Jointly distributed random variables

So far we have been only dealing with the probability distributions of single random variables. However, we are often interested in probability statements concerning two or random variables. In order to do this, we define the joint (cumulative) distribution functions of these random variables.

**Definition 1** Suppose that \( X \) and \( Y \) are two random variables. The joint (cumulative) distribution function of \( X \) and \( Y \) is the function on \( \mathbb{R}^2 \) defined by

\[
F(x, y) = \mathbb{P}(X \leq x, Y \leq y), \quad (x, y) \in \mathbb{R}^2.
\]

We can get the distribution function of \( X \) easily from the joint distribution function of \( X \) and \( Y \):

\[
F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \leq x, Y < \infty) = \mathbb{P}( \lim_{y \to \infty} \{X \leq x, Y \leq y\}) = \lim_{y \to \infty} F(x, y) = F(x, \infty).
\]

Similarly we can get the distribution function of \( Y \) easily from the joint distribution function of \( X \) and \( Y \):

\[
F_Y(y) = \lim_{x \to \infty} F(x, y) = F(\infty, y).
\]

The distribution functions \( F_X \) and \( F_Y \) are sometimes called the marginal distribution functions of \( X \) and \( Y \) respectively.

The joint distribution function \( F \) of \( X \) and \( Y \) contains all the statistical information about \( X \) and \( Y \). In particular, given the joint distribution function \( F \) of \( X \) and \( Y \), we can calculate the probability of any event defined in terms of \( X \) and \( Y \). For instance, for any real numbers \( a \leq b \) and \( c \leq d \), we have

\[
\mathbb{P}(a < X \leq b, c < Y \leq d) = F(b, d) - F(a, d) - F(b, c) + F(a, c).
\]

In the case when \( X \) and \( Y \) are both discrete random variables, it is more convenient to use the joint probability mass function of \( X \) and \( Y \).

**Definition 2** Suppose that \( X \) and \( Y \) are two discrete random variables. The joint probability mass function of \( X \) and \( Y \) is the function on \( \mathbb{R}^2 \) defined by

\[
p(x, y) = \mathbb{P}(X = x, Y = y), \quad (x, y) \in \mathbb{R}^2.
\]

The mass function \( p_X \) of \( X \) can be easily obtained from the joint mass function of \( X \) and \( Y \):

\[
p_X(x) = \mathbb{P}(X = x) = \sum_{y : p(x, y) > 0} p(x, y), \quad x \in \mathbb{R}.
\]
Similarly,
\[ p_Y(y) = P(Y = y) = \sum_{x: p(x,y) > 0} p(x,y), \quad y \in \mathbb{R}. \]

The mass functions \( p_X \) and \( p_Y \) are sometimes called the marginal mass functions of \( X \) and \( Y \) respectively.

**Example 3** A box contains 3 balls labeled 1, 2 and 3. 2 balls are randomly drawn from the box without replacement. Let \( X \) be the number on the first ball and \( Y \) the number on the second ball. Then the joint mass function of \( X \) and \( Y \) is given by \( p(1,2) = p(1,3) = p(2,1) = p(2,3) = p(3,1) = p(3,2) = \frac{1}{6} \) and \( p(x,y) = 0 \) elsewhere.

This joint mass function can be expressed by the following table

<table>
<thead>
<tr>
<th>( x )</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1/6</td>
<td>1/6</td>
</tr>
<tr>
<td>2</td>
<td>1/6</td>
<td>0</td>
<td>1/6</td>
</tr>
<tr>
<td>3</td>
<td>1/6</td>
<td>1/6</td>
<td>1/6</td>
</tr>
</tbody>
</table>

**Example 4** Suppose that \( X \) and \( Y \) are two discrete random variables with joint mass function given by the following table

<table>
<thead>
<tr>
<th>( x )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/4</td>
<td>1/8</td>
<td>1/16</td>
<td>1/16</td>
</tr>
<tr>
<td>2</td>
<td>1/16</td>
<td>1/16</td>
<td>1/4</td>
<td>1/8</td>
</tr>
</tbody>
</table>

Then \( p_X(1) = \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{1}{2} \). Similarly, \( p_Y(1) = \frac{1}{4} + \frac{1}{16} = \frac{5}{16} \), \( p_Y(2) = \frac{3}{16} \), \( p_Y(3) = \frac{5}{16} \), \( p_Y(4) = \frac{3}{16} \).

