Tema 4: Variables aleatorias multidimensionales

Los contenidos a desarrollar en este tema son los siguientes:

- Distribución conjunta de probabilidad
- Probabilidad/densidad marginales y condicionadas
- Independencia
- Covarianza, correlación
- Esperanza de $g(X, Y)$, esperanza y varianza de suma/diferencia

Lecturas recomendadas:

- Capítulos 5 y 6 del libro de Newbold, Carlson y Thorne (2009).
- Capítulo 6 del libro de Peña (2001).

Distribución conjunta de probabilidad

Distribución conjunta de las variables aleatorias **discretas** X e Y es la función $p(x, y)$ que expresa la probabilidad simultánea de que X tome el valor x e Y tome el valor y

$$p(x, y) = P(X = x, Y = y)$$

Propiedades

- $0 \leq p(x, y) \leq 1$ para todos los valores de x e y
- $\sum_x \sum_y p(x, y) = 1$

Distribución conjunta de probabilidad - Ejemplo

Ejemplo

X = número de tarjetas de crédito

Y = número de compras por semana

La función de **masa conjunta** de probabilidad es

$X \setminus Y$	0	1	2	3
1	0,08	0,07	0,04	0,01
2	0,10	0,35	0,26	0,09

Distribución conjunta de probabilidad

Sean X e Y variables aleatorias **continuas**, su función de **densidad conjunta** de probabilidad, $f(x, y)$, es una función con las **propiedades**

- $f(x, y) \geq 0$ para todos los valores de x e y
- el volumen debajo de la función de densidad de probabilidad cuando se abarcan todos los valores e X e Y es 1 ($\int \int f(x, y) dx dy = 1$)

Ejemplo $f(x, y) = x + y$, donde $0 \leq x \leq 1$, $0 \leq y \leq 1$

$$\int \int f(x, y) dx dy = \int_0^1 \int_0^1 (x + y) dx dy = \int_0^1 \left[\frac{x^2}{2} + yx \right]_0^1 dy$$

$$= \int_0^1 \left(\frac{1}{2} + y \right) dy = \left[\frac{y}{2} + \frac{y^2}{2} \right]_0^1 = 1$$

Probabilidad marginal de variables discretas

La función de **probabilidad marginal** de la variable **discreta** X, es la función de probabilidad que se obtiene **sumando** las probabilidades conjuntas correspondientes a todos los valores posibles de la variable Y:

$$p_x(x) = \sum_y p(x, y)$$

Analógamente, para la variable Y, se obtiene sumando en todos los valores posibles de la variable X:

$$p_y(y) = \sum_x p(x, y)$$

Probabilidad marginal de variables discretas - Ejemplo

Ejemplo

$X =$ número de tarjetas de crédito

$Y =$ número de compras por semana

Las funciones de masa **marginales** son

X	1	2	
x	0,2	0,8	1

Y	0	1	2	3	
y	0,18	0,42	0,30	0,10	1

Densidad marginal de variables continuas

La función de **densidad marginal** de la variable **continua** X es

$$f_x(x) = \int f(x,y)dy$$

Al igual con la variable Y

$$f_y(y) = \int f(x,y)dx$$

Ejemplo $f(x,y) = x + y$, donde $0 \le x \le 1$, $0 \le y \le 1$

$$f_x(x) = \int_0^1 (x+y)dy = \left[xy + \frac{y^2}{2}\right]_0^1 = x + \frac{1}{2},\ \ 0 \le x \le 1$$

$$f_y(y) = \int_0^1 (x+y)dx = \left[\frac{x^2}{2} + yx\right]_0^1 = \frac{1}{2} + y,\ \ 0 \le y \le 1$$

Esperanza y varianza de la variable aleatoria

- Para las variables **discretas**

$$
\begin{aligned}
E[X] &= \mu_x = \sum_x x p_x(x) \\
V[X] &= \sigma_x^2 = \sum_x (x - E[X])^2 p_x(x) = E[X^2] - (E[X])^2 \\
DT[X] &= \sigma_x = \sqrt{\sigma_x^2}
\end{aligned}
$$

Ejemplo

$$
\begin{aligned}
E[X] &= 1(0,2) + 2(0,8) = 1,8 \\
V[X] &= (1^2(0,2) + 2^2(0,8)) - (1,8)^2 = 0,16 \\
DT[X] &= \sqrt{0,16} = 0,4
\end{aligned}
$$

Esperanza y varianza de la variable aleatoria

- Para las variables **continuas**

$$E[X] = \int_{-\infty}^{\infty} x f_x(x) dx$$

$$V[X] = \int_{-\infty}^{\infty} (x - E[X])^2 f_x(x) dx = E[X^2] - (E[X])^2$$

Ejemplo

$$E[X] = \int_0^1 x \overbrace{(x + \frac{1}{2})}^{f_x(x)} dx = 7/12$$

$$V[X] = \int_0^1 x^2 \overbrace{(x + \frac{1}{2})}^{f_x(x)} dx - (7/12)^2 = 11/144$$

Propiedades de esperanza y varianza

Esperanza (a, b, c constantes)

