# Random Variables

In probabilistic models, a random variable is a variable whose possible values are numerical outcomes of a random phenomenon. As a function or a map, it maps from an element (or an outcome) of a sample space to real values. The underlying probability over sample space makes it possible to define its cumulative distribution, which allows us to calculate quantities such as its expected value and variance, and the moments of its distribution.

**Random variables.** A numerically valued map $X$ of an outcome $\omega$ from a sample space $\Omega$ to the real line $\mathbb{R}$

$$X : \Omega \to \mathbb{R} : \omega \to X(\omega)$$

is called a **random variable (r.v.)**, and usually determined by an experiment. We conventionally denote random variables by uppercase letters $X, Y, Z, U, V$, etc., from the end-of-alphabet letters. In particular, a **discrete random variable** is a random variable that can take values on a finite set $\{a_1, a_2, \ldots, a_n\}$ of real numbers (usually integers), or on a countably infinite set $\{a_1, a_2, a_3, \ldots\}$. The statement such as "$X = a_i$" is an event since

$$\{\omega : X(\omega) = a_i\}$$

is a subset of a sample space $\Omega$.

**Frequency function.** We can consider the probability of the event $\{X = a_i\}$, denoted by $P(X = a_i)$. The function

$$p(a_i) := P(X = a_i)$$

over the possible values of $X$, say $a_1, a_2, \ldots$, is called a **frequency function**, or a **probability mass function**. The frequency function $p$ must satisfy

$$\sum_i p(a_i) = 1,$$

where the sum is over the possible values of $X$. The frequency function will completely describe the probabilistic nature of random variable.

**Joint distributions of discrete random variables.** When two discrete random variables $X$ and $Y$ are obtained in the same experiment, we can define their **joint frequency function** by

$$p(a_i, b_j) = P(X = a_i, Y = b_j)$$

where $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ are the possible values of $X$ and $Y$, respectively. The **marginal frequency function** of $X$, denoted by $p_X$, can be calculated by

$$p_X(a_i) = P(X = a_i) = \sum_j p(a_i, b_j),$$

where the sum is over the possible values $b_1, b_2, \ldots$ of $Y$. Similarly, the marginal frequency function $p_Y(b_j) = \sum_i p(a_i, b_j)$ of $Y$ is given by summing over the possible values $a_1, a_2, \ldots$ of $X$.

**Example 1.** An experiment consists of throwing a fair coin three times. Let $X$ be the number of heads, and let $Y$ be the number of heads before the first tail.

(a) List the sample space $\Omega$.

(b) Describe the events $\{X = 0\}$, $\{Y = 0\}$, and $\{X = 0, Y = 0\}$.

(c) Find the frequency function $p$ for $X$ and $Y$. And compute the joint frequency $p(0, 0)$.

**Solution.**

(a) $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

(b) $\{X = 0\} = \{TTT\}$, $\{Y = 0\} = \{TTT, THH, THT, TTH\}$, and $\{X = 0, Y = 0\} = \{TTT\}$,

(c) $p_X(0) = \frac{1}{8}$; $p_X(1) = \frac{3}{8}$; $p_X(2) = \frac{3}{8}$; $p_X(3) = \frac{1}{8}$.
$p_y(0) = \frac{1}{2}$; $p_y(1) = \frac{1}{4}$; $p_y(2) = \frac{1}{8}$; $p_y(3) = \frac{1}{8}$.
$p(0, 0) = P(X = 0, Y = 0) = \frac{1}{8}$.

**Cumulative distribution function.** Another useful function is the **cumulative distribution function (cdf)**, and it is defined by

$$F(x) = P(X \leq x), \qquad -\infty < x < \infty.$$

The cdf of a discrete r.v. is a nondecreasing step function. It jumps wherever $p(x) > 0$, and the jump at $a_i$ is $p(a_i)$. cdf's are usually denoted by uppercase letters, while frequency functions are usually denoted by lowercase letters.

