DISTRIBUCIÓN CONDICIONAL

Nexus-Probability

CURSO 2 (PROBABILIDAD II)

PARTE 2 / LECCIÓN 2

Como hemos revisado en lecciones pasadas, la probabilidad condicional de un evento A dado un evento B está dada por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Esta definición es relevante para el caso de funciones de probabilidad o de densidad y también para el caso de funciones de distribución.

Definición 1 *Sea (X, Y) un vector aleatorio continuo (o discreto) y con función de probabilidad (o densidad) $f_{X,Y}(x, y)$. Sea y un valor de la variable Y tal que $f_Y(y) \neq 0$. A la función $x \to f_{X|Y}(x|y)$ definida a continuación se le llama la **función de probabilidad (o densidad** de X dado que Y=y,*

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

Observe que esta función es considerada como una función de x y que el valor de y es fijo y puede considerarse como un parámetro de dicha función, es decir, para cada valor fijo de y se tiene una función diferente.
Para el caso discreto,

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

Es importante mencionar que en el caso continuo las espresiones $f_{X,Y}(x, y)$ y $f_Y(y)$ no son probabilidades.
Para el caso donde X y Y son independientes,

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

Definición 2 *Sea (X, Y) un vector aleatorio con función de probabilidad o densidad $f_{X,Y}(x, y)$. Sea y un valor de Y tal que $f_Y(y) \neq 0$. La* ***función de distribución condicional*** *de X dado $Y = y$ es la función,*

$$F_{X|Y}(x|y) = \begin{cases} \sum_{u \leq x} f_{X|Y}(u|y) & \text{en el caso discreto,} \\ \int_{-\infty}^{x} f_{X|Y}(u|y) & \text{en el caso continuo} \end{cases}$$

De esta manera la función de distribución condicional se calcula como la suma o integral de la correspondiente función de probabilidad o densidad condicional. Veamos que cuando X y Y son independientes,

$$F_{X|Y}(x|y) = F_X(x)$$

Definimos a continuación el valor esperado de una v.a dado que un evento ha ocurrido para otra v.a cuando se conoce la distribución conjunta de las dos variables.

Definición 3 (Esperanza condicional) *Sea (X, Y) un vector aleatorio continuo con función de densidad $f_{X,Y}(x, y)$ y sea y un valor tal que $f_y(y) \neq 0$. La* ***esperanza condicional*** *de X dado $Y = y$ es la esperanza de la función de densidad condicional $f_{X|Y}(x|y)$, cuando existe, es decir,*

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

Efectuando un cambio en el orden de las integrales es inmediato comporbar que

$$E[X|Y = y] = \sum_x x f_{X|Y}(x|y) = \sum_x x P(X = x|Y = y)$$

si suponemos que $f_Y(y) \neq 0$ y que la suma es absolutamente convergente. Por otro lado, efectuando un cambio en el orden de las sumas tenemos la siguiente expresión

$$E[X] = \sum_y x f_{X|Y}(x|y) = \sum_x x E[X|Y = y] P(Y = y)$$

En el caso donde X y Y son independientes,

$$E[X|Y = y] = E[X]$$

Es importante hablar de la **covarianza** entre variables aleatorias. Sean X y Y dos variables aleatorias con esperanza finita y con función de densidad o de probabilidad conjunta $f(x, y)$. La covarianza entre X y Y es un número real que se denota por $Cov(X, Y)$ y se define como la esperanza de la variable aleatoria $(X - E[X])(Y - E[Y])$.

Definición 4 *La covarianza entre las variables aleatorias X y Y es el número real*

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

La covarianza está estrechamente relacionada con otro concepto que se define para dos variables aleatorias llamado coeficiente de correlación, y para el cual se cuente con una interpretación. Cuando X y Y son variables aleatorias discretas la covarianza se calcula como sigue,

$$Cov(X,Y) = \sum_{x,y}(x - E(X))(y - E(y))f(x,y)dxdy$$

Por otro lado, el caso donde las variables aleatorias son continuas, tenemos que

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - E(X))(y - E(Y))f(x,y)dxdy$$

De otra manera, también puede ser expresado de la siguiente forma:

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

Es importante mencionar que la covarianza es simétrica, es decir, $Cov(X,Y) = Cov(Y,X)$ En el caso donde X y Y son independientes, entonces $Cov(X,Y) = 0$, el recíproco es falso.
Por otro lado, cuando hablamos de varianza,

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

Además,

$$-\sqrt{Var(X)Var(Y)} \leq Cov(X,Y) \leq \sqrt{Var(X)Var(Y)}$$

Definición 5 (Coeficiente de correlación) *El coeficiente de correlación entre las variables aleatorias X y Y con varianzas finitas distintas de cero se define como el número*

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

A este número también podemos denortarlo como $\rho_{X,Y}$.

