

Introducción a la Probabilidad y la Estadística



Martes y Jueves Aula B17
Dra Ana Georgina Flesia

Distribución Conjunta: Definición

Densidad

Recordemos que una variable aleatoria se dice continua con densidad o absolutamente continua cuando existe una función $f_X : \mathbb{R} \rightarrow \mathbb{R}$ no negativa que nos permite calcular la probabilidad de que X este en A de la siguiente forma

$$P(X \in A) = \int_A f_X(x) dx$$

Distribución Conjunta: Definición

Densidad conjunta

Sea (X_1, \dots, X_n) un vector aleatorio continuo, es decir, una n -upla donde cada coordenada es una variable aleatoria continua con densidad sobre el espacio de probabilidad (Ω, \mathcal{A}, P) , se dice que la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ no negativa es una función densidad conjunta del vector (X_1, \dots, X_n) si ocurre que

$$P[(X_1, \dots, X_n) \in A] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) I_A(x_1, \dots, x_n) dx_1 \cdots dx_n$$

para todo $A \subset \mathbb{R}^n$.

Distribución Conjunta: Definición

Distribución acumulada conjunta

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

$$\begin{aligned} F(x_1, \dots, x_n) &= P[(X_1 \leq x_1, \dots, X_n \leq x_n)] \\ &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(x_1, \dots, x_n) dx_1 \cdots dx_n \end{aligned}$$

Distribución Conjunta: Propiedades

Propiedades

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

1. f no es única, pues cualquier función que no coincida con f en un número finito o numerable de puntos también es una densidad conjunta para el vector aleatorio
2. $f \geq 0$, es decir, es no negativa
3. $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \cdots dx_n = 1$
4. Si F es derivable entonces $f = \frac{\partial^n F}{\partial x_1 \cdots \partial x_n}$

Marginales

Definición

Sean (X_1, \dots, X_n) un vector aleatorio continuo con función densidad conjunta f , entonces las densidades marginales de las variables componentes del vector son

$$f_{X_j}(x_j) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_1 \cdots, dx_{j-1}, dx_{j+1} \cdots dx_n$$

esto es, integro sobre todas las coordenadas salvo la coordenada x_j .

Independencia

Definición

Sean (X_1, \dots, X_n) un vector aleatorio continuo con función densidad conjunta f , decimos que las componentes del vector son independientes si

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$$

Ejemplo

Considere la siguiente función:

$$f(x, y) = \begin{cases} k(x + y^3) & \forall x, y \in [0, 1] \\ 0 & cc \end{cases}$$

1. Determine el valor de la constante k para que f sea la función de densidad conjunta de un vector (X, Y)
2. Encuentre las densidades marginales f_X y f_Y
3. Diga si las variables son independientes o no.
4. Encuentre la esperanza y varianza de X e Y

Resolución

Para calcular la integral doble $\int_0^1 \int_0^1 (x + y^3) dx dy$, vamos a dividir la integral en dos partes:

$$\int_0^1 \int_0^1 (x + y^3) dx dy = \int_0^1 \int_0^1 x dx dy + \int_0^1 \int_0^1 y^3 dx dy$$

Calculando la primera parte

$$\int_0^1 \int_0^1 x \, dx \, dy$$

Primero, calculamos la integral con respecto a x :

$$\int_0^1 x \, dx = \left[\frac{x^2}{2} \right]_0^1 = \frac{1^2}{2} - 0 = \frac{1}{2}$$

Ahora, la integral con respecto a y :

$$\int_0^1 \frac{1}{2} \, dy = \frac{1}{2} \cdot [y]_0^1 = \frac{1}{2} \cdot (1 - 0) = \frac{1}{2}$$

Calculando la segunda parte

$$\int_0^1 \int_0^1 y^3 dx dy$$

La integral con respecto a x es:

$$\int_0^1 y^3 dx = y^3 \cdot [x]_0^1 = y^3 \cdot (1 - 0) = y^3$$

Ahora integremos con respecto a y :

$$\int_0^1 y^3 dy = \left[\frac{y^4}{4} \right]_0^1 = \frac{1^4}{4} - 0 = \frac{1}{4}$$

Sumando las partes

Ahora sumamos los resultados de ambas partes

$$\int_0^1 \int_0^1 (x + y^3) dx dy = \frac{1}{2} + \frac{1}{4} = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}$$

