Distribución de Probabilidad Conjunta

Jhon J. Padilla A., PhD.

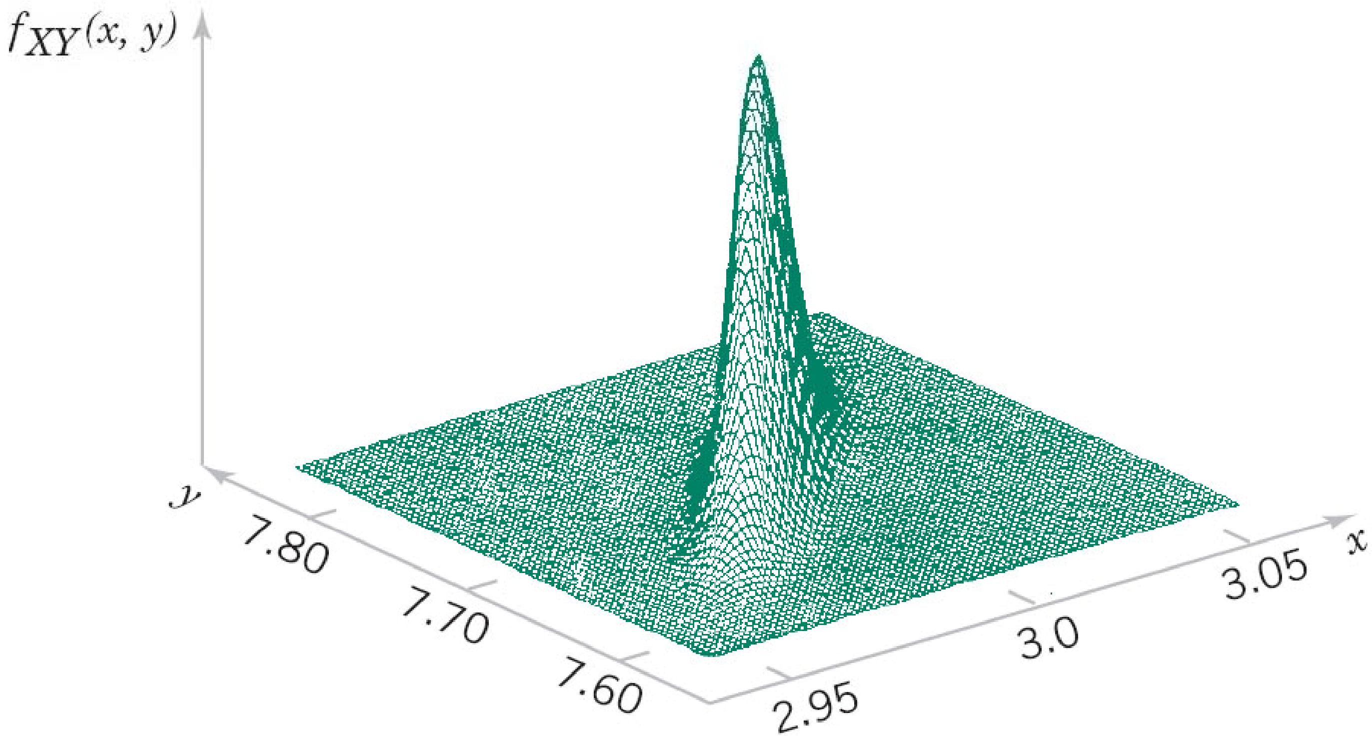
Introducción

- Los resultados de un experimento pueden ser causa de múltiples variables.
- En estas situaciones se requiere de tener una f.d.p que describa la variación de la probabilidad de ocurrencia con respecto a la variación de estas variables.
- Esta f.d.p que tiene en cuenta el efecto de múltiples variables aleatorias se denomina **distribución de probabilidad conjunta**.
- Una distribución de probabilidad conjunta puede ser **discreta o continua** dependiendo del tipo de v.a.'s que describe.

Definición

- La función de masa de probabilidad conjunta de dos variables aleatorias contínuas X y Y se denota como $f_{XY}(x,y)$, y satisface las siguientes propiedades
- 1) $f_{XY}(x,y) \geq 0$ para toda x,y
- 2) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1$$
- 3) Para cualquier región R del espacio bidimensional: $$P([X,Y \in R]) = \iint_R f_{XY}(x,y)\,dx\,dy$$

Distribución de probabilidad bidimensional



Variables aleatorias Múltiples

- La función de distribución de probabilidad conjunta para las variables X_1, X_2,…,X_p, se denota como:

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots x_p)$$

Definición

- Una función de distribución de probabilidad conjunta para las variables aleatorias continuas X_1, X_2,…,X_p, denotada como $f_{X_1 X_2 \dots X_p}(x_1, x_2, \dots x_p)$, satisface las siguientes propiedades
- 1) $f_{X_1 X_2 \dots X_p}(x_1, x_2, \dots x_p) \geq 0$
- 2) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{X_1 X_2 \dots X_p}(x_1, x_2, \dots x_p) dx_1 dx_2 \dots dx_p = 1$
- 3) Para cualquier región B del espacio de p dimensiones,

$$P([X_1, X_2, \dots, X_p] \in B) = \int \int_B \dots \int f_{X_1 X_2 \dots X_p}(x_1, x_2, \dots x_p) dx_1 dx_2 \dots dx_p$$

Independencia

- Las variables aleatorias contínuas X_1, X_2,…,X_p, son independientes si y sólo si

$$f_{X_1 X_2 \ldots X_p}(x_1, x_2, \ldots x_p) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots. f_{X_p}(x_p)$$

- Para toda x_1, x_2,…,x_p

COVARIANZA

- Cuando la probabilidad depende de más de una variable aleatoria, resulta conveniente describir la relación entre las variables.
- La **covarianza** es una medida común de la relación entre dos variables aleatorias
- La covarianza es una medida de asociación lineal entre las v.a's.
- Si la relación entre las v.a's no es lineal, la covarianza podría no ser sensible a la relación.

