

Estadística matemática

Wilfredo Gallegos

25 de julio de 2023

Jueves 6 de julio

1. Contenido

5.1

Definición 1.1. Sean Y_1 y Y_2 variables aleatorias discretas. La función de probabilidad conjunta para Y_1 y Y_2 es $P(y_1, y_2) = P(Y_1 = y_1) - \inf < y_2 < \inf$

martes 11 de julio

Definición 1.2. Si Y_1 y Y_2 son variables aleatorias discretas conjuntas con función de probabilidad conjunta $P(y_1, y_2)(P[Y_1 = y_1] \cap [Y_2 = y_2])$ y función marginal $P_1(y_1)(P[Y_1 = y_1])$ y $P_2(y_2)(P[Y_2 = y_2])$

$$\Rightarrow p(y_1|y_2) = P([Y_1 = y_1] | [Y_2 = y_2]) = \frac{P([Y_1 = y_1] \cap [Y_2 = y_2])}{P([Y_2 = y_2])} = \frac{P(y_1, y_2)}{P_2(y_2)}, \quad \text{Siempre que } P_2(y_2) > 0$$

Nota:

$$p(y_2|y_1) = \frac{P(y_1, y_2)}{P_1(y_1)} \geq 0 \text{ y } \sum p(y_1|y_1) = 1$$

5.7 ej. 5.5

Encontrar la distribución condicional de Y_1 dado $Y_2 = 1$

Y_1 : # de republicanos en el comite de 2 personas

Y_2 : # de demócratas en el comite de 2 personas

$$p(Y_2 = 0 | Y_2 = 1) = \frac{\frac{\binom{3}{0}\binom{2}{1}\binom{1}{1}}{\binom{6}{2}}}{\frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}}} = \frac{2/15}{8/15} = \frac{1}{4}$$
$$p(Y_2 = 1 | Y_2 = 1) = \frac{\frac{\binom{3}{1}\binom{2}{1}\binom{0}{0}}{\binom{6}{2}}}{\frac{\binom{2}{1}\binom{4}{1}}{\binom{6}{2}}} = \frac{6/15}{8/15} = \frac{3}{4}$$

5.6

Definición 1.3. Si Y_1 y Y_2 son variables aleatorias continuas con función densidad conjunta $f(y_1, y_2) \Rightarrow$ la función de distribución acumulada condicional de Y_1 , dado $Y_2 = y_2$ es $F(y_1|y_2) = P(Y_1 \leq y_1 | Y_2 = y_2)$

5.7

Definición 1.4. Sean Y_1 y Y_2 son variables aleatorias continuas con función densidad conjunta $f(y_1, y_2)$ y densidades marginales $f_1(y_1)$ y $f_2(y_2)$. Para cualquier y_2 tal que $f_2(y_2) > 0$, la densidad condicional de Y_1 , dado $Y_2 = y_2$ es

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)}$$

y para cualquier y_1 tal que $f_1(y_1) > 0$, la densidad condicional de Y_2 dada $Y_1 = y_1$ es

$$f(y_2|y_1) = \frac{f(y_1, y_2)}{f_1(y_1)}$$

5.8 ej.

Y_2 : existencia de bebidas al inicio de un día determinado

Y_1 : despacho durante el día

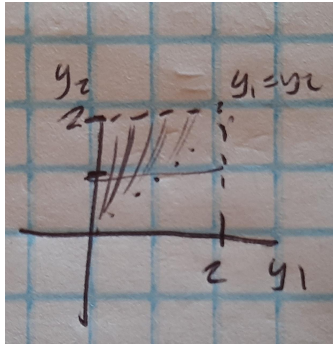
$$Y_1 \leq Y_2$$

$$f(y_1, y_2) = \begin{cases} 1/2, & 0 \leq y_1 \leq y_2 \leq 2 \\ 0, & \text{en caso contrario} \end{cases}$$

Encontrar la densidad condicional de Y_1 dado $Y_2 = y_2$ y calcular la probabilidad de que se venda menos o igual a 1/2 galón, dado que la máquina contiene 1.5 galones al empezar el día.