**Example 5** Suppose that \( X \) and \( Y \) are two discrete random variables with joint mass function given by the following table

<table>
<thead>
<tr>
<th>( x )</th>
<th>-2</th>
<th>0</th>
<th>2</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td>1/9</td>
<td>1/27</td>
<td>1/27</td>
<td>1/9</td>
</tr>
<tr>
<td>1</td>
<td>2/9</td>
<td>0</td>
<td>1/9</td>
<td>1/9</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1/9</td>
<td>4/27</td>
</tr>
</tbody>
</table>

Find (a) \( P(Y \text{ is even}) \); (b) \( P(XY \text{ is odd}) \).

**Solution** (a) \( P(Y \text{ is even}) = P(Y = -2) = \frac{8}{27} \). (b) \( P(XY \text{ is odd}) = P(X = -1, Y = 1) + P(X = -1, Y = 3) = \frac{2}{9} \).

Suppose that \( X \) and \( Y \) are two random variables. We say that \( X \) and \( Y \) are jointly absolutely continuous if there is a nonnegative function \( f(x,y) \) on \( \mathbb{R}^2 \) such that for any region \( C \) in \( \mathbb{R}^2 \),
\[
\mathbb{P}((X,Y) \in C) = \iint_{(x,y) \in C} f(x,y) \, dx \, dy.
\]
The function $f(x, y)$ is called the joint probability density function of $X$ and $Y$.

If $A$ and $B$ are subset of $\mathbb{R}$, then as a special case of the display above, we have

$$
P(X \in A, Y \in B) = \int_B \int_A f(x, y) dx dy.
$$

Because

$$
F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv,
$$

it follows, upon differentiation, that

$$
f(x, y) = \frac{\partial^2 F}{\partial x \partial y}(x, y).
$$

If $X$ and $Y$ are jointly absolutely continuous, then both $X$ and $Y$ are absolutely continuous and their density functions can be obtained as follows:

$$
f_X(x) = \int_{-\infty}^\infty f(x, y) dy, \quad f_Y(y) = \int_{-\infty}^\infty f(x, y) dx.
$$

The proofs of these two formulae are exactly the same, we give give the proof of the first one: for any subset $A$ of $\mathbb{R}$,

$$
P(X \in A) = P(X \in A, Y \in (-\infty, \infty)) = \int_A \int_{-\infty}^\infty f(x, y) dy dx.
$$

The densities $f_X$ and $f_Y$ are sometimes called the marginal densities of $X$ and $Y$ respectively.

**Example 6** Suppose that $X$ and $Y$ are jointly absolutely continuous with density function given by

$$
f(x, y) = \begin{cases} 
2e^{-x}e^{-2y}, & 0 < x < \infty, 0 < y < \infty \\
0, & \text{otherwise}
\end{cases}
$$

Find (a) $P(X > 1, Y < 1)$ (b) $P(X < Y)$.

**Solution.** (a)

$$
P(X > 1, Y < 1) = \int_1^\infty \left( \int_0^1 2e^{-x}e^{-2y} dy \right) dx = \int_1^\infty e^{-x} dx \int_0^1 2e^{-2y} dy
$$

$$
= e^{-1}(1 - e^{-2}).
$$

(b)

$$
P(X < Y) = \int_0^\infty \left( \int_x^\infty 2e^{-x}e^{-2y} dy \right) dx = \int_0^\infty e^{-x} \left( \int_x^\infty 2e^{-2y} dy \right) dx
$$

$$
= \int_0^\infty e^{-x}e^{-2x} dx = \int_0^\infty e^{-3x} dx = \frac{1}{3}.
$$

**Example 7** A point $(X, Y)$ is randomly chosen from a disk of radius $R$ centered at the origin so that each point in the disk is equally likely. Then $X$ and $Y$ are jointly absolutely continuous with joint density given by

$$
f(x, y) = \begin{cases} 
\frac{1}{\pi R^2}, & x^2 + y^2 < R^2 \\
0, & \text{otherwise}
\end{cases}
$$

Find the marginal densities of $X$ and $Y$. Find the density of $Z = \sqrt{X^2 + Y^2}$.

