Probabilidad y estadística

MAT 041, Primer semestre

Francisco Cuevas Pacheco

11 de octubre de 2022

Contenidos

✓ Independencia

✓ Covarianza y Correlación

Definición (Independencia)

Sea $(X, Y)^\top$ *un vector aleatorio con función de probabilidad conjunta* $f_{XY}(x, y)$*. Diremos que* X *e* Y *son independientes si y sólo si para cada* $(x, y) \in rec(X) \times rec(Y)$*,*

$$f_{XY}(x, y) = f_X(x)\, f_Y(y).$$

Propiedad

Si las variables aleatorias X e Y son independientes, entonces

$$\boxed{E(XY) = E(X)E(Y)}$$

Ejercicio

Sean X e Y variables aleatorias con función de densidad conjunta

$$f(x,y) = \begin{cases} \frac{9}{4}(x^2y^2 - x^2 - y^2 + 1) & \textit{si } 0 \leq x \leq 1 \quad 0 \leq y \leq 1 \\ 0 & \textit{en otro caso} \end{cases}$$

verifique que X e Y son independientes.

Ejercicio

Sean X e Y variables aleatorias independientes, cada una con función de densidad (marginal)

$$g(x) = \begin{cases} 2x & si\ 0 \leq x \leq 1 \\ 0 & en\ otro\ caso \end{cases}$$

Calcule $P(X + Y \leq 1)$.

Ejercicio

Sean X e Y variables aleatorias independientes, cada una con función de densidad (marginal)

$$g(x) = \begin{cases} 2x & \textit{si } 0 \leq x \leq 1 \\ 0 & \textit{en otro caso} \end{cases}$$

Calcule $P(X + Y \leq 1)$.

Solución: La función de densidad conjunta es

$$f_{XY}(x, y) = g(x)g(y) = \begin{cases} 4xy & \text{si } 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Luego,

$$P(X + Y \leq 1) = \int_0^1 \int_0^{1-x} 4xy \, dy dx = \frac{1}{6}$$

La covarianza mide el grado de asociación lineal entre dos variables.

Definición (Covarianza)

Sean X e Y variables aleatorias con varianza finita. La covarianza entre X e Y es

$$cov(X, Y) = \sigma_{XY} = E\left([X - E(X)][Y - E(Y)]\right).$$

La covarianza mide el grado de asociación lineal entre dos variables.

Definición (Covarianza)

Sean X e Y variables aleatorias con varianza finita. La covarianza entre X e Y es

$$cov(X, Y) = \sigma_{XY} = E\left([X - E(X)][Y - E(Y)]\right).$$

La correlación es una versión normalizada de la covarianza.

Definición (Correlación)

La correlación entre X y Y se define como

$$cor(X, Y) = \rho_{XY} = \frac{cov(X, Y)}{\sqrt{var(X) var(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Observación

La desigualdad de Cauchy-Schwarz implica que

$$-1 \leq \text{cor}(X, Y) \leq 1.$$

- Cuando $\text{cor}(X, Y)$ es cercana a 1, se tiene asociación lineal directa entre las variables.
- Cuando $\text{cor}(X, Y)$ cercana a -1 se tiene asociación lineal inversa entre las variables.
- Cuando $\text{cor}(X, Y)$ es cercana a 0, no existe evidencia de asociación lineal entre las variables.

Propiedades

- Si X e Y son independientes, entonces $\text{cov}(X, Y) = 0$. La implicancia opuesta no es cierta en general.
- Simetría: $\text{cov}(X, Y) = \text{cov}(Y, X)$.
- Bilinealidad: considere las variables aleatorias $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ y las constantes $a_1, \ldots, a_n, b_1, \ldots, b_m$, entonces

$$\text{cov}\left(\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{m} b_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \text{cov}\left(X_i, Y_j\right).$$

Fórmula alternativa para la covarianza

Es posible desarrollar la expresión de la covarianza:

$$\begin{aligned}
\operatorname{cov}(X, Y) &= E([X - E(X)][Y - E(Y)]) \\
&= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) \\
&= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\
&= E(XY) - E(X)E(Y)
\end{aligned}$$

En particular, observe que $\operatorname{cov}(X, X) = \operatorname{var}(X)$.