**Independent random variables.** Let $X$ and $Y$ be discrete random variables with joint frequency function $p(x, y)$. Then $X$ and $Y$ are said to be **independent**, if they satisfy

$$p(x, y) = p_X(x)p_Y(y)$$

for all possible values of $(x, y)$.

**Example 2.** Continue the same experiment of throwing a fair coin three times. Let $X$ be the number of heads, and let $Y$ be the number of heads before the first tail.

(a) Find the cdf $F(x)$ for $X$ at $x = -1, 0, 1, 2, 2.5, 3, 4.5$.

(b) Are $X$ and $Y$ independent?

**Solution.**

(a) $F(-1) = 0$, $F(0) = \frac{1}{8}$, $F(1) = \frac{1}{2}$, $F(2) = F(2.5) = \frac{7}{8}$, and $F(3) = F(4.5) = 1$.

(b) Since $p_X(0) = \frac{1}{8}$, $p_Y(0) = \frac{1}{2}$, and $p(0,0) = \frac{1}{8}$, we find that $p(0,0) \neq p_X(0)p_Y(0)$. Thus, $X$ and $Y$ are not independent.

**Continuous random variables.** A **continuous random variable** is a random variable whose possible values are real values such as 78.6, 5.7, 10.24, and so on. Examples of continuous random variables include temperature, height, diameter of metal cylinder, etc. In what follows, a random variable means a "continuous" random variable, unless it is specifically said to be discrete. The **probability distribution** of a random variable $X$ specifies how its values are distributed over the real numbers. This is completely characterized by the **cumulative distribution function (cdf)**. The cdf

$$F(t) := P(X \leq t).$$

represents the probability that the random variable $X$ is less than or equal to $t$. Then we say that "the random variable $X$ is distributed as $F(t)$."

**Properties of cdf.** The cdf $F(t)$ must satisfy the following properties:

(a) The cdf $F(t)$ is a positive function.

$$F(t) \geq 0 \text{ for all } t.$$

(b) The cdf $F(t)$ increases monotonically.

$$F(s) \leq F(t) \text{ whenever } s < t.$$

(c) The cdf $F(t)$ must tend to zero as $t$ goes to negative infinity "$-\infty$", and must tend to one as $t$ goes to positive infinity "$+\infty$."

$$\lim_{t \to -\infty} F(t) = 0$$
$$\lim_{t \to +\infty} F(t) = 1$$

**Probability density function (pdf).** It is often the case that the probability that the random variable $X$ takes a value in a particular range is given by the area under a curve over that range of values. This curve is called the **probability density function (pdf)** of the random variable $X$, denoted by $f(x)$. Thus, the probability that "$a \leq X \leq b$" can be expressed by

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

The pdf $f(x)$ must satisfy the following properties:

(a) The pdf $f(x)$ is a positive function [that is, $f(x) \geq 0$].

(b) The area under the curve of $f(x)$ (and above the $x$-axis) is one [that is, $\int_{-\infty}^{\infty} f(x)dx = 1$].

**Relation between cdf and pdf.** The pdf $f(x)$ is related to the cdf $F(t)$ via

$$F(t) = P(X \le t) = \int_{-\infty}^{t} f(x)dx.$$

This implies that such cdf $F(t)$ is a continuous function, and that the pdf $f(x)$ is the derivative of the cdf $F(x)$ if $f$ is continuous at $x$, that is,

$$f(x) = \frac{dF(x)}{dx}.$$

**Example 3.** A random variable $X$ is called a **uniform** random variable on $[a, b]$, when $X$ takes any real number in the interval $[a, b]$ equally likely. Then the pdf of $X$ is given by

$$f(x) = \begin{cases} 1/(b-a) & \text{if } a \le x \le b; \\ 0 & \text{otherwise [that is, if } x < a \text{ or } b < x]. \end{cases}$$

Find the cdf $F(t)$.