Definición

- La covarianza entre dos v.a. X y Y, denotada como cov(X,Y) ó σ_{xy}, es

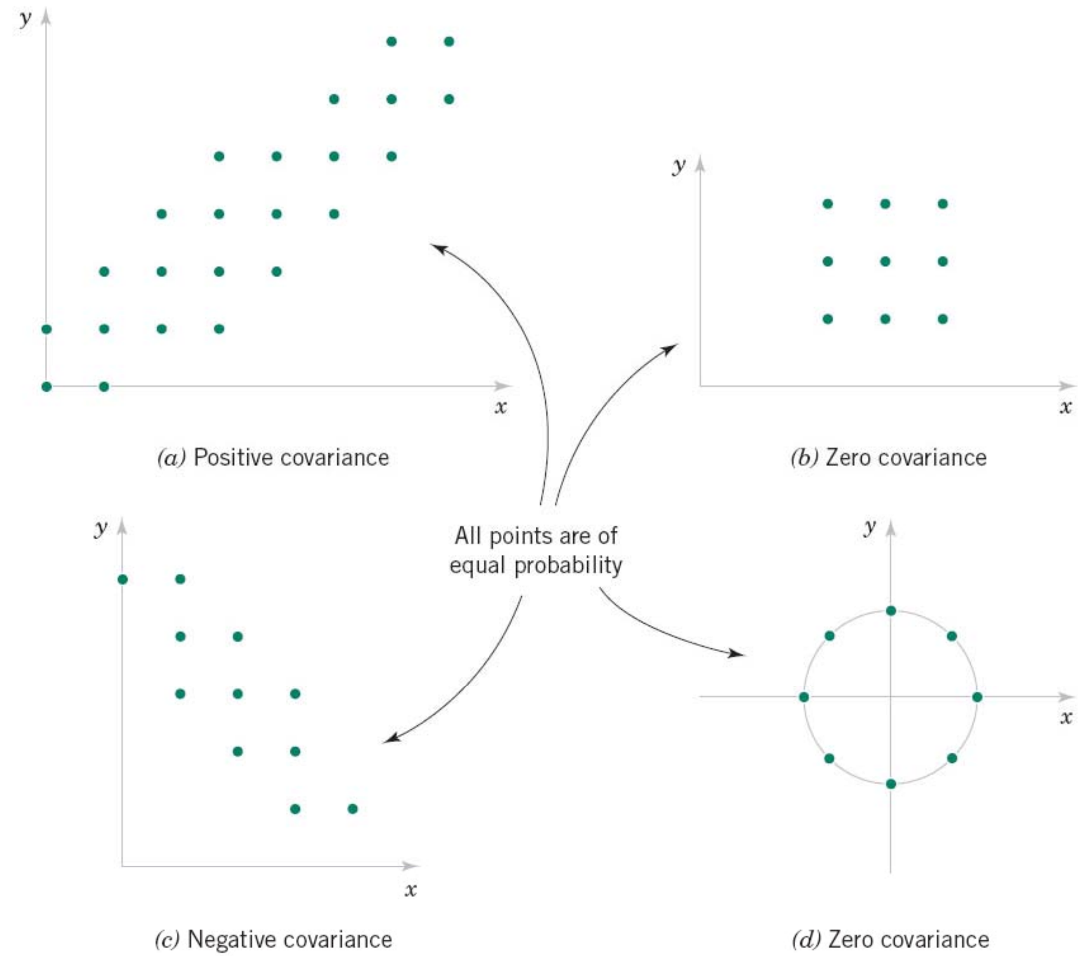
$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)(y-\mu_Y)\,f_{XY}(x,y)\,dxdy$$

- Y se puede demostrar que también

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dxdy - \mu_X\mu_Y$$

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - \mu_X\mu_Y$$

Distribuciones de probabilidad conjuntas y el signo de la covarianza entre X y Y



CORRELACION

- La correlación mide el grado de relación lineal entre dos variables.
- La correlación escala la covarianza por la desviación estándar de cada variable.
- Es una cantidad adimensional

Definición

- La correlación entre las variables aleatorias X y Y, denotada como ρ_{xy}, es

$$\rho_{XY} = \frac{\operatorname{cov}(X,Y)}{V(X)V(Y)} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

- Además,

$$-1 \leq \rho_{XY} \leq +1$$

Interpretación de la correlación

- Si los puntos de la distribución de probabilidad conjunta de X y Y tienden a caer en una recta con pendiente positiva (negativa), entonces la correlación estará cerca de +1 (o de -1).
- Si la correlación da +1 ó -1, entonces los puntos de la distribución de probabilidad conjunta caen exactamente en una línea recta.
- Se dice que dos v.a con correlación diferente de cero están correlacionadas.
- Finalmente, si X y Y son v.a's independientes, entonces

$$\sigma_{XY} = \rho_{XY} = 0$$