**Solution.**
We can also define the joint distributions for $n$ random variables in exactly the same way as for we did for $n = 2$. For instance, the joint distribution function $F$ of the $n$ random variables $X_1, \ldots, X_n$ is defined by

$$F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n.$$ 

Further, the $n$ random variables are said to be jointly absolutely continuous if there is a nonnegative function $f$ on $\mathbb{R}^n$ such that for any region $C$ is $\mathbb{R}^n$,

$$P((X_1, \ldots, X_n) \in C) = \int \cdots \int_{\{x_1, \ldots, x_n\} \in C} f(x_1, \ldots, x_n) dx_1 \cdots dx_n.$$

In particular, for any $n$ subsets $A_1, \ldots, A_n$ of $\mathbb{R}$,

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = \int_{A_n} \cdots \int_{A_1} f(x_1, \ldots, x_n) dx_1 \cdots dx_n.$$

One can similarly define the joint mass function of $n$ discrete random variables.

**Independent random variables**

Two random variables $X$ and $Y$ are said to be independent if for any any subsets $A$ and $B$ of $\mathbb{R}$,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

It can be shown that the following results are true.

**Theorem 8** Two random variables $X$ and $Y$ are independent if and only if

$$F(x, y) = F_X(x)F_Y(y), \quad (x, y) \in \mathbb{R}^2.$$

**Theorem 9** Two discrete random variables $X$ and $Y$ are independent if and only if

$$p(x, y) = p_X(x)p_Y(y), \quad (x, y) \in \mathbb{R}^2.$$

**Theorem 10** Suppose that $X$ and $Y$ are jointly absolutely continuous. Then $X$ and $Y$ are independent if and only if

$$f(x, y) = f_X(x)f_Y(y), \quad (x, y) \in \mathbb{R}^2.$$

It is easy to see that the random variables $X$ and $Y$ in Example 3, in Example 4 and in Example 5 are dependent, the random variables in Example 6 are independent, and the random variables in Example 7 are dependent.

One can define the independence of $n > 2$ random variables similar to the case $n = 2$. Similarly we have the following

$n$ random variables $X_1, \ldots, X_n$ are independent if

$$F_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n).$$

$n$ discrete random variables $X_1, \ldots, X_n$ are independent if

$$p_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n).$$

$n$ jointly absolutely continuous random variables $X_1, \ldots, X_n$ are independent if

$$f_{X_1, \ldots, X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n), \quad (x_1, \ldots, x_n) \in \mathbb{R}^n).$$
Example 11  Suppose that \( n + m \) independent trials, having a common success probability \( p \), are performed. If \( X \) is the number of successes in the first \( n \) trials, \( Y \) is the number of successes in the last \( m \) trials and \( Z \) is the total number of successes, then \( X \) and \( Y \) are independent, but \( X \) and \( Z \) are dependent.

Example 12  Suppose that the number of people entering a post office on a given day is a Poisson random variables with parameter \( \lambda \). Suppose that each person entering the post office is a male with probability \( p \) and a female with probability \( 1 - p \). Let \( X \) be the number of males entering the post office on that day and \( Y \) the number of females entering the post office on that day. Find the joint mass function of \( X \) and \( Y \).

Example 13  A man and a woman decides to meet at a certain location. If each person arrives independently at a time uniformly distributed between noon and 1 pm, find the probability that the first to arrive has to wait longer than 10 minutes.

Solution  Using geometric considerations, we can easily find the answer is \( \frac{25}{36} \).

Example 14  Suppose that \( X \) and \( Y \) are independent geometric random variables with parameter \( p \). (a) Find the distribution of \( \min(X, Y) \). (b) Find \( \mathbb{P}(Y \geq X) \). (c) Find the distribution of \( X + Y \). (d) Find \( \mathbb{P}(Y = y|X + Y = z) \) for \( z \geq 2 \) and \( y = 1, \ldots, z - 1 \).