Definición 6 *El coeficiente de correlación entre dos variables aleatorias X y Y con varianzas finitas distintas de cero satisface las siguientes dos desigualdades:*

$$-1 \leq \rho(X,Y) \leq 1$$

Ejercicios

Algunos de los siguientes ejercicios propuestos tendrán solución en **Python**, por lo que te invitamos a ejecutar el código en tu computadora, cuando sea el caso.

Ejercicio 1 *Sea (X, Y) un vector aleatorio continuo con función de probabilidad dada por la siguiente tabla,*

x/y	0	1	2	3
0	0.1	0.1	0.2	0.1
1	0.1	0.2	0.1	0.1

Calcule la función de probabilidad condicional $f_{X|Y}(x|y)$ para $y = 1$.

Solución.

Sumando las probabilidades de la columna correspondientes a $y = 1$ se encuentra que $f_Y(y) = 0.3$. Por lo tanto, siguiendo la fórmula de la expresión tenemos que,

$$f_{X|Y}(x|1) = \frac{f_{X,Y}(x,1)}{f_Y(1)} = \begin{cases} 1/3 & \text{si } x = 0 \\ 2/3 & \text{si x=1} \\ 0 & \text{en otro caso.} \end{cases}$$

De manera análoga pueden calcularse $f_{X|Y}(x|0), f_{X|Y}(x|2), f_{X|Y}(x|3)$ y también $f_{X|Y}(y|0)$ y $f_{X|Y}(y|1)$.

Ejercicio 2 *Sea (X,Y) un vector aleatorio continuo con función de densidad dada por,*

$$f_{X,Y}(x,y) = \begin{cases} x + y & \textit{si } 0 < x < 1, 0 < y < 1 \\ 0 & \textit{en otro caso.} \end{cases}$$

Calcularemos la función de probabilidad condicional $f_{X|Y}(x|y)$ para cada y en el intervalo $(0,1)$. Integrando sobre x tenemos que la función de densidad marginal de Y es

$$f_Y(y) = \begin{cases} \frac{1}{2} + y & \textit{si } 0 < y < 1 \\ 0 & \textit{en otro caso.} \end{cases}$$

Por lo tanto, para cada $y \in (0,1)$ fijo, la función de densidad condicional X dado $Y = y$ está dada por

$$x \mapsto f_{X|Y}(y|x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{x+y}{1/2+y} & \textit{si } 0 < x < 1 \\ 0 & \textit{en otro caso.} \end{cases}$$

De manera análoga puede calcularse $y \mapsto f_{Y|X}(y|x)$ oara cada $x \in (0,1)$

Ejercicio 3 *Dada la función de probabilidad conjunta:*

$$f(x,y)$$

x/y	*0*	*1*
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$

Calculamos las distribuciones marginales:

$$P(X=0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \quad P(X=1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(Y=0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, \quad P(Y=1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Cálculo de las esperanzas:

$$E[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

$$E[Y] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$$

Cálculo de las varianzas:

$$E[X^2] = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\textit{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

$$E[Y^2] = 0^2 \cdot \frac{1}{2} + 1^2 \cdot \frac{1}{2} = \frac{1}{2}$$

$$\textit{Var}(Y) = E[Y^2] - (E[Y])^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

Cálculo de la covarianza:

$$E[XY] = 0 \cdot 0 \cdot \frac{1}{4} + 0 \cdot 1 \cdot \frac{1}{4} + 1 \cdot 0 \cdot \frac{1}{4} + 1 \cdot 1 \cdot \frac{1}{4} = \frac{1}{4}$$

$$\textit{Cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{4} - \left(\frac{1}{2} \cdot \frac{1}{2}\right) = 0$$

Finalmente, el coeficiente de correlación:

$$\rho_{X,Y} = \frac{\textit{Cov}(X,Y)}{\sqrt{\textit{Var}(X) \cdot \textit{Var}(Y)}} = \frac{0}{\sqrt{\frac{1}{4} \cdot \frac{1}{4}}} = 0$$

Por lo tanto, X y Y no están correlacionadas.