1.10.5 Covariance and Correlation

Covariance and correlation are two measures of the strength of a relationship between two r.vs.

We will use the following notation.
$\mathrm{E}(X_1) = \mu_{X_1}$
$\mathrm{E}(X_2) = \mu_{X_2}$
$\mathrm{var}(X_1) = \sigma^2_{X_1}$
$\mathrm{var}(X_2) = \sigma^2_{X_2}$

Also, we assume that $\sigma^2_{X_1}$ and $\sigma^2_{X_2}$ are finite positive values. A simplified notation $\mu_1,\ \mu_2,\ \sigma_1^2,\ \sigma_2^2$ will be used when it is clear which rvs we refer to.

Definition 1.19. *The covariance of X_1 and X_2 is defined by*

$$\mathrm{cov}(X_1, X_2) = \mathrm{E}[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})]. \tag{1.16}$$

□

Some useful properties of the covariance and correlation are given in the following two theorems.

Theorem 1.15. *Let X_1 and X_2 denote random variables and let $a, b, c, \dots$ denote some constants. Then, the following properties hold.*

1. $\mathrm{cov}(X_1, X_2) = \mathrm{E}(X_1X_2) - \mu_{X_1}\mu_{X_2}$.

2. *If random variables X_1 and X_2 are independent then*

$$\mathrm{cov}(X_1, X_2) = 0.$$

3. $\mathrm{var}(aX_1 + bX_2) = a^2\,\mathrm{var}(X_1) + b^2\,\mathrm{var}(X_2) + 2ab\,\mathrm{cov}(X_1, X_2)$.

4. *For $U = aX_1 + bX_2 + e$ and for $V = cX_1 + dX_2 + f$ we can write*

$$\mathrm{cov}(U, V) = ac\,\mathrm{var}(X_1) + bd\,\mathrm{var}(X_2) + (ad + bc)\,\mathrm{cov}(X_1, X_2).$$

Proof. We will show property 4.

$$\begin{aligned}
\operatorname{cov}(U,V) &= \mathrm{E}\left[\big((aX_1+bX_2+e)-\mathrm{E}(aX_1+bX_2+e)\big)\right.\\
&\qquad \left.\times\big((cX_1+dX_2+f)-\mathrm{E}(cX_1+dX_2+f)\big)\right]\\
&= \mathrm{E}\left[\big(a(X_1-\mathrm{E}\,X_1)+b(X_2-\mathrm{E}\,X_2)\big)\big(c(X_1-\mathrm{E}\,X_1)+d(X_2-\mathrm{E}\,X_2)\big)\right]\\
&= \mathrm{E}\left[ac(X_1-\mathrm{E}\,X_1)^2+bd(X_2-\mathrm{E}\,X_2)^2+(ad+bc)(X_1-\mathrm{E}\,X_1)(X_2-\mathrm{E}\,X_2)\right]\\
&= ac\operatorname{var}X_1+bd\operatorname{var}X_2+(ad+bc)\operatorname{cov}(X_1,X_2).
\end{aligned}$$

□

Property 2 says that if two variables are independent, then their covariance is zero. This does not always work both ways, that is it does not mean that if the covariance is zero then the variables must be independent. The following small example shows this fact.

Example 1.27. Let $X \sim U(-1,1)$ and let $Y = X^2$. Then

$$\begin{aligned}
&\mathrm{E}(X) = 0\\
&\mathrm{E}(Y) = \mathrm{E}(X^2) = \int_{-1}^{1} x^2\frac{1}{2}dx = \frac{1}{3}\\
&\mathrm{E}(XY) = \mathrm{E}(X^3) = 0.
\end{aligned}$$

Hence,

$$\operatorname{cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = 0,$$

but Y is a function of X, so these two variables are not independent.

□

Another measure of strength of relationship of two rvs is correlation. It is defined as

Definition 1.20. *The correlation of X_1 and X_2 is defined by*

$$\rho_{(X_1,X_2)} = \frac{\operatorname{cov}(X_1,X_2)}{\sigma_{X_1}\sigma_{X_2}}. \tag{1.17}$$

□

Correlation is a dimensionless measure and it expresses the strength of linearly related variables as shown in the following theorem.