Sol:

$$f_2(y_2) = \begin{cases} \int_{-\infty}^{\infty} f(y_1, y_2) dy_1 = \int_0^{y_2} \frac{1}{2} dy_1 = \frac{1}{2} y_1 \Big|_0^{y_2} = \frac{1}{2} y_2 & , \quad 0 \leq y_2 \leq 2 \\ 0 & , \quad \text{Cualquier otro caso} \end{cases}$$



Note que $f_2(y_2) > 0$ en $0 < y_1 \leq 2$

$$f(y_1|y_2) = \frac{f(y_1, y_2)}{f_2(y_2)} \frac{1/2}{1/2 y_2} = \frac{1}{y_2}, 0 < y_1 \leq y_2 \leq 2$$

Además, $f(y_1|y_2)$ es indefinida si $y_2 \leq 0$ o $y_2 > 2$

5.8 ej.

$$\begin{aligned} P(Y_1 \leq 1/2 | Y_2 = 1,5) &= \int_{-\infty}^{1/2} f(y_1 | y_2 = 1,5) dy_1 \\ &= \int_{-\infty}^{1/2} \frac{1}{1,5} dy_1 = \frac{1}{1,5} y_1 \Big|_0^{1/2} = \frac{1/2}{1,5} = \frac{1}{3} \end{aligned}$$

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

5.7 Variables aleatorias independientes

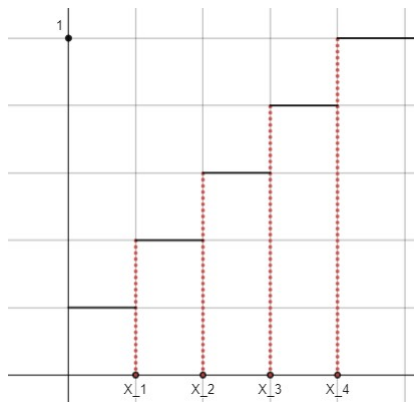
5.8

Definición 1.5. Sean Y_1 con una función distribución acumulada $F_1(y_1)$ y Y_2 con función de distribución acumulada $F_2(y_2)$ y $F(y_1, y_2)$ es la función de distribución acumulada conjunta de Y_1 y Y_2 , entonces Y_1 y Y_2 son independientes ssi $F(y_1, y_2) = F_1(y_1)F_2(y_2) \forall (y_1, y_2) \in \mathbb{R}^2$ Si Y_1 y Y_2 no son independientes son dependientes

Teorema 1.1. Dado una variable aleatoria $X \Rightarrow P([X = a]) = \lim_{n \rightarrow 0} P([a - n < x \leq a])$

Demostración. Sea $a \in \mathbb{R}$

$$\text{Como } \lim_{n \rightarrow 0} a - n = a \left(\lim_{n \rightarrow \infty} a - \frac{1}{h} \right) = a$$



Tomese una sucesión

$$\frac{t_1}{a-1} < \frac{t_2}{a-\frac{1}{2}} < \dots < \frac{t_n}{a-\frac{1}{n}} < a$$

tal que $\lim_{n \rightarrow \infty} t_n = 0$ y $A_n = [t_n < t \leq a]$

$$\begin{aligned} A_1 &= [a - 1 < t \leq a] \\ A_2 &= [a - 1/2 < t \leq a], A_2 \subseteq A_1 \\ &\vdots \\ &\vdots \\ A_n &= [a - 1/n < t \leq a], A_n \subseteq A_{n-1} \end{aligned}$$

Luego

$$\begin{aligned} \lim_{n \rightarrow \infty} P([a - n < x \leq a]) &= \lim_{n \rightarrow \infty} P([t_n < x \leq a]) \\ &= \lim_{n \rightarrow \infty} P([x \leq a]) - P([x \leq t_n]) = P([x \leq a]) - \underbrace{\lim_{n \rightarrow \infty} P([X \leq t_n])}_{P([X < a])} \\ &= P([X \leq a]) - P([X < a]) = P([X = a]) \end{aligned}$$