Solution. (a) For any positive integer \( z \), we have

\[
\mathbb{P}(\min(X, Y) \geq z) = \mathbb{P}(X \geq z, Y \geq z) = \mathbb{P}(X \geq z)\mathbb{P}(Y \geq z) = (1 - p)^{2(z-1)}.
\]

Thus \( \min(X, Y) \) is a geometric random variable with parameter

\[
1 - (1 - p)^2 = 2p - p^2.
\]

(b) \[
\mathbb{P}(Y \geq X) = \sum_{x=1}^{\infty} \mathbb{P}(X = x, Y \geq X) = \sum_{x=1}^{\infty} \mathbb{P}(X = x, Y \geq x)
= \sum_{x=1}^{\infty} \mathbb{P}(X = x)\mathbb{P}(Y \geq x) = \sum_{x=1}^{\infty} p(1 - p)^{x-1}(1 - p)^{x-1}
= p \sum_{x=1}^{\infty} (1 - p)^{2(x-1)} = p \frac{2p - p^2}{2p - p^2}.
\]

(c) Let \( z \geq 2 \) be an integer. Then

\[
\mathbb{P}(X + Y = z) = \sum_{x=1}^{z-1} \mathbb{P}(X = x, X + Y = z) = \sum_{x=1}^{z-1} \mathbb{P}(X = x, Y = z - x)
= \sum_{x=1}^{z-1} \mathbb{P}(X = x)\mathbb{P}(Y = z - x) = \sum_{x=1}^{z-1} p(1 - p)^{x-1}p(1 - p)^{z-x-1}
= (z - 1)p^2(1 - p)^{z-2}.
\]

(d)
\begin{align*}
P(Y = y | X + Y = z) &= \frac{P(Y = y, X + Y = z)}{P(X + Y = z)} = \frac{P(X = z - y, Y = y)}{P(X + Y = z)} \\
&= \frac{P(X = z - y)P(Y = y)}{P(X + Y = z)} = \frac{p(1 - p)^{z-y-1}p(1-p)^y - 1}{(z-1)p^2(1-p)^{z-2}} \\
&= \frac{1}{z - 1}.
\end{align*}

**Sums of Independent Random Variables**

We often have to compute the distribution of \(X + Y\) from the distributions of \(X\) and \(Y\) when \(X\) and \(Y\) are independent. We start with the case when both \(X\) and \(Y\) are absolutely continuous.

Suppose that \(X\) and \(Y\) are independent absolutely continuous random variables with densities \(f_X\) and \(f_Y\) respectively. Then the distribution of \(X + Y\) is

\[
F_{X+Y}(z) = \mathbb{P}(X + Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f_X(x)f_Y(y) dy dx
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_X(u-y)f_Y(y) dy du
\]

\[
= \int_{-\infty}^{z} \int_{-\infty}^{\infty} f_X(u-y)f_Y(y) dy du.
\]

Thus \(X + Y\) is also an absolutely continuous random variable and its density is given by

\[
f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y)f_Y(y) dy = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x) dx.
\]

If, furthermore, \(X\) and \(Y\) are nonnegative, then

\[
f_{X+Y}(z) = \begin{cases} 
\int_{0}^{z} f_X(x)f_Y(z-x) dx, & z > 0 \\
0, & \text{otherwise}
\end{cases}
\]

**Example 15** Suppose that \(X\) and \(Y\) are independent random variables and both are uniformly distributed on \((0, 1)\). Find the density of \(X + Y\).

**Solution** We could solve this problem by the formula above. Here I give a direct solution. By geometric considerations, we can easily find that for \(z \in (0, 1]\),

\[
P(X + Y \leq z) = \frac{z^2}{2},
\]

and for \(z \in (1, 2),

\[
P(X + Y \leq z) = 1 - \frac{(2 - z)^2}{2}.
\]

Thus

\[
f_{X+Y}(z) = \begin{cases} 
z, & 0 < z \leq 1, \\
2 - z, & 1 < z < 2, \\
0, & \text{otherwise}
\end{cases}
\]
**Theorem 16** Suppose that $X$ and $Y$ are independent random variables, $X$ is a Gamma random variable with parameters $(\alpha_1, \lambda)$ and $Y$ is a Gamma random variable with parameters $(\alpha_2, \lambda)$. Then $X + Y$ is a Gamma random variable with parameters $(\alpha_1 + \alpha_2, \lambda)$.