Por lo tanto, el valor de la integral es $\frac{3}{4}$: y la densidad conjunta resulta

$$f(x, y) = \begin{cases} \frac{4}{3}(x + y^2) & \forall x, y \in [0, 1] \\ 0 & cc \end{cases}$$

Marginales f_X y f_Y

Integrando la densidad $f(x, y)$ en la variable y

$$f_X(x) = \frac{4}{3} \int_0^1 (x+y^3) dy = \frac{4}{3} \left[\int_0^1 x dy + \int_0^1 y^3 dy \right] = \frac{4}{3} \left[x + \frac{1}{4} \right] = \frac{4}{3}x + \frac{1}{3}$$

Integrando la densidad $f(x, y)$ en la variable x

$$f_Y(y) = \frac{4}{3} \int_0^1 (x+y^3) dx = \frac{4}{3} \left[\int_0^1 x dx + \int_0^1 y^3 dx \right] = \frac{4}{3} \left[\frac{1}{2} + y^3 \right] = \frac{4}{3}y^3 + \frac{2}{3}$$

Independencia

$$f(x, y) = \begin{cases} \frac{4}{3}(x + y^3) & \forall x, y \in [0, 1] \\ 0 & \text{cc} \end{cases}$$

$$f_X(x) = \begin{cases} \frac{4}{3}x + \frac{1}{3} & \forall x \in [0, 1] \\ 0 & \text{cc} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{4}{3}y^3 + \frac{2}{3} & \forall y \in [0, 1] \\ 0 & \text{cc} \end{cases}$$

$$\left(\frac{4}{3}x + \frac{1}{3}\right) \times \left(\frac{4}{3}y^3 + \frac{2}{3}\right) \neq \frac{4}{3}(x + y^3)$$

por lo cual X e Y no son independientes

$E(X)$ y $E(Y)$

$$\begin{aligned} E(X) &= \int_0^1 x f_X(x) dx = \int_0^1 \frac{4}{3} x^2 dx + \int_0^1 \frac{1}{3} x dx \\ &= \frac{4}{3} \left[\frac{x^3}{3} \right]_0^1 + \frac{1}{3} \left[\frac{x^2}{2} \right]_0^1 \\ &= \frac{4}{3} \frac{1}{3} + \frac{1}{3} \frac{1}{2} = \frac{11}{18} \end{aligned}$$

$$\begin{aligned} E(Y) &= \int_0^1 y f_Y(y) dy = \int_0^1 \frac{4}{3} y^4 dy + \int_0^1 \frac{2}{3} y dy \\ &= \frac{4}{3} \left[\frac{y^5}{5} \right]_0^1 + \frac{2}{3} \left[\frac{y^2}{2} \right]_0^1 \\ &= \frac{4}{3} \frac{1}{5} + \frac{2}{3} \frac{1}{2} = \frac{3}{5} \end{aligned}$$

Ejemplo II

Considere la siguiente función:

$$f(x, y) = \begin{cases} \frac{12}{7}(x^2 + xy) & \forall x, y \in [0, 1] \\ 0 & \text{cc} \end{cases}$$

1. Encuentre las densidades marginales f_X y f_Y
2. Calcule $P(X > Y)$

Marginales f_X y f_Y

Integrando la densidad $f(x, y)$ en la variable y

$$f_X(x) = \frac{12}{7} \int_0^1 (x^2 + xy) dy = \frac{12}{7} \left(x^2 + \frac{x}{2} \right)$$


Integrando la densidad $f(x, y)$ en la variable x


$$f_Y(y) = \frac{12}{7} \int_0^1 (x^2 + xy) dx = \frac{12}{7} \left(\frac{1}{3} + \frac{y}{2} \right)$$


$$P(X > Y)$$

Integrando la densidad $f(x, y)$ sobre el conjunto
 $\{(x, y) | 0 \leq y \leq x \leq 1\}$

$$P(X > Y) = \frac{12}{7} \int_0^1 \int_0^x (x^2 + xy) dy dx = \frac{9}{14}$$


$$\int_0^x (x^2 + xy) dy = x^2 [y]_0^x + x \left[\frac{y^2}{2} \right]_0^x = x^3 + \frac{x^3}{2} = \frac{3x^3}{2}$$


$$\int_0^1 \frac{3x^3}{2} dx = \frac{3}{2} \int_0^1 x^3 dx = \frac{3}{2} \left[\frac{x^4}{4} \right]_0^1 = \frac{3}{2} \frac{1}{4} = \frac{3}{8}$$


$$\frac{12}{7} \cdot \frac{3}{8} = \frac{36}{56} = \frac{9}{14}$$

Esperanza de funciones de vectores continuos

Proposición

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

$$E(h(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f_{X,Y}(x, y) dx dy$$

Esperanza de funciones de vectores continuos

Ejemplo

Sea (X, Y) un vector aleatorio con densidad conjunta

$$f_{XY}(x, y) = e^{-y} \quad 0 \leq x \leq y$$

Hallar la esperanza de $Z = XY$.