□

Por otro lado:

$$\begin{aligned} C_1 &= [x \leq t_1] \\ C_2 &= [x \leq t_2], C_1 \subseteq C_2 \\ &\vdots \\ &\vdots \\ C_n &= [x \leq t_n], C_{n-1} \subseteq C_n \end{aligned}$$

$$\lim_{n \rightarrow \infty} [x \leq t_n] = \bigcup_{i=1}^{\infty} [x \leq t_i] = [x < a]$$

Luego

$$\begin{aligned} B_1 &= [x \leq t_1] \\ B_2 &= [x \leq t_2] \setminus [x \leq t_1] \\ &\vdots \\ &\vdots \\ B_n &= [x \leq t_n] \setminus [x \leq t_{n-1}] \end{aligned}$$

Notese que $B_i \cap B_j = \emptyset$ para $i \neq j$ y $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n [x \leq t_{i-1}]$ entonces

$$\begin{aligned} P(\bigcup_{i=1}^n B_i) &= \sum_{i=1}^n P(B_i) \\ \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) &= \lim_{n \rightarrow \infty} \sum_{i=2}^n P([x \leq t_i] \setminus [x \leq t_{i-1}]) + P([x \leq t_1]) \\ &= \sum_{i=2}^n [P([x \leq t_i]) - P([x \leq t_{i-1}])] + P([x \leq t_1]) \\ &= \lim_{n \rightarrow \infty} P([x \leq t_n]) \end{aligned}$$

Teorema 1.2.

1. Si Y_1 y Y_2 son variables aleatorias discretas con función de probabilidad conjunta $P[y_1, y_2]$ y funciones marginales $P_1(y_1)$ y $P_2(y_2)$, entonces Y_1 y Y_2 son independientes ssi $P(y_1, y_2) = P_1(y_1)P_2(y_2), \forall y_1, y_2 \in \mathbb{R}$
2. Si Y_1 y Y_2 son variables aleatorias continuas con función densidad conjunta $f(y_1, y_2)$ y funciones densidades marginales $P_1(y_1)$ y $P_2(y_2)$, entonces Y_1 y Y_2 son independientes ssi $f(y_1, y_2) = f(y_1)f(y_2), \forall y_1, y_2 \in \mathbb{R}$

martes 18 de julio

Ej.

Sea X y Y variables aleatorias continuas con función de densidad

$$f(x, y) = \begin{cases} 6e^{-2x}e^{-3y}, & 0 < x, y < \infty \\ 0, & \text{En cualquier otro caso} \end{cases}$$

¿ X y Y son variables aleatorias independientes?

$$I(x, y) = \begin{cases} 1, & 0 < x, y < \infty \\ 0, & \text{en cualquier otro caso} \end{cases}$$

$$I(x) = \begin{cases} 1, & 0 < x < \infty \\ 0, & \text{en cualquier otro caso} \end{cases}$$

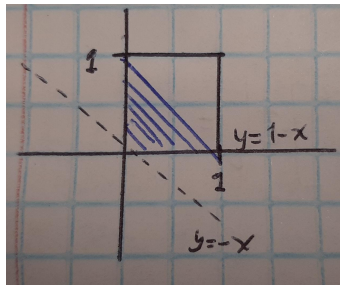
$$\begin{aligned} f(x, y) &= 6e^{-2x}e^{-3y} \cdot I(x, y) = (2e^{-2x} \cdot I(x)) \cdot (3e^{-3y} \cdot I(y)) = \\ &= f_x(x)f_y(y), \quad \forall x, y \in \mathbb{R} \end{aligned}$$

Por el teorema 5.8, X y Y son variables aleatorias independientes.

Ej.

Sea X y Y variables aleatorias con función de densidad

$$f(x, y) = \begin{cases} 24xy, & 0 < x, y < 1; 0 < x + y < 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$



Claramente no se puede factorizar en 2 partes que de x y otro que depende de y . Por lo tanto, X y Y son variables aleatorias dependientes.