**Proof** $X$ and $Y$ are both positive random variables and

$$f_X(x) = \begin{cases} \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1}e^{-\lambda x} & x > 0, \\
0 & x \leq 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} y^{\alpha_2-1}e^{-\lambda y} & y > 0, \\
0 & y \leq 0. \end{cases}$$

Thus, $f_{X+Y}(z) = 0$ for $z \leq 0$ and for $z > 0$,

$$f_{X+Y}(z) = \frac{\lambda^{\alpha_1+\alpha_2} e^{-\lambda z}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} \int_0^z x^{\alpha_1-1}(z-x)^{\alpha_2-1}dx.$$  

Make the change of variable $x = zu$ in the above integral we get

$$f_{X+Y}(z) = c\lambda^{\alpha_1+\alpha_2+\alpha_1+\alpha_2-1} e^{-\lambda z}, \quad z > 0,$$

where

$$c = \int_0^1 u^{\alpha_1-1}(1-u)^{\alpha_2-1}du \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}.$$  

The proof is now finished.

**Theorem 17** Suppose that $X$ and $Y$ are independent random variables, $X$ is a normal random variable with parameters $(\mu_1, \sigma_1^2)$ and $Y$ is a normal random variable with parameters $(\mu_2, \sigma_2^2)$. Then $X + Y$ is a normal random variable with parameters $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

**Example 18** Suppose that $X$ and $Y$ are independent random variables and both of them are normal random variables with parameters $(0, \sigma^2)$. Find the density of $X^2 + Y^2$.

**Solution** $X^2$ and $Y^2$ are independent random variables and each is a Gamma random variable with parameters $(1, \frac{1}{2\sigma^2})$. Thus by Theorem 16, we know that $X^2 + Y^2$ is a Gamma random variable with parameters $(1, \frac{1}{2\sigma^2})$, which is the same as an exponential random variable with parameter $\frac{1}{2\sigma^2}$.

Now we deal with the case when $X$ and $Y$ are both discrete random variables.

Suppose that $X$ and $Y$ are independent discrete random variables with mass functions $p_X$ and $p_Y$ respectively. Let $x_1, x_2, \ldots$ be the possible values of $X$. For any real number $z$,

$$\mathbb{P}(X + Y = z) = \sum_i \mathbb{P}(X = x_i, X + Y = z) = \sum_i \mathbb{P}(X = x_i, Y = z - x_i)$$

$$= \sum_i \mathbb{P}(X = x_i) \mathbb{P}(Y = z - x_i) = \sum_i p_X(x_i) p_Y(z - x_i).$$

Thus the mass function of $X + Y$ is given by

$$p_{X+Y}(z) = \sum_x p_X(x) p_Y(z - x).$$
If, furthermore, $X$ and $Y$ are integer valued random variables, then the mass function of $X + Y$ is given by

$$p_{X+Y}(z) = \begin{cases} \sum_{x=-\infty}^{\infty} p_X(x)p_Y(z - x), & \text{when } z \text{ is an integer} \\ 0, & \text{otherwise} \end{cases}$$

And if $X$ and $Y$ are nonnegative integer valued random variables, then the mass function of $X + Y$ is given by

$$p_{X+Y}(z) = \begin{cases} \sum_{x=0}^{z} p_X(x)p_Y(z - x), & z = 0, 1, \ldots \\ 0, & \text{otherwise} \end{cases}$$

**Theorem 19** Suppose that $X$ and $Y$ are independent random variables, $X$ is a binomial random variable with parameters $(n, p)$ and $Y$ is a binomial random variable with parameters $(m, p)$. Then $X + Y$ is a binomial random variable with parameters $(n + m, p)$.

**Theorem 20** Suppose that $X$ and $Y$ are independent random variables, $X$ is a Poisson random variable with parameter $\lambda_1$ and $Y$ is a Poisson random variable with parameter $\lambda_2$. Then $X + Y$ is a Poisson random variable with parameter $\lambda_1 + \lambda_2$.

### Expectations of random variables
Given the joint distribution of $X$ and $Y$, we want to find the expectation of a function of $X$ and $Y$. We will deal with this in two separate cases: when $X$ and $Y$ are both discrete or when $X$ and $Y$ are jointly absolutely continuous.

