, Y = y, W = 1) + P(X = x, Y = y, W = 2)$$

	$p_{X,Y}$		
	$Y = 1$	$Y = 2$	$Y = 3$
$X = 1$	0,07	0,04	0,08
$X = 2$	0,18	0,12	0,19
$X = 3$	0,07	0,09	0,16

Distribución Conjunta:Ejemplo

- Observemos que el centro de la tabla suma 1, como debe ser y todos los valores son no negativos.

	$p_{X,Y}$		
	$Y=1$	$Y=2$	$Y=3$
$X=1$	0,07	0,04	0,08
$X=2$	0,18	0,12	0,19
$X=3$	0,07	0,09	0,16

Distribución Conjunta:Ejemplo

- También podemos observar que la función de masa de X puede ser obtenida sumando cada columna de la tabla, y los valores de la función de masa de Y pueden ser obtenidos sumando a través de las filas de la tabla.

	$p_{X,Y}$		
	$Y=1$	$Y=2$	$Y=3$
$X=1$	0,07	0,04	0,08
$X=2$	0,18	0,12	0,19
$X=3$	0,07	0,09	0,16

Distribución Marginal: Definición

Marginales

Sea $(x_1, \cdots, X_n)$ un vector aleatorio discreto sobre el espacio de probabilidad $(\Omega, \mathcal{A}, P)$, se definen las funciones densidad de masa marginales de X_i como

$$p_{X_i}(x_i) = \sum_{x_j / j \neq i} p(x_1, \cdots, x_n)$$

	$p_{X,Y}$			
	$Y = 1$	$Y = 2$	$Y = 3$	p_X
$X = 1$	0,07	0,04	0,08	0.19
$X = 2$	0,18	0,12	0,19	0.49
$X = 3$	0,07	0,09	0,16	0.32
p_Y	0.32	0.25	0.43	1

Distribución Marginal:Ejemplo

- Como encontramos la varianza y esperanza de las variables X e Y?

	$p_{X,Y}$			
	$Y=1$	$Y=2$	$Y=3$	p_X
$X=1$	0,07	0,04	0,08	0.19
$X=2$	0,18	0,12	0,19	0.49
$X=3$	0,07	0,09	0,16	0.32
p_Y	0.32	0.25	0.43	1

- $E(Y) = \sum x p_Y(y) = 1 \times 0.32 + 2 \times 0.25 + 3 \times 0.43 = 2.11$
- $Var(Y) = E(Y^2) - [E(Y)]^2 =$
  $[1^2 \times 0.32 + 2^2 \times 0.25 + 3^2 \times 0.43] - (2.11)^2 = 0.73$

Esperanza de una función de un vector

Teorema

Sea $(X_1, \cdots, X_n)$ un vector aleatorio n-dimensional. Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función, entonces $g(X_1, \cdots, X_n)$ es una variable aleatoria cuya esperanza es

$$E[g(X_1, \cdots, X_n)] = \sum_{(x_1, \cdots, x_n)} g(x_1, \cdots, x_n) p(X_1, \cdots, X_n)$$

Esperanza de una función de un vector discreto

Corolario

$$E(X+Y) = E(X) + E(Y)$$

Mas generalmente

$$E(g_1(X) + g_2(Y)) = E(g_1(X)) + E(g_2(Y))$$

Este resultado se extiende a un número finito de variables

$$E(X_1 + \cdots + X_n) = \sum_{i=1}^{n} E(X_i)$$

Mas generalmente

$$E(g_1(X_1) + \cdots + g_n(X_n)) = \sum_{i=1}^{n} E(g_i(X_i))$$

Distribución Conjunta: Ejemplo

Supongamos tirar una moneda honesta tres veces. Definimos las siguientes variables aleatorias, X_1 representa el número de caras en la primer tirada, X_2 representa el número de caras en las primeras dos tiradas y X_3 es el total de caras en las tres tiradas.

1. Encuentre la densidad discreta conjunta de X_1 y X_3 y las densidades marginales p_{X_1} y p_{X_3}.
2. Calcule la esperanza de $X_1 + X_3$
3. Encuentre la densidad discreta conjunta de las tres variables.
4. Calcule $P(X_1 + X_3 = 2)$ y $P(X_1 + 1 \leq X_3)$.

Resolución

- Si llamamos cara por c y número por s, el espacio muestral Ω es

$$\Omega = \{(ccc), (ccs), (csc), (scc), (ssc), (scs), (css), (sss)\}$$

- La densidad conjunta entre X_1 y X_3 se calcula de la siguiente forma:

$$\begin{aligned}
p_{X_1,X_3}(0,0) &= P(X_1 = 0, X_3 = 0) = P((sss)) = \frac{1}{8} \\
p_{X_1,X_3}(0,1) &= P(X_1 = 0, X_3 = 1) = P((ssc),(scs)) = \frac{2}{8} \\
p_{X_1,X_3}(0,2) &= P(X_1 = 0, X_3 = 2) = P((scc)) = \frac{1}{8}
\end{aligned}$$

los demás valores se calculan de forma similar.

