- The variance of a random variable X with mean μ is given by

$$\begin{aligned}
\text{var}(X) \equiv \sigma^2 &\equiv E\left[(X - E(X))^2 \right] \\
&\equiv E\left[(X-\mu)^2 \right] \\
&\equiv \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx \\
&\equiv \int_{-\infty}^{\infty} x^2 f(x)\, dx - \left[\int_{-\infty}^{\infty} x f(x) dx \right]^2 \\
&\equiv E(x^2) - E^2(x)
\end{aligned} \tag{39}$$

- The variance is a measure of the dispersion of the random variable about the mean.

Definition

Let X and Y be any two random variables defined in the same probability space. The covariance of X and Y, denoted $\text{cov}[X, Y]$ or $\sigma_{X,Y}$, is defined as

$$\begin{aligned}
\text{cov}[X, Y] &\equiv E\left[(X-\mu_X)(Y-\mu_Y)\right] \\
&\equiv E[XY] - E[X]E[Y] \\
&\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\,dx\,dy \\
&- \left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x, y)\,dx\,dy \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x, y)\,dx\,dy\right] \qquad (40) \\
&\equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\,dx\,dy \\
&- \left[\int_{-\infty}^{\infty} xf_X(x, y)\,dx \cdot \int_{-\infty}^{\infty} yf_Y(x, y)\,dy\right]
\end{aligned}$$

- The covariance measures the interaction between two random variables
  - Note: its numerical value is not independent of the units of measurement of X and Y.
- Positive values of the covariance imply that X increases when Y increases; negative values indicate X decreases as Y decreases.
- Example: Let the joint density of two random variables X_1 and X_2 be given by

$$f(x_{1,}x_2) = \begin{cases} 2x_2 e^{-x_1} & x_1 \geq 0,\ 0 \leq x_2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

- We already showed that

$$E[X_1] = 1,\ E[X_2] = \frac{2}{3},\ \text{and } E[X_1 X_2] = \frac{2}{3}$$

- The covariance can the be found:

$$\begin{aligned}\mathrm{cov}[X_1, X_2] &\equiv E[X_1 X_2] - E[X_1]E[X_2] \\ &\equiv \frac{2}{3} - (1)\left(\frac{2}{3}\right) = 0\end{aligned}$$

Variance, Covariance and Correlation

Example

- Let the joint density of two random variables X_1 and X_2 be given by

$$f(x_1, x_2) = \begin{cases} \frac{1}{6}x_1 & 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

- First compute the expected value of $X_1 X_2$ as follows.

$$E[X_1 X_2] = \int_0^3 \int_0^2 \frac{1}{6} x_1^2 x_2 \, dx_1 \, dx_2 = \int_0^3 \left(\frac{1}{18} x_1^3 x_2 \Big|_0^2 \right) dx_2$$

$$= \int_0^3 \frac{8}{18} x_2 \, dx_2 = \int_0^3 \frac{4}{9} x_2 \, dx_2$$

$$= \frac{4}{18} x_2^2 \Big|_0^3 = \frac{36}{18} = 2$$

- Then compute expected value of X_1 as follows

$$E[X_1] = \int_0^3 \int_0^2 \frac{1}{6} x_1^2 \, dx_1 \, dx_2 = \int_0^3 \left(\frac{1}{18} x_1^2 \Big|_0^2 \right) dx_2$$
$$= \int_0^3 \frac{4}{9} \, dx_2 = \frac{4}{9} x_2 \Big|_0^3 = \frac{4}{3}$$

- Then compute the expected value of X_2 as follows.

$$E[X_2] = \int_0^3 \int_0^2 \frac{1}{6} x_1 x_2 \, dx_1 \, dx_2 = \int_0^3 \left(\frac{1}{12} x_1^2 x_2 \Big|_0^2 \right) dx_2$$
$$= \int_0^3 \frac{4}{12} x_2 \, dx_2 = \int_0^3 \frac{1}{3} x_2 \, dx_2 = \frac{1}{6} x_2^2 \Big|_0^3 = \frac{9}{6} = \frac{3}{2}$$

- The covariance is then given by

$$\begin{aligned}\operatorname{cov}[X_1, X_2] &\equiv E[X_1 X_2] - E[X_1]E[X_2] \\ &\equiv 2 - \left(\frac{4}{3}\right)\left(\frac{3}{2}\right) = 2 - 2 = 0\end{aligned}$$

- New example: Let the joint density of two random variables X_1 and X_2 be given by

$$f(x_1, x_2) = \begin{cases} \frac{3}{8}x_1 & 0 \leq x_2 \leq x_1 \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

- First compute the expected value of $X_1 X_2$ as follows.

$$E\left[X_1\,X_2\right] = \int_0^2 \int_{x_2}^2 \frac{3}{8} x_1^2 x_2 \, dx_1 \, dx_2 = \int_0^2 \left(\frac{3}{24} x_1^3 x_2 \Big|_{x_2}^{2} \right) \, dx_2$$

$$= \int_0^2 \left(\frac{24}{24} x_2 - \frac{3}{24} x_2^4 \right) \, dx_2 = \int_0^2 \left(x_2 - \frac{1}{8} x_2^4 \right) \, dx_2$$

$$= \left(\frac{x_2^2}{2} - \frac{1}{40} x_2^5 \right) \Big|_0^2 = \frac{4}{2} - \frac{32}{40} = 2 - \frac{4}{5} = \frac{6}{5}$$