**Solution.**

$$F(t) = \int_{-\infty}^{t} f(x)dx = \begin{cases} 0 & \text{if } t < a; \\ \frac{t-a}{b-a} & \text{if } a \le t \le b; \\ 1 & \text{if } t > b. \end{cases}$$

**Quantile function.** Given a cdf $F(x)$, we can define the *quantile function* $Q(u) = \inf\{x : F(x) \ge u\}$ for $u \in (0, 1)$. In particular, if $F$ is continuous and strictly increasing, then we have $Q(u) = F^{-1}(u)$. When $F(x)$ is the cdf of a random variable $X$, and $F(x)$ is continuous, we can find

$$x_p = Q(p) \quad \Leftrightarrow \quad P(X \le x_p) = p$$

for every $p \in (0, 1)$. The value $x_{1/2} = Q(\frac{1}{2})$ is called the *median* of $F$, and the values $x_{1/4} = Q(\frac{1}{4})$ and $x_{3/4} = Q(\frac{3}{4})$ are called the *lower quartile* and the *upper quartile*, respectively.

**Example 4.** Compute the quantile function $Q(u)$ for a uniform random variable $X$ on $[a, b]$. Find the median of distribution.

**Solution.** Restricting the domain $[a, b]$ for $F$, we find that $F(x) = \frac{x-a}{b-a}$ is strictly increasing function from $(a, b)$ to $(0, 1)$, and therefore, that $F(x)$ has the inverse function $Q(u)$. By solving $u = \frac{x-a}{b-a}$ in terms of $x$, we obtain $Q(u) = a + (b-a)u$ for $0 < u < 1$. The median is given by $Q\left(\frac{1}{2}\right) = \frac{a+b}{2}$.

**Joint density function.** Consider a pair $(X, Y)$ of continuous random variables. A **joint density function** $f(x, y)$ is a nonnegative function satisfying

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dy\, dx = 1,$$

and is used to compute probabilities constructed from the random variables $X$ and $Y$ simultaneously by

$$P(a \le X \le b, c \le Y \le d) = \int_a^b \left[ \int_c^d f(x,y)\, dy \right] dx$$

In particular,

$$F(u,v) = P(X \le u,\ Y \le v) = \int_{-\infty}^{u} \int_{-\infty}^{v} f(x,y)\, dy\, dx.$$

is called the **joint distribution function**.

**Marginal densities and independence.** Given the joint density function $f(x,y)$, the distribution for each of $X$ and $Y$ is called the **marginal distribution**, and the **marginal density functions** of $X$ and $Y$, denoted by $f_X(x)$ and $f_Y(y)$, are given respectively by

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

If the joint density function $f(x,y)$ for continuous random variables $X$ and $Y$ is expressed in the form

$$f(x,y) = f_X(x) f_Y(y) \quad \text{for all } x, y,$$

then $X$ and $Y$ are said to be independent.

**Example 5.** Consider the following joint density function for random variables $X$ and $Y$:

$$f(x,y) = \begin{cases} \lambda^2 e^{-\lambda y} & \text{if } 0 \le x \le y; \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

(a) Find $P(X \le 1, Y \le 1)$. (b) Find the marginal density functions $f_X(x)$ and $f_Y(y)$. (c) Are $X$ and $Y$ independent?

**Solution.** For (a) we can calculate it as follows.

$$P(X \le 1, Y \le 1) = \int_0^1 \int_0^y \lambda^2 e^{-\lambda y} dx dy = \int_0^1 \lambda^2 e^{-\lambda y} y dy$$
$$= \int_0^1 \lambda e^{-\lambda y} dy - \left[ \lambda y e^{-\lambda y} \right]_0^1 = \left[ -e^{-\lambda y} \right]_0^1 - \lambda e^{-\lambda} = 1 - e^{-\lambda} - \lambda e^{-\lambda}$$

**Continue the solution.** (b) We obtain

$$f_X(x) = \int_x^{\infty} \lambda^2 e^{-\lambda y} dy = \left[ -\lambda e^{-\lambda y} \right]_{y=x}^{y=\infty} = \lambda e^{-\lambda x}$$
$$f_Y(y) = \int_0^y \lambda^2 e^{-\lambda y} dx = \left[ \lambda^2 e^{-\lambda y} x \right]_{x=0}^{x=y} = \lambda^2 y e^{-\lambda y}$$

for $0 \le x \le \infty$ and $0 \le y \le \infty$.

