Probabilidad conjunta

Gamaliel Moreno Chávez

MCPI

Ago-Dic
2020

Distribuciones de probabilidad conjunta

Hay situaciones en las que se busque registrar los resultados simultáneos de diversas variables aleatorias. Por ejemplo, en un experimento químico controlado podríamos medir la cantidad del precipitado P y la del volumen V de gas liberado, lo que daría lugar a un espacio muestral bidimensional que consta de los resultados (p, v); o bien, podríamos interesarnos en la dureza d y en la resistencia a la tensión T de cobre estirado en frío que produciría los resultados (d, t).

Distribuciones de probabilidad conjunta

Solo consideraremos el caso de dos variables aleatorias discretas X e Y que forman una variable aleatoria bidimensional denotada por (X, Y), que puede tomar pares de valores (x_i, y_j): por ejemplo, podríamos asumir $(i = 1, 2, \ldots; j = 1, 2, \ldots)$ (Cualquier secuencia puede ser finita o infinita, o comenzar para algunos otros valores) con la función de probabilidad conjunta o función de masa de probabilidad $p(x_i, y_j)$, que ahora es una función de dos variables.

Distribuciones de probabilidad conjunta

Las probabilidades deben satisfacer

- $0 \leq p(x_i, y_j) \leq 1 \quad \forall (i, j)$
- $\displaystyle\sum_{i \in I} \sum_{j \in J} p(x_i, y_j) = 1$

Los dominios de i y j están definidos por $i \in I$ y $j \in J$ donde I y J pueden ser secuencias limitadas o ilimitadas de los números enteros.

Las variables aleatorias X y Y son independientes si y sólo si

$$p(x_i, y_j) = q(x_i) r(y_j) \quad \forall (i, j)$$

donde

$$\sum_{i \in I} q(x_i) = 1 \quad \sum_{j \in J} r(y_j) = 1$$

Valor esperado condicional

Si la variable aleatoria $Z = H(X, Y)$ es una función de las variables aleatorias X y Y, entonces el valor esperado de Z viene dado por

$$E(Z) = \sum_{i \in I} \sum_{j \in J} H(x_i, y_j) p(x_i, y_j)$$

Consideramos la distribución conjunta de dos variables aleatorias X e Y con función de masa conjunta $p(x_i, y_j)$. Suponemos que podemos asociarnos con probabilidades condicionales X e Y

$$P(X = x_i | Y = y_j) \text{ y } P(Y = y_j | X = x_i).$$

otra notación equivalente

$$p_X(x_i | y_j) \text{ o simplemente } P(X|Y).$$

$$p_Y(y_j | x_i) \text{ o simplemente } P(Y|X).$$

Valor esperado condicional

El valor condicional de X para cada elemento de Y es

$$E(X|Y = y_j) = \sum_{i \in I} x_i P(X = x_i|Y = y_j) \text{ o } = \sum_{i \in I} x_i p_X(x_i|y_j).$$

La importancia de esta expectativa es que tiene valor para cada elemento de J. La totalidad de estos valores es $E[X|Y]$, que es una variable aleatoria. Tendrá el mismo número de elementos que J. Del mismo modo

$$E(Y|X = x_i) = \sum_{j \in J} y_j p_Y(y_j|x_i).$$

Probabilidad condicional

Las probabilidades condicionales están dadas por

$$p_X(x_i|y_j) = p(x_i, y_j)/\sum_{i \in I} p(x_i, y_j)$$

$$p_Y(y_j|x_i) = p(x_i, y_j)/\sum_{j \in J} p(x_i, y_j)$$

En la primera expresión, $p(x_i, y_j)$ es la probabilidad de que ocurran X e Y, y el denominador es la probabilidad de que ocurra Y.Se aplica una interpretación similar a la segunda. Las probabilidades $\sum_{i \in I} p(x_i, y_j)$ y $\sum_{j \in J} p(x_i, y_j)$ se conocen como distribuciones de probabilidad marginal.