**Theorem 21** If $X$ and $Y$ are discrete random variables with joint mass function $p(x, y)$, then for any function $\phi$ on $\mathbb{R}^2$,

$$E[\phi(X, Y)] = \sum_x \sum_y \phi(x, y)p(x, y).$$

If $X$ and $Y$ are jointly absolutely continuous with joint density $f(x, y)$, then then for any function $\phi$ on $\mathbb{R}^2$,

$$E[\phi(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x, y)f(x, y)dxdy.$$  

By taking the function $\phi$ to be $\phi(x, y) = x + y$, we get the following result:

$$E[X + Y] = E[X] + E[Y].$$

Using induction we immediately get the following

$$E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i].$$

By taking the function $\phi$ to be $\phi(x, y) = xy$, we get the following result:

**Theorem 22** If $X$ and $Y$ are independent random variables, then

$$E[XY] = E[X] \cdot E[Y].$$
More generally, we have the following:

**Theorem 23** If $X$ and $Y$ are independent random variables, then for any functions $\phi$ and $\psi$ on $\mathbb{R}$,

$$E[\phi(X)\psi(Y)] = E[\phi(X)] \cdot E[\psi(Y)].$$

Just as the expectation and variance of a single random variable give us information about the random variable, so does the covariance of two random variables give us information about the relationship between the two random variables.

**Definition 24** The covariance of two random variables $X$ and $Y$, denoted by $\text{Cov}(X, Y)$, is defined as

$$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])].$$

Expanding the right hand side of the above definition, we see that


Note that if $X$ and $Y$ are independent, then $\text{Cov}(X, Y) = 0$. However, the converse is not true. A simple example is as follows. Let $X$ be a random variable such that $P(X = 0) = P(X = 1) = P(X = -1) = \frac{1}{3}$, and let $Y$ be a random variable defined by

$$Y = \begin{cases} 0, & \text{if } X \neq 0 \\ 1, & \text{if } X = 0. \end{cases}$$

$X$ and $Y$ are obviously dependent, and yet $\text{Cov}(X, Y) = 0$.

One measure of degree of dependence between the random variables $X$ and $Y$ is the correlation coefficient $\rho(X, Y)$ defined by

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

The correlation coefficient $\rho(X, Y)$ is always between $-1$ and $1$. $\rho(X, Y) = 1$ if and only if $P(X = aY + b) = 1$ for some $a > 0$ and $b \in \mathbb{R}$, and $\rho(X, Y) = -1$ if and only if $P(X = aY + b) = 1$ for some $a < 0$ and $b \in \mathbb{R}$.

The following result list some properties of covariance.

**Theorem 25**

(i) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$;

(ii) $\text{Cov}(X, X) = \text{Var}(X)$;

(iii) $\text{Cov}(aX, Y) = a\text{Cov}(X, Y)$;

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

9
As a particular case of the last theorem we get

\[ \text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j). \]

In particular, if \( X \) and \( Y \) are independent, then

\[ \text{Var} \left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \text{Var}(X_i). \]

**Example 26** An accident occurs at a point \( X \) that is uniformly distributed on a road of length \( L \). At the time of the accident, an ambulance is at a location \( Y \) that is also uniformly distributed on that road. Assuming that \( X \) and \( Y \) are independent, find the expected distance between the ambulance and the location of the accident.

**Example 27** A box contains 3 balls labeled 1, 2, and 3. 2 balls are randomly selected from the box without replacement. Let \( X \) be the number on the first ball and let \( Y \) be the number on the second ball. Find \( \text{Cov}(X, Y) \).

**Example 28** The joint density of \( X \) and \( Y \) is given by

\[ f(x, y) = \begin{cases} \frac{1}{4}e^{-y/2}, & \text{if } 0 < x < y < \infty, \\ 0, & \text{otherwise}. \end{cases} \]

Find \( \text{Cov}(X, Y) \).

**Example 29** Suppose that \( X \) is a binomial random variable with parameters \((n, p)\).

For \( i = 1, \ldots, n \), let \( X_i = 1 \) if the \( i \)-th trial is a success and \( X_0 = 0 \) otherwise. Then \( X_1, \ldots, X_n \) are independent Bernoulli random variables with parameter \( p \), and \( X = X_1 + \cdots + X_n \). Thus \( E[X] = np \) and \( \text{Var}(X) = np(1-p) \).