Resolución

La tabla queda

	p_{X_1,X_3}				
	$X_3 = 0$	$X_3 = 1$	$X_3 = 2$	$X_3 = 3$	p_{X_1}
$X_1 = 0$	1/8	2/8	1/8	0	1/2
$X_1 = 1$	0	1/8	2/8	1/8	1/2
p_{X_3}	1/8	3/8	3/8	1/8	1

La esperanza de $X_1 + X_3$ es

$$\begin{aligned} E[X_1 + X_3] &= [0+0]1/8 + [0+1]2/8 + [0+2]1/8 + [0+3]0 \\ &+ [1+0]0 + [1+1]1/8 + [1+2]2/8 + [1+3]1/8 \\ &= 2 \end{aligned}$$

Resolución

Para calcular la densidad conjunta de las tres variables operamos de la misma forma, solo que tenemos que armar una tabla más complicada

$$
\begin{aligned}
p_{X_1,X_2,X_3}(0,0,0) &= P(X_1=0, X_2=0, X_3=0) = P((sss)) = \frac{1}{8} \\
p_{X_1,X_2,X_3}(1,1,1) &= P(X_1=1, X_2=1, X_3=1) = P((css)) = \frac{1}{8} \\
p_{X_1,X_2,X_3}(1,2,2) &= P(X_1=1, X_2=2, X_3=2) = P((ccs)) = \frac{1}{8}
\end{aligned}
$$

Resolución

La tabla de la distribución conjunta de las tres variables queda

	$X_3 = 0$		$X_3 = 1$		$X_3 = 2$		$X_3 = 3$	
$X_2 \backslash X_1$	0	1	0	1	0	1	0	1
0	1/8	0	1/8	0	0	0	0	0
1	0	0	1/8	1/8	1/8	1/8	0	0
2	0	0	0	0	0	1/8	0	1/8

Resolución

Para calcular $P(X_1 + X_3 = 2)$ y $P(X_1 + 1 \leq X_3)$, basta con ver cuantos duplas de valores satisfacen estos eventos y buscar su probabilidad en la tabla de p_{X_1,X_3}.

$$\begin{aligned} P(X_1 + X_3 = 2) &= P(X_1 = 0, X_3 = 2) + P(X_1 = 1, X_3 = 1) \\ &= p_{X_1,X_3}(0,2) + p_{X_1,X_3}(1,1) \\ &= 1/8 + 1/8 = 1/4 \end{aligned}$$

$$\begin{aligned} P(X_1 + 1 \leq X_3) &= P(X_1 = 0, 1 \leq X_3) + P(X_1 = 1, 2 \leq X_3) \\ &= p_{X_1,X_3}(0,1) + p_{X_1,X_3}(0,2) + p_{X_1,X_3}(0,3) \\ &+ p_{X_1,X_3}(1,2) + p_{X_1,X_3}(1,3) \\ &= 2/8 + 1/8 + 0 + 2/8 + 1/8 = 1/4 \end{aligned}$$

Variables independientes

Independencia

Sea $(X_1, \cdots, X_n)$ un vector aleatorio discreto sobre el espacio de probabilidad $(\Omega, \mathcal{A}, P)$, se dice que las variables X_i son independientes si

$$P[X_1 \in B_1, \cdots, X_n \in B_n] = P(X_1 \in B_1) \times \cdots \times P(X_n \in B_n)$$

Caracterización de la independencia

Independencia

Las variables aleatorias $X_1, \cdots, X_n$ son independientes si y solo si la función de distribución conjunta del vector $(X_1, \cdots, X_n)$ se escribe como producto de las funciones de distribución de cada una de las coordenadas del vector:

$$F(x_1, \cdots, x_n) = F_{X_1}(x_1) \times \cdots \times F_{X_n}(x_n)$$

Criterio de independencia para variables discretas

Independencia

Las variables aleatorias $X_1, \cdots, X_n$ son independientes si y solo si la función de masa (probabilidad puntual) conjunta del vector $(X_1, \cdots, X_n)$ se escribe como producto de las funciones de masa (probabilidad puntual) de cada una de las coordenadas del vector:

$$p(x_1, \cdots, x_n) = p_{X_1}(x_1) \times \cdots \times p_{X_n}(x_n)$$

Lema

Independencia y Esperanza

Si X e Y son independientes entonces

$$E(XY) = E(X)E(Y)$$

Mas generalmente

$$E(h_1(X)h_2(Y)) = E(h_1(X))E(h_2(Y))$$

Covarianza

Definición

Dadas dos variables aleatorias X e Y definimos la covarianza entre ellas mediante la fórmula