Theorem 1.16. *For any random variables X_1 and X_2*

1. $-1 \leq \rho_{(X_1,X_2)} \leq 1$,
2. $|\rho_{(X_1,X_2)}| = 1$ *iff there exist numbers* $a \neq 0$ *and* b *such that*

$$P(X_2 = aX_1 + b) = 1.$$

If $\rho_{(X_1,X_2)} = 1$ *then* $a > 0$, *and if* $\rho_{(X_1,X_2)} = -1$ *then* $a < 0$. □

Exercise 1.17. Prove Theorem 1.16. Hint: Consider roots (and so the discriminant) of the quadratic function of t:

$$g(t) = \mathrm{E}[(X - \mu_X)t + (Y - \mu_Y)]^2.$$

Example 1.28. Let the joint pdf of X, Y be

$$f_{X,Y}(x, y) = 1 \quad \text{on the support } \{(x, y) : 0 < x < 1, x < y < x + 1\}.$$

The two rvs are not independent as the range of Y depends on x. We will calculate the correlation between X and Y. For this we need to obtain the marginal pdfs, $f_X(x)$ and $f_Y(y)$. The marginal pdf for X is

$$f_X(x) = \int_x^{x+1} 1 dy = y|_x^{x+1} = 1, \quad \text{on support } \{x : 0 < x < 1\}.$$

Note that to obtain the marginal pdf for the rv Y the range of X has to be considered separately for $y \in (0, 1)$ and for $y \in [1, 2)$. When $y \in (0, 1)$ then $x \in (0, y)$. When $y \in [1, 2)$, then $x \in (y - 1, 1)$. Hence,

$$f_Y(y) = \begin{cases} \int_0^y 1dx = x|_0^y = y, & \text{for } y \in (0,1); \\ \int_{y-1}^1 1dx = x|_{y-1}^1 = 2 - y, & \text{for } y \in [1,2); \\ 0, & \text{otherwise.} \end{cases}$$

These give

$$\mu_X = \frac{1}{2}, \quad \sigma_X^2 = \frac{(b-a)^2}{12} = \frac{1}{12};$$
$$\mu_Y = 1 \quad \sigma_Y^2 = \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2 = \frac{7}{6} - 1 = \frac{1}{6},$$

where
$$\mathrm{E}(Y^2) = \int_0^1 y^2 y dy + \int_1^2 y^2(2-y)dy = \frac{7}{6}.$$
Also,
$$\mathrm{E}(XY) = \int_0^1 \int_x^{x+1} xy1dydx = \frac{7}{12}.$$
Hence,
$$\mathrm{cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = \frac{7}{12} - \frac{1}{2} \times 1 = \frac{1}{12}.$$
Finally,
$$\rho_{(X,Y)} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{1}{12}}{\sqrt{\frac{1}{12}\frac{1}{6}}} = \frac{1}{\sqrt{2}}.$$
The linear relationship between X and Y is not very strong.

Note: We can make an interesting comparison of this value of the correlation with the correlation of X and Y having a joint uniform distribution on $\{(x,y) : 0 < x < 1, x < y < x + 0.1\}$, which is a 'narrower strip' of values then previously. Then, $f_{X,Y}(x,y) = 10$ and it can be shown, that $\rho(X,Y) = 10/\sqrt{101}$, which is close to 1. The linear relationship between X and Y is very strong in this case.

$\square$

1.10.6 Bivariate Normal Distribution

Here we use matrix notation. A bivariate rv is treated as a random vector
$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}.$$
The expectation of a bivariate random vector is written as
$$\boldsymbol{\mu} = \mathrm{E}\,\boldsymbol{X} = \mathrm{E}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$
and its variance-covariance matrix is
$$\boldsymbol{V} = \begin{pmatrix} \mathrm{var}(X_1) & \mathrm{cov}(X_1,X_2) \\ \mathrm{cov}(X_2,X_1) & \mathrm{var}(X_2) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