Resolución

Aplicando el teorema anterior

$$\begin{aligned}E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x, y) dx dy \\&= \int_0^{\infty} y e^{-y} \left[\int_0^y x dx \right] dy = \int_0^{\infty} y e^{-y} \left[\frac{y^2}{2} \right] dy \\&= \int_0^{\infty} \frac{y^3}{2} e^{-y} dy = \frac{\Gamma(4)}{2} \int_0^{\infty} \frac{1^4}{\Gamma(4)} y^{4-1} e^{-y} dy \\&= \frac{\Gamma(4)}{2} = \frac{(4-1)!}{2} \\&= 6/2 = 3\end{aligned}$$

usando el hecho de que

$$1 = \int_0^{\infty} \frac{\lambda^{\alpha}}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$$

para todo α y λ positivos.

Esperanza de funciones de vectores continuos

Propiedades

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

1. $E(aX + bY) = aE(X) + bE(Y)$
2. Si X e Y son independientes entonces $E(XY) = E(X)E(Y)$

Covarianza

Definición

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

$$Cov(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (X - E(X))(y - E(Y))f_{X,Y}(x, y)dx dy$$

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. $Cov(aX + b, cY + d) = acCov(X, Y) \quad \forall a, b, c, d \in \mathbb{R}$
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
 $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$

Correlación

Definición

Sean X e Y variables aleatorias continuas 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

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

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 $a < 0$.

Práctica Esencial

1. Sea A una variable aleatoria Exponencial($\lambda = 1.5$) y sea Θ una variable aleatoria Uniforme($a = -\pi, b = \pi$). ¿Cuál es $\text{Cov}[A \cos(\Theta + 2\pi s), A \cos(\Theta + 2\pi t)]$?
2. En un sistema en espera, un componente se usa hasta que se desgasta y luego es reemplazado inmediatamente por otro, no necesariamente idéntico, componente. (El segundo componente se dice que está "en modo de espera", es decir, esperando ser usado). El tiempo de vida total de un sistema en espera es simplemente la suma de los tiempos de vida de sus componentes individuales. Sea X y Y los tiempos de vida de los dos componentes de un sistema en espera, y suponga que X y Y son variables aleatorias distribuidas exponencialmente de forma independiente con tiempos de vida esperados de 3 semanas y 4 semanas, respectivamente. Sea $T = X + Y$, el tiempo de vida del sistema en espera. ¿Cuál es la desviación estándar del tiempo de vida del sistema?
3. Sea U_1, U_2, \dots, U_n variables aleatorias independientes e idénticamente distribuidas (i.i.d.) Uniforme($a = 0, b = 1$) aleatorias. Sea $S_n = U_1 + \dots + U_n$ denotar su suma. Calcule $E[S_n]$ y $SD[S_n]$ en términos de n .

1. Calcular $\text{Cov}[A \cos(\Theta + 2\pi s), A \cos(\Theta + 2\pi t)]$

- A es una variable aleatoria exponencial con $\lambda = 1.5$, por lo que $E[A] = 1/\lambda = 1/1.5 = 2/3$ y $\text{Var}(A) = 1/\lambda^2 = 1/(1.5)^2 = 4/9$.
- Θ es una variable aleatoria uniforme en $(-\pi, \pi)$, por lo que $E[\cos(\Theta)] = 0$ (ya que $\cos(\Theta)$ es simétrico alrededor de 0) y $\text{Var}(\cos(\Theta))$ puede derivarse, pero aquí usamos la covarianza.