**Example 30** Independent trials, each results in a success with probability \( p \), are performed. Let \( X \) be the number trials needed to get a total of \( n \) successes.

Let \( X_1 \) be the number of trials needed to get the first success. For \( i > 1 \), let \( X_i \) be the number of additional trials, after the \((i-1)\)-th success, needed to get the \( i \)-th success. Then \( X_1, \ldots, X_n \) are independent geometric random variables with parameter \( p \). It follows from the definition that \( X = X_1 + \cdots + X_n \). Thus \( E[X] = \frac{n}{p} \) and \( \text{Var}(X) = \frac{n(1-p)}{p^2} \).

**Example 31** Suppose that \( N \) people throw their hats into the center of the room. The hats are mixed up, and each person randomly selects. Let \( X \) be the number of people that get their own hat. Find \( E[X] \) and \( \text{Var}(X) \).

**Solution** For \( i = 1, \ldots, N \), let \( X_i = 1 \) if the \( i \)-th person gets his own hat and \( X_i = 0 \) otherwise. Then \( X = X_1 + \cdots + X_N \). For each \( i = 1, \ldots, N \), \( \mathbb{P}(X_i = 1) = \frac{1}{N} \). Thus \( E[X] = 1 \). For \( i \neq j \), \( \mathbb{P}(X_i = 1, X_j = 1) = \frac{1}{N(N-1)} \). Thus

\[ \text{Cov}(X_i, X_j) = \frac{1}{N(N-1)} - \frac{1}{N^2} = \frac{1}{N} \frac{1}{N(N-1)}. \]
\[ \text{Hence} \]
\[ \text{Var}(X) = \sum_{i=1}^{N} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \]
\[ = N \cdot \frac{1}{N} \left( 1 - \frac{1}{N} \right) + \left( N^2 - N \right) \frac{1}{N} \frac{1}{N(N-1)} = 1 \]

**Example 32** If \( n \) balls are randomly selected (without replacement) from a box containing \( N \) balls of which \( m \) are white. Let \( X \) be the number of white balls selected. Find \( E[X] \) and \( \text{Var}(X) \).

**Solution** For \( i = 1, \ldots, m \), let \( X_i = 1 \) if the \( i \)-th white ball is among the selected and \( X_i = 0 \) otherwise. Then \( X = X_1 + \cdots + X_m \).

Now \( E[X_i] = \mathbb{P}(X_i = 1) = \frac{n}{N} \). Thus \( E[X] = \frac{mn}{N} \). For \( i \neq j \),

\[ E[X_i X_j] = \mathbb{P}(X_i = 1, X_j = 1) = \binom{N-2}{n-2} \binom{N}{n} \frac{n(n-1)}{N(N-1)} \]

and

\[ \text{Cov}(X_i, X_j) = \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2} \]

Therefore

\[ \text{Var}(X) = \sum_{i=1}^{m} \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j) \]
\[ = m \frac{n}{N} \left( 1 - \frac{n}{N} \right) + (m^2 - m) \left( \frac{n(n-1)}{N(N-1)} - \frac{n^2}{N^2} \right) \]
\[ = n \left( \frac{m}{N} \right) \left( 1 - \frac{m}{N} \right) \left( 1 - \frac{n-1}{N-1} \right) \]

**Example 33** Suppose that there are \( N \) different types of coupons, and each time one obtains a coupon, it is equally likely to be any one of the \( N \) types. Find the expected number of coupons one need to amass in order to get a complete set of at least one of each type.

**Solution** Let \( X \) be the number of coupons one need to amass in order to get a complete set of at least one of each type. For \( i = 0, 1, \ldots, N-1 \), let \( X_i \) be the number of additional coupons that need to be obtained after \( i \) distinct types have been collected in order to obtain another distinct type. Then

\[ X = X_0 + X_1 + \cdots + X_{N-1} \]

Obviously \( X_0 = 1 \), and for \( i = 1, \ldots, N_1 \), \( X_i \) is a geometric random variable with parameter \( (N-i)/N \) and thus

\[ E[X_i] = \frac{N}{N-i} \]

Therefore

\[ E[X] = 1 + \frac{N}{N-1} + \cdots + \frac{1}{N} = N \left( 1 + \cdots + \frac{1}{N-1} + \frac{1}{N} \right) \]

**Remark on Notations**