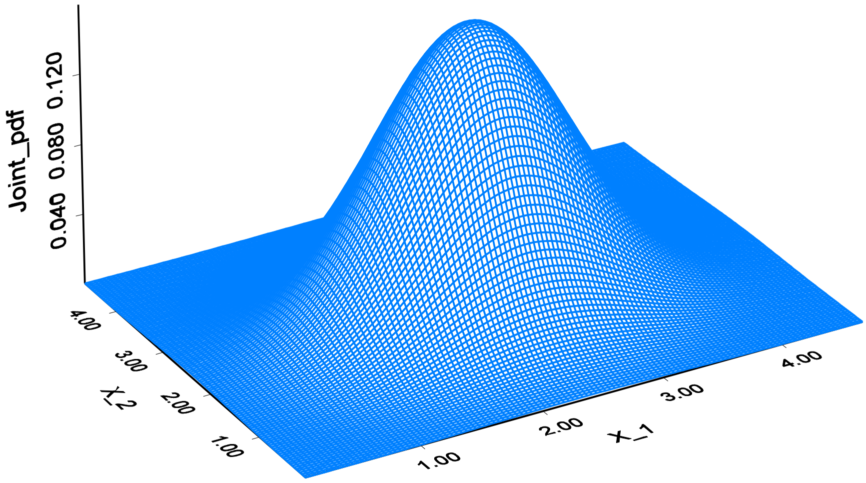

Figure 1.2: Bivariate Normal pdf

Then the joint pdf of a normal bi-variate rv $\boldsymbol{X}$ is given by

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{\det(\boldsymbol{V})}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \tag{1.18}$$

where $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$.

The determinant of $\boldsymbol{V}$ is

$$\det \boldsymbol{V} = \det\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = (1-\rho^2)\sigma_1^2\sigma_2^2.$$

Hence, the inverse of $\boldsymbol{V}$ is

$$\boldsymbol{V}^{-1} = \frac{1}{\det \boldsymbol{V}}\begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} = \frac{1}{1-\rho^2}\begin{pmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{pmatrix}.$$

Then the exponent in formula (1.18) can be written as

$$\begin{aligned}
&-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{V}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \\
&= -\frac{1}{2(1-\rho^2)}(x_1-\mu_1, x_2-\mu_2)\begin{pmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{pmatrix}\begin{pmatrix} x_1-\mu \\ x_2-\mu \end{pmatrix} \\
&= -\frac{1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right).
\end{aligned}$$

So, the joint pdf of the two-dimensional normal rv $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}$$
$$\times \exp\left\{\frac{-1}{2(1-\rho^2)}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right\}.$$

Note that when $\rho = 0$ it simplifies to

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_1^2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right)\right\},$$

which can be written as a product of the marginal distributions of X_1 and X_2. Hence, if $\boldsymbol{X} = (X_1, X_2)^{\mathrm{T}}$ has a bivariate normal distribution and $\rho = 0$ then the variables X_1 and X_2 are independent.

1.10.7 Bivariate Transformations

Theorem 1.17. *Let X and Y be jointly continuous random variables with joint pdf $f_{X,Y}(x, y)$ which has support on $\mathcal{S} \subseteq \mathbb{R}^2$. Consider random variables $U = g(X, Y)$ and $V = h(X, Y)$, where $g(\cdot, \cdot)$ and $h(\cdot, \cdot)$ form a one-to-one mapping from $\mathcal{S}$ to $\mathcal{D}$ with inverses $x = g^{-1}(u, v)$ and $y = h^{-1}(u, v)$ which have continuous partial derivatives. Then, the joint pdf of (U, V) is*

$$f_{U,V}(u, v) = f_{X,Y}\left(g^{-1}(u, v), h^{-1}(u, v)\right)|J|,$$

where, the Jacobian of the transformation J is

$$J = \det\begin{pmatrix} \frac{\partial g^{-1}(u,v)}{\partial u} & \frac{\partial g^{-1}(u,v)}{\partial v} \\ \frac{\partial h^{-1}(u,v)}{\partial u} & \frac{\partial h^{-1}(u,v)}{\partial v} \end{pmatrix}$$

for all $(u, v) \in \mathcal{D}$ □

Example 1.29. Let X, Y be independent rvs and $X \sim \operatorname{Exp}(\lambda)$ and $Y \sim \operatorname{Exp}(\lambda)$. Then, the joint pdf of (X, Y) is

$$f_{X,Y}(x, y) = \lambda e^{-\lambda x}\lambda e^{-\lambda y} = \lambda^2 e^{-\lambda(x+y)}$$

on support $\mathcal{S} = \{(x, y) : x > 0, y > 0\}$.