5.9 Teorema

Sean Y_1 y Y_2 variables aleatorias independientes y sean $g(Y_1)$ y $h(Y_2)$ funciones solo de Y_1 y Y_2 . Entonces $E(g(Y_1) \cdot h(Y_2)) = E(g(Y_1)) \cdot E(h(Y_2))$ siempre que existan los valores esperados.

Ej.

En el ej. 5.19

Y_2 : Proporción de impurezas en la muestra

Y_1 : Proporción de impurezas de tipo I entre todas las impurezas encontradas

$$f(y_1, y_2) = \begin{cases} 2(1 - y_1), & 0 < y_1, y_2 < 1 \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Encontrar $E(y_1, y_2) = E(y_1) \cdot E(y_2)$ si Y_1 y Y_2 son independientes.

$$\begin{aligned}
 E(y_1, y_2) &= \int_0^1 \int_0^1 \underbrace{y_1 \cdot y_2}_{\text{funciones}} \cdot \underbrace{2(1-y_1)}_{\text{densidad conjunta}} dy_2 dy_1 \\
 &= 2 \int_0^1 y_1(1-y_1) \left. \frac{y_2^2}{2} \right|_0^1 dy_1 \\
 &= 2 \int_0^1 y_1(1-y_1) \left(\frac{1}{2} \right) dy_1 \\
 &= 2 \int_0^1 \frac{y_1}{2} - \frac{y_1^2}{2} dy_1 = \left[\frac{y_1^2}{4} - \frac{y_1^3}{6} \right]_0^1 = \\
 &= 2 \left(\frac{1}{4} - \frac{1}{6} \right) = \frac{1}{6}
 \end{aligned}$$

Marginales

$$\begin{aligned}
 f(y_1) &= \begin{cases} \int_0^1 2(1-y_1) dy_2 = 2(1-y_1), & 0 \leq y_1 \leq 1 \\ 0, & \text{en cualquier otro caso} \end{cases} \\
 f(y_2) &= \begin{cases} \int_0^1 2(1-y_1) dy_1 = 2y_2 - y_2^2 \Big|_0^1 = 1, & 0 \leq y_2 \leq 1 \\ 0, & \text{en cualquier otro caso} \end{cases}
 \end{aligned}$$

Valores Esperados

$$\begin{aligned}
 E(Y_1) &= \int_0^1 y_1(2)(1-y_1) dy_1 = \int_0^1 2y_1 - 2y_1^2 dy_1 = y_1^2 - \frac{2}{3}y_1^3 \Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3} \\
 E(Y_2) &= \int_0^1 y_2(1) dy_2 = \left. \frac{y_2^2}{2} \right|_0^1 = \frac{1}{2}
 \end{aligned}$$

Valores esperados y varianzas de funciones lineales de variables aleatorias

5.8) Teorema: Sean Y_1, \dots, Y_n y X_1, \dots, X_n variables aleatorias con $E(Y_i) = \mu_i$ y $E(X_j) = \xi_j$. Se define $U_1 = \sum_{i=1}^n a_i Y_i$ y $U_2 = \sum_{j=1}^m b_j X_j$ para $a_1, \dots, a_n, b_1, \dots, b_m \in \mathbb{R}$. Entonces,

1. $E(U) = \sum_{i=1}^n a_i E(Y_i)$
2. $Var(U_1) = \sum_{i=1}^n a_i^2 Var(Y_i) + 2 \sum_{1 \leq i < j \leq n} a_i a_j Cov(Y_i, Y_j)$
 - $Cov(X, Y) = E(xy) - (E(x))(E(y))$
3. a) $Cov(U_1, U_2) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Cov(Y_i, X_j)$
 b) $Cov(X, Y) = (E(x - E(x)))(E(y - E(y)))$

Demostración.

1.

$$\begin{aligned} E(U_1) &= E(\sum_i^n a_i Y_i) = \sum_i^n E(a_i Y_i) = (\text{Por teorema 5.8}) = \\ &\sum_i^n a_i E(Y_i) \end{aligned}$$

□