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

Puede probarse usando las propiedades de la esperanza que

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

Lema

Si X e Y son independientes entonces

$$Cov(X,Y) = 0$$

Covarianza

Propiedades

Sea (X, Y) un vector aleatorio con densidad conjunta f_{XY}, entonces

1. $Cov(X, X) = V(X)$
2. $Cov(X, Y) = E(XY) - E(X)E(Y)$
3. Si X e Y son independientes, entonces $Cov(X, Y) = 0$.
4. $Cov(aX + b, cY + d) = acCov(X, Y)\ \forall a, b, c, d \in \mathbb{R}$
5. $Var(aX + bY) = a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$ $\forall a, b \in \mathbb{R}$
6. Si X e Y son independientes entonces $Var(aX + bV) = a^2Var(X) + b^2Var(Y)$

Correlación

Definición

Sean X e Y variables aleatorias con $0 < Var(X) < \infty, 0 < Var(Y) < \infty$. Definimos el coeficiente de correlación lineal como

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

Correlación

Propiedades

1. El coeficiente de correlación lineal es independiente de posición y escala, pues

$$\rho(X+a, Y+a) = \rho(X, Y) \qquad \rho(aX, aY) = \rho(X, Y), a > 0$$

2. $-1 \leq \rho(X, Y) \leq 1$
3. $\rho(X, Y) = 1$ si y solamente si $P(Y = aX + b) = 1$, con $a > 0$.
   $\rho(X, Y) = -1$ si y solamente si $P(Y = aX + b) = 1$, con $a < 0$.

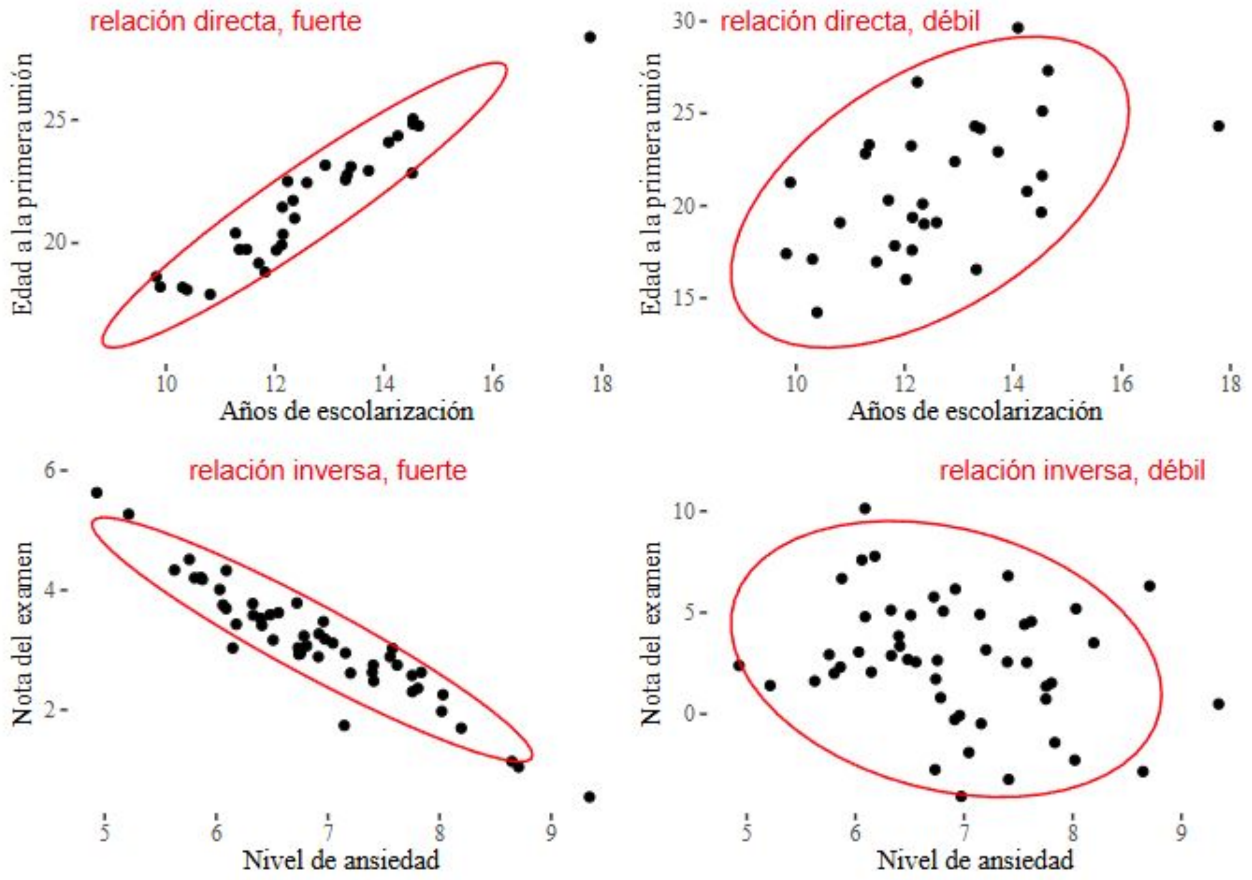

Figura 5.4: Comparación de la forma de las nubes de puntos según la intensidad de la relación