We will find the joint pdf for (U, V), where $U = g(X, Y) = X + Y$ and $V = h(X, Y) = X/Y$. This transformation and the support for (X, Y) give the support for (U, V). This is $\{(u, v) : u > 0, v > 0\}$.

The inverse functions are

$$x = g^{-1}(u, v) = \frac{uv}{1+v} \quad \text{and } y = h^{-1}(u, v) = \frac{u}{1+v}.$$

The Jacobian of the transformation is equal to

$$J = \det\begin{pmatrix} \frac{\partial g^{-1}(u,v)}{\partial u} & \frac{\partial g^{-1}(u,v)}{\partial v} \\ \frac{\partial h^{-1}(u,v)}{\partial u} & \frac{\partial h^{-1}(u,v)}{\partial v} \end{pmatrix} = \det\begin{pmatrix} \frac{v}{1+v} & \frac{u}{(1+v)^2} \\ \frac{1}{1+v} & -\frac{u}{(1+v)^2} \end{pmatrix} = \frac{-u}{(1+v)^2}.$$

Hence, by Theorem 1.17 we can write

$$\begin{aligned} f_{U,V}(u, v) &= f_{X,Y}\left(g^{-1}(u, v), h^{-1}(u, v)\right)|J| \\ &= \lambda^2 \exp\left\{-\lambda\left(\frac{uv}{1+v} + \frac{u}{1+v}\right)\right\} \times \frac{u}{(1+v)^2} \\ &= \frac{\lambda^2 u e^{-\lambda u}}{(1+v)^2}, \end{aligned}$$

for $u, v > 0$. □

These transformed variables are independent. In a simpler situation where $g(x)$ is a function of x only and $h(y)$ is function of y only, it is easy to see the following very useful result.

Theorem 1.18. *Let X and Y be independent rvs and let $g(x)$ be a function of x only and $h(y)$ be function of y only. Then the functions $U = g(X)$ and $V = h(Y)$ are independent.* □

Proof. (Continuous case) For any $u \in \mathbb{R}$ and $v \in \mathbb{R}$, define

$$A_u = \{x : g(x) \le u\} \quad \text{and} \quad A_v = \{y : h(y) \le v\}.$$

Then, we can obtain the joint cdf of (U, V) as follows

$$\begin{aligned} F_{U,V}(u, v) &= P(U \le u, V \le v) = P(X \in A_u, Y \in A_v) \\ &= P(X \in A_u)P(Y \in A_v) \quad \text{as } X \text{ and } Y \text{ are independent.} \end{aligned}$$

The mixed partial derivative with respect to u and v will give us the joint pdf for (U, V). That is,

$$f_{U,V}(u,v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u,v) = \left(\frac{d}{du} P(X \in A_u)\right)\left(\frac{d}{dv} P(Y \in A_v)\right)$$

as the first factor depends on u only and the second factor on v only. Hence, the rvs $U = g(X)$ and $V = h(Y)$ are independent. $\square$

Exercise 1.18. Let (X, Y) be a two-dimensional random variable with joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 8xy, & \text{for } 0 \leq x < y \leq 1; \\ 0, & \text{otherwise.} \end{cases}$$

Let $U = X/Y$ and $V = Y$.

(a) Are the variables X and Y independent? Explain.

(b) Calculate the covariance of X and Y.

(c) Obtain the joint pdf of (U, V).

(d) Are the variables U and V independent? Explain.

(e) What is the covariance of U and V?

Exercise 1.19. Let X and Y be independent random variables such that

$$X \sim \text{Exp}(\lambda) \quad \text{and} \quad Y \sim \text{Exp}(\lambda).$$

(a) Find the joint probability density function of (U, V), where

$$U = \frac{X}{X+Y} \quad \text{and} \quad V = X + Y.$$

(b) Are the variables U and V independent? Explain.

(c) Show that U is uniformly distributed on $(0, 1)$.

(d) What is the distribution of V?