- La covarianza se calcula como:

$$\text{Cov}[A \cos(\Theta + 2\pi s), A \cos(\Theta + 2\pi t)] = E[(A \cos(\Theta + 2\pi s))(A \cos(\Theta + 2\pi t))] - E[A \cos(\Theta + 2\pi s)]E[A \cos(\Theta + 2\pi t)]$$

- Dado que A y Θ son independientes, $E[A \cos(\Theta + 2\pi s)] = E[A]E[\cos(\Theta + 2\pi s)]$. Como $\Theta + 2\pi s$ y $\Theta + 2\pi t$ son desplazamientos de Θ , y \cos es periódico con período 2π , $E[\cos(\Theta + 2\pi s)] = E[\cos(\Theta)] = 0$, así que el segundo término es 0.

- El primer término es:

$$E[A^2 \cos(\Theta + 2\pi s) \cos(\Theta + 2\pi t)]$$

Dado que A y Θ son independientes, esto se descompone en $E[A^2]E[\cos(\Theta + 2\pi s) \cos(\Theta + 2\pi t)]$.

- $E[A^2] = \text{Var}(A) + (E[A])^2 = 4/9 + (2/3)^2 = 4/9 + 4/9 = 8/9$.
- Usamos la identidad trigonométrica: $\cos(a) \cos(b) = \frac{1}{2}[\cos(a - b) + \cos(a + b)]$. Entonces:

$$E[\cos(\Theta + 2\pi s) \cos(\Theta + 2\pi t)] = E\left[\frac{1}{2}[\cos(2\pi(s - t)) + \cos(2\pi(s + t) + 2\Theta)]\right]$$

Como Θ es uniforme en $(-\pi, \pi)$, $E[\cos(2\pi(s + t) + 2\Theta)] = 0$ (simetría), y $E[\cos(2\pi(s - t))]$ depende de si $s = t$ o no. Si $s = t$, $\cos(0) = 1$; si $s \neq t$, el promedio sobre un período completo es 0.

- Por lo tanto, la covarianza es:
 - Si $s = t$: $E[A^2] \cdot \frac{1}{2} = (8/9) \cdot (1/2) = 4/9$.
 - Si $s \neq t$: 0.

Sea X y Y los tiempos de vida de los dos componentes de un sistema en espera, y suponga que X y Y son variables aleatorias distribuidas exponencialmente de forma independiente con tiempos de vida esperados de 3 semanas y 4 semanas, respectivamente. Sea $T = X + Y$, el tiempo de vida del sistema en espera. ¿Cuál es la desviación estándar del tiempo de vida del sistema?

2. Desviación estándar del tiempo de vida del sistema $T = X + Y$

- X y Y son exponenciales independientes con $E[X] = 3$ semanas y $E[Y] = 4$ semanas.
- Para una variable exponencial, $\text{Var}(X) = (E[X])^2 = 9$ y $\text{Var}(Y) = 16$.
- Como X y Y son independientes, $\text{Var}(T) = \text{Var}(X) + \text{Var}(Y) = 9 + 16 = 25$.
- La desviación estándar es $SD(T) = \sqrt{\text{Var}(T)} = \sqrt{25} = 5$ semanas.
- Respuesta: La desviación estándar del tiempo de vida del sistema es 5 semanas.

3. Sea U_1, U_2, \dots, U_n variables aleatorias independientes e idénticamente distribuidas (i.i.d.) Uniforme($a = 0, b = 1$) aleatorias. Sea $S_n = U_1 + \dots + U_n$ denotar su suma. Calcule $E[S_n]$ y $SD[S_n]$ en términos de n .

3. Calcular $E[S_n]$ y $SD[S_n]$ en términos de n

- U_1, U_2, \dots, U_n son i.i.d. uniformes en $(0, 1)$, por lo que $E[U_i] = \frac{0+1}{2} = 0.5$ y $\text{Var}(U_i) = \frac{(1-0)^2}{12} = \frac{1}{12}$.
- $S_n = U_1 + \dots + U_n$, así que $E[S_n] = n \cdot E[U_i] = n \cdot 0.5 = \frac{n}{2}$.
- Como las variables son independientes, $\text{Var}(S_n) = n \cdot \text{Var}(U_i) = n \cdot \frac{1}{12} = \frac{n}{12}$.
- $SD[S_n] = \sqrt{\text{Var}(S_n)} = \sqrt{\frac{n}{12}} = \frac{\sqrt{n}}{\sqrt{12}} = \frac{\sqrt{n}}{2\sqrt{3}}$.
- Respuesta: $E[S_n] = \frac{n}{2}$, $SD[S_n] = \frac{\sqrt{n}}{2\sqrt{3}}$.