Probabilidad condicional

Si X e Y son independientes, entonces las ecuaciones anteriores se convierten en

$$p_X(x_i, y_j) = q(x_i), \quad p_Y(y_j, x_i) = r(y_j)$$

con valores esperados

$$E(X|Y) = \sum_{i \in I} x_i q(x_i) = E(x)$$

$$E(Y|X) = \sum_{j \in J} y_j r(y_j) = E(y)$$

Distribuciones de probabilidad conjunta

Las esperanzas condicionales $E(X|Y)$ y $E(Y|X)$, al ser variables aleatorias, también tendrán expectativas.

$$\begin{aligned} E(E(X|Y)) &= \sum_{j\in J} E(X|Y) \sum_{k\in I} p(x_k, y_j) \\ &= \sum_{j\in J} \sum_{i\in I} x_i p_X(x_i, y_j) \sum_{k\in I} p(x_k, y_j) \\ &= \sum_{j\in J} \sum_{i\in I} x_i p(x_i, y_j) \\ &= E(X) \end{aligned}$$

Distribuciones de probabilidad conjunta

Ejemplo

Sea la variable aleatoria (X, Y) tome los valores (x_i, y_j) $(i = 1, 2, 3; j = 1, 2, 3)$. Las funciones de masa $p(x_i, y_j)$ se dan en la tabla.

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Encontrar la variable aleatoria $E(X|Y)$, y verifique $E(E(X|Y)) = E(X)$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Notese de la tabla

$$I = \{1, 2, 3\}$$

$$J = \{1, 2, 3\}$$

$$p(x_1, y_3) = 0.05 \quad p(x_3, y_2) = 0.25$$

$$\sum_{i=1}^{3} \sum_{j=1}^{3} p(x_i, y_j) = 1$$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Las distribuciones de probabilidad marginal para Y tendrán cada una tres componentes, a saber

$$\sum_{i=1}^{3} p(x_i, y_1) = p(x_1, y_1) + p(x_2, y_1) + p(x_3, y_1) = 0.25 + 0.05 + 0.05 = 0.35$$

$$\sum_{i=1}^{3} p(x_i, y_2) = 0 + 0.01 + 0.25 = 0.35$$

$$\sum_{i=1}^{3} p(x_i, y_3) = 0.05 + 0.15 + 0.10 = 0.3.$$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

DE manera similar para X

$$\sum_{j=1}^{3} p(x_1, y_j) = 0.3$$

$$\sum_{j=1}^{3} p(x_2, y_j) = 0.3$$

$$\sum_{j=1}^{3} p(x_3, y_j) = 0.4$$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Calculo de las probabilidades condicionales

$$p_X(x_1|y_1) = p(x_1, y_1)/\sum_{i=1}^{3} p(x_i, y_1) = 0.25/0.35 = 5/7,$$

$$p_X(x_2|y_1) = 0.05/0.35 = 1/7, \quad p_X(x_3|y_1) = 0.05/0.35 = 1/7,$$

$$p_X(x_1|y_2) = 0, \quad p_X(x_2|y_2) = 2/7, \quad p_X(x_3|y_2) = 5/7,$$

$$p_X(x_1|y_3) = 1/6, \quad p_X(x_2|y_3) = 1/2, \quad p_X(x_3|y_3) = 1/3,$$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Con esto podemos encontrar la esperanza condicional

$$E(X|Y=y_1)=\sum_{i=1}^{3} x_i p(x_i|y_1)=(1)(7/5)+(2)(1/7)+(3)(1/7)=10/7,$$

$$E(X|Y=y_2)=\sum_{i=1}^{3} x_i p(x_i|y_2)=19/7,$$

$$E(X|Y=y_3)=\sum_{i=1}^{3} x_i p(x_i|y_3)=13/6,$$

Distribuciones de probabilidad conjunta

$p(x_i, y_j)$	y_1	y_2	y_3
x_1	0.25	0	0.05
x_2	0.05	0.10	0.15
x_3	0.05	0.25	0.10

Así la variable aleatoria $E(X|Y)$ toma los valores

$$\left\{\frac{10}{7}\frac{19}{7}\frac{13}{6}\right\} = Z = \{z_1, z_2, z_3\}$$

Entonces el valor esperado de $E(X|Y)$ o de Z es

$$\begin{aligned} E(Z) &= \sum_{j=1}^{3} z_j \sum_{i=1}^{3} p(x_i, y_j) \\ &= (10/7)(0.35) + (19/7)(0.35) + (13/6)(0.3) = 21/10 \end{aligned}$$

igualmente

$$E(X) = \sum_{i=1}^{3}\sum_{j=1}^{3} x_i p(x_i, y_j) = (1)(0.3) + (2)(0.3) + (3)(0.4) = 21/10$$