- $E[c] = c$
- $E[cX] = cE[X]$
- $E[X + Y] = E[X] + E[Y]$

Varianza

- $V[c] = 0$
- $V[\mathbf{c}X] = \mathbf{c^2}V[X]$
- $V[\mathbf{a} + \mathbf{b}X] = \mathbf{b^2}V[X]$

Probabilidad condicionada de variables discretas

La función de **probabilidad condicionada** de la variable aleatoria Y **discreta**, dado que la variable aleatoria discreta X toma el valor x, expresa la probabilidad de que Y tome el valor y cuando se especifica al valor x para X

$$p(y|x) = \frac{p(x,y)}{p_x(x)} \qquad \left(p(x|y) = \frac{p(x,y)}{p_y(y)} \right)$$

Ejemplo
Calcularemos la distribución del número de tarjetas de crédito entre los individuos que realizan 2 compras a la semana $p(x|y=2)$

X	1	2	
$p(x\|y=2)$	0,13 $\left(= \frac{0{,}04}{0{,}30}\right)$	0,87 $\left(= \frac{0{,}26}{0{,}30}\right)$	1

Probabilidad condicionada de variables continuas

La función de **densidad condicionada** de la variable aleatoria Y **continua** dado X,

$$f(y|x) = \frac{f(x,y)}{f_x(x)} \qquad \left(f(x|y) = \frac{f(x,y)}{f_y(y)} \right)$$

Ejemplo

$$f(x|y=0,2) = \frac{f(x;0,2)}{f_y(0,2)} = \frac{x+0,2}{0,7}, \;\; 0 \leq x \leq 1$$

- A partir de éstas funciones se pueden calcular la media y la varianza condicionadas.

Esperanza y varianza condicionada

- Variables **discretas**

Ejemplo

$$E[X|Y=2] = 1(0,13) + 2(0,87) = 1,87$$

$$V[X|Y=2] = (1^2(0,13) + 2^2(0,87)) - (1,87)^2 = 0,11$$

- Variables **continuas**

Ejemplo

$$E[X|Y=0,2] = \int_0^1 x \left(\overbrace{\frac{x+0,2}{0,7}}^{f(x|y=0,2)} \right) dx = 0,61$$

$$V[X|Y=0,2] = \int_0^1 x^2 \left(\frac{x+0,2}{0,7} \right) dx - (0,61)^2 = 0,08$$

Independencia

- Se dice que las variables aleatorias X e Y son **independientes** si y sólo si su función de probabilidad/densidad conjunta es el producto de su probabilidad/densidad marginal

$$p(x,y) = p_x(x)p_y(y) \qquad \Big(f(x,y) = f_x(x)f_y(y)\Big)$$

para todos los valores x e y

- Alternativamente

$$\begin{array}{ll} p(y|x) = p_y(y) & \big(f(y|x) = f_y(y)\big) \\ p(x|y) = p_x(x) & \big(f(x|y) = f_x(x)\big) \end{array}$$

para todos los valores x e y

Asociación lineal

Covarianza mide la asociación **lineal** entre las variables

$$Cov(X,Y) = \sigma_{xy} = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

Ejemplo

$$\begin{aligned} Cov(X,Y) &= 1 \cdot 0 \cdot 0,08 + 1 \cdot 1 \cdot 0,07 + 1 \cdot 2 \cdot 0,04 + 1 \cdot 3 \cdot 0,01 \\ &+ 2 \cdot 0 \cdot 0,1 + 2 \cdot 1 \cdot 0,35 + 2 \cdot 2 \cdot 0,26 + 2 \cdot 3 \cdot 0,09 \\ &- \overbrace{(1,8)(1,32)}^{E[X]E[Y]} = 0,33 - 0,44 = -0,11 \end{aligned}$$

Asociación lineal

Sean X e Y variables aleatorias, la **correlación** entre X e Y es

$$Cor(X,Y) = r_{xy} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$$

- $-1 \leq Cor(X,Y) \leq 1$
- Si la correlación es 0, indica que no existe relación **lineal** entre las variables.
- Si las variables son independientes, la correlación es 0. El recíproco no es cierto.
- Una correlación **positiva** indica que, si una variable aleatoria es **alta** (**baja**), la otra tiene una probabilidad mayor de ser **alta** (**baja**).
- Una correlación **negativa** indica que, si una variable aleatoria es **alta** (**baja**), la otra tiene una probabilidad mayor de ser **baja** (**alta**).

Esperanza de funciones de variables

Sean X e Y variables aleatorias, la esperanza de cualquier función $g(X,Y)$ de estas variables se define en el caso de variables discretas como:

$$E[g(X,Y)] = \sum_{x} \sum_{y} g(x,y)p(x,y)$$

y en el caso de variables continuas como:

$$E[g(X,Y)] = \int \int g(x,y)f(x,y)dxdy$$

Esperanza y varianza de funciones lineales

Propiedades

- $E[X+Y] = E[X] + E[Y]$
- $E[X-Y] = E[X] - E[Y]$
- $V[X \pm Y] = V[X] + V[Y] \pm Cov(X,Y)$
  Sí $Cov(X,Y) = 0$
  $$V[X \pm Y] = V[X] + V[Y]$$
- $E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n]$
- $V[X_1 + X_2 + \ldots + X_n] = V[X_1] + V[X_2] + \ldots + V[X_n]$ si la **covarianza entre cada par de variables es 0**