Propiedad

Si X e Y son independientes $\implies \rho_{XY} = 0$

Demostración

Directo ya que Si X e Y son independientes entonces se cumple que $E(XY) = E(X)E(Y)$.

Definición (Matriz de Covarianza)

La matriz de covarianza del vector aleatorio $(X, Y)^\top$ *es*

$$\begin{bmatrix} var(X) & cov(X, Y) \\ cov(X, Y) & var(Y) \end{bmatrix}$$

Esta matriz es simétrica y definida positiva.

Propiedad Adicional

Sean $X_1, \ldots, X_n$ variables aleatorias y $a_1, \ldots, a_n$ constantes, entonces

$$\begin{aligned}
\operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{cov}(X_i, X_j) \\
&= \sum_{i=1}^{n} a_i^2 \operatorname{var}(X_i) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} a_i a_j \operatorname{cov}(X_i, X_j).
\end{aligned}$$

En particular,

$$\operatorname{var}(X \pm Y) = \operatorname{var}(X) + \operatorname{var}(Y) \pm 2\operatorname{cov}(X, Y).$$

En otras palabras, la varianza de una suma no es la suma de las varianzas, salvo casos especiales (por ejemplo, cuando las variables aleatorias son independientes).

Ejercicio

Sea $(X, Y)^\top$ *un vector aleatorio con función de densidad conjunta*

$$f_{XY}(x,y) = \begin{cases} \dfrac{3x^2}{2} + y & \text{si } 0 < x < 1,\ 0 < y < 1 \\ \\ 0 & \text{en otro caso} \end{cases}$$

Muestre que la matriz de covarianza del vector $(X, Y)^\top$ *está dada por*

$$\begin{bmatrix} \frac{73}{960} & -\frac{1}{96} \\ \\ -\frac{1}{96} & \frac{11}{144} \end{bmatrix}$$

Solución: Procedemos por definición:

$$E(XY) = \int_0^1 \int_0^1 xy\left(\frac{3x^2}{2} + y\right) \mathrm{d}x\mathrm{d}y = \frac{34}{96}$$

$$E(X) = \int_0^1 \int_0^1 x\left(\frac{3x^2}{2} + y\right) \mathrm{d}x\mathrm{d}y = \frac{5}{8}$$

$$E(Y) = \int_0^1 \int_0^1 y\left(\frac{3x^2}{2} + y\right) \mathrm{d}x\mathrm{d}y = \frac{7}{12}$$

$$E(Y^2) = \int_0^1 \int_0^1 y^2\left(\frac{3x^2}{2} + y\right) \mathrm{d}x\mathrm{d}y = \frac{5}{12}$$

$$E(X^2) = \int_0^1 \int_0^1 x^2\left(\frac{3x^2}{2} + y\right) \mathrm{d}x\mathrm{d}y = \frac{7}{15}$$

Luego

$$
\begin{aligned}
Var(X) &= E(X^2) - E(X)^2 = 7/15 - (5/8)^2 = 73/960 \\
Var(Y) &= E(Y^2) - E(Y)^2 = 5/12 - (7/12)^2 = 11/144 \\
Cov(X, Y) &= E(XY) - E(X)E(Y) = 34/96 - (5 \times 7)/(8 \times 12) = -1/96.
\end{aligned}
$$

Ejercicio

1. Calcule $E(X/Y = 1/4)$ y $Var(X/Y = 1/4)$
2. Calcule $E(Y/X = 1/2)$ y $Var(Y/X = 1/2)$

Solución: Calculamos la distribucion condicional:

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

Entonces: $f_Y(y) = \int_0^1 f_{XY}(x,y)\mathrm{d}x = \frac{1}{2} + y$

Luego

$$f_{X/Y=y}(x) = \frac{\frac{3x^2}{2} + y}{\frac{1}{2} + y}$$

En consecuencia $f_{X/Y=1/4}(x) = 2x^2 + 1/2$

Luego

$$E(X/Y = 1/4) = \int_0^1 x f_{X/Y=1/4}(x)\mathrm{dx} = \frac{5}{8}$$
$$E(X^2/Y = 1/4) = \int_0^1 x^2 f_{X/Y=1/}(x)\mathrm{dx} = \frac{29}{60}$$
$$V(X^2/Y = 1/4) = \frac{29}{60} - \frac{25}{64} = \frac{89}{960}$$