- Then compute expected value of X_1 as follows

$$E\left[X_1\right] = \int_0^2 \int_0^{x_1} \frac{3}{8} x_1^2 \, dx_2 \, dx_1 = \int_0^2 \left(\frac{3}{8} x_1^2 x_2 \Big|_0^{x_1} \right) \, dx_1$$

$$= \int_0^2 \frac{3}{8} x_1^3 \, dx_1 = \frac{3}{32} x_1^4 \Big|_0^2 = \frac{48}{32} = \frac{3}{2}$$

- Then compute the expected value of X_2 as follows.

$$E\left[X_2\right] = \int_0^2 \int_{x_2}^2 \frac{3}{8} x_1 x_2 \, dx_1 \, dx_2 = \int_0^2 \left(\frac{3}{16} x_1^2 x_2 \Big|_{x_2}^2 \right) dx_2$$

$$= \int_0^2 \left(\frac{12}{16} x_2 - \frac{3}{16} x_2^3 \right) dx_2 = \int_0^2 \left(\frac{3}{4} x_2 - \frac{3}{16} x_2^3 \right) dx_2$$

$$= \left(\frac{3}{8} x_2^2 - \frac{3}{64} x_2^4 \right) \Big|_0^2 = \frac{12}{8} - \frac{48}{64}$$

$$= \frac{96}{64} - \frac{48}{64} = \frac{48}{64} = \frac{3}{4}$$

- The covariance is then given by

$$cov[X_{1,}X_2] = E[X_1X_2] - E[X_1]E[X_2] \equiv \frac{6}{5} - \left(\frac{3}{2}\right)\left(\frac{3}{4}\right) = \frac{3}{40}$$

Definition

The correlation coefficient, denoted by $\rho[X, Y]$, or $\rho_{X,Y}$ of random variables X and Y is defined to be

$$\rho_{X,Y} = \frac{\mathrm{cov}[X, Y]}{\sigma_X \sigma_Y} \tag{41}$$

provided that $cov[X, Y]$, σ_X and σ_Y exist, and σ_X, σ_Y are positive.

- The correlation coefficient is independent of the units of measurement.
- The correlation coefficient is bounded between negative one and one.
- The sign of the correlation coefficient is the same as the sign of the covariance.

Theorem

If X and Y are independent random variables, then

$$cov[X,\ Y] = 0. \tag{42}$$

Proof: We know from the definition of covariance that

$$\text{cov}[X,\ Y] = E[XY] - E[X]E[Y] \tag{43}$$

We also know that if X and Y are independent, then

$$E\left[g(X)h(Y)\right] = E\left[g(X)\right]E\left[h(Y)\right] \tag{44}$$

Let $g(X) = X$ and $h(Y) = Y$ to obtain

$$E\left[XY\right] = E\left[X\right]E\left[Y\right] \tag{45}$$

Substituting into equation 43 we obtain

$$\text{cov}[X,\ Y] = E[X]E[Y] - E[X]E[Y] = 0 \tag{46}$$

Variance, Covariance and Correlation

Independence and Covariance

- The converse is not true, i.e., $\text{cov}[X, Y] = 0$ does not imply X and Y are independent.
- Example: Consider the following discrete probability distribution.

		x_1			
		-1	0	1	
x_2	-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{5}{16}$
	0	$\frac{3}{16}$	0	$\frac{3}{16}$	$\frac{6}{16} = \frac{3}{8}$
	1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{5}{16}$
		$\frac{5}{16}$	$\frac{6}{16} = \frac{3}{8}$	$\frac{5}{16}$	1

- These random variables are not independent because the joint probabilities are not the product of the marginal probabilities.

- For example

$$p_{X_1X_2}[-1,\,-1] = \frac{1}{16} \neq p_{X_1}(-1)p_{X_1}(-1) = \left(\frac{5}{16}\right)\left(\frac{5}{16}\right) = \frac{25}{256}$$

- Now compute the covariance between X_1 and X_2. First find $E[X_1]$ as follows

$$E[X_1] = (-1)\left(\frac{5}{16}\right) + (0)\left(\frac{6}{16}\right) + (1)\left(\frac{5}{16}\right) = 0$$

- Similarly for the expected value of X_2.

$$E[X_2] = (-1)\left(\frac{5}{16}\right) + (0)\left(\frac{6}{16}\right) + (1)\left(\frac{5}{16}\right) = 0$$

- Now compute $E[X_1 X_2]$ as follows

$$
\begin{aligned}
E[X_1 X_2] &= (-1)(-1)\left(\frac{1}{16}\right) + (-1)(0)\left(\frac{3}{16}\right) + (-1)(1)\left(\frac{1}{16}\right) \\
&\quad + (0)(-1)\left(\frac{3}{16}\right) + (0)(0)(0) + (0)(1)\left(\frac{3}{16}\right) \\
&\quad + (1)(-1)\left(\frac{1}{16}\right) + (1)(0)\left(\frac{3}{16}\right) + (1)(1)\left(\frac{1}{16}\right) \\
&= \frac{1}{16} - \frac{1}{16} - \frac{1}{16} + \frac{1}{16} = 0
\end{aligned}
$$

- The covariance is then

$$\begin{aligned} \text{cov}[X,\ Y] &\equiv E[XY] - E[X]E[Y] \\ &\equiv 0 - (0)\,(0) = 0 \end{aligned}$$

- In this case the covariance is zero, but the variables are not independent.

$$
\begin{aligned}
\text{var}[a_1 x_1 + a_2 x_2] &= a_1^2 \text{var}(x_1) + a_2^2 \text{var}(x_2) + 2 a_1 a_2 \text{cov}(x_1,\, x_2) \\
&= a_1^2 \sigma_1^2 + 2 a_1 a_2 \sigma_{12} + a_2^2 \sigma_2^2 \\
&= \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\
&= \text{var}\left[\begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right]
\end{aligned} \tag{47}
$$