(c) Since $f(x,y) \ne f_X(x) f_Y(y)$, they are not independent.

**Conditional density functions.** Suppose that two random variables $X$ and $Y$ has a joint density function $f(x,y)$. If $f_X(x) > 0$, then we can define the *conditional density function* $f_{Y|X}(y|x)$ given $X = x$ by

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}.$$

Similarly we can define the *conditional density function* $f_{X|Y}(x|y)$ given $Y = y$ by

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

if $f_Y(y) > 0$. Then, clearly we have the following relation

$$f(x,y) = f_{Y|X}(y|x)f_X(x) = f_{X|Y}(x|y)f_Y(y).$$

**Example 6.** Find the conditional density functions $f_{X|Y}(x|y)$ and $f_{Y|X}(y|x)$ for the joint density function (2.1) in the previous example.

**Solution.**

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \begin{cases} \frac{\lambda^2 e^{-\lambda y}}{\lambda^2 y e^{-\lambda y}} = \frac{1}{y} & \text{if } 0 \le x \le y; \\ 0 & \text{if } x > y \text{ or } x < 0. \end{cases}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \begin{cases} 0 & \text{if } y < x; \\ \frac{\lambda^2 e^{-\lambda y}}{\lambda e^{-\lambda x}} = \lambda e^{-\lambda(y-x)} & \text{if } y \ge x. \end{cases}$$

**Independence of random variables.** In theory we can consider a joint frequency function $p(x,y,z)$ for three discrete random variables $X$, $Y$, and $Z$. If it satisfies

$$p(x,y,z) = p_X(x)p_Y(y)p_Z(z)$$

then they are said to be independent. In the same manner, if a joint density function $f(x,y,z)$ for continuous random variables $X$, $Y$ and $Z$ is expressed in the form

$$f(x,y,z) = f_X(x)f_Y(y)f_Z(z)$$

then they are independent. In general if $X_1, X_2, \ldots, X_n$ are **independent** random variables then we can construct the joint density function by

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) \times f_{X_2}(x_2) \times \cdots \times f_{X_n}(x_n)$$

**Expectation.** Let $X$ be a discrete random variable whose possible values are $a_1, a_2, \ldots$, and let $p(x)$ is the frequency function of $X$. Then the **expectation** (**expected value** or **mean**) of the random variable $X$ is given by

$$E[X] = \sum_i a_i p(a_i).$$

We often denote the expected value $E(X)$ of $X$ by $\mu$ or $\mu_X$. For a function $g$, we can define the expectation of function of random variable by

$$E[g(X)] = \sum_i g(a_i)p(a_i).$$

**Variance.** The **variance** of a random variable $X$, denoted by $\mathrm{Var}(X)$ or $\sigma^2$, is the expected value of "the squared difference between the random variable and its expected value $E(X)$," and can be defined as

$$\mathrm{Var}(X) := E[(X - E(X))^2] = E[X^2] - (E[X])^2.$$

The square-root $\sqrt{\mathrm{Var}(X)}$ of the variance $\mathrm{Var}(X)$ is called the **standard error (SE)** (or **standard deviation (SD)**) of the random variable $X$.

**Example 7.** A random variable $X$ takes on values 0, 1, 2 with the respective probabilities $P(X = 0) = 0.2$, $P(X = 1) = 0.5$, $P(X = 2) = 0.3$. Compute

(a) $E[X]$

(b) $E[X^2]$

(c) $\mathrm{Var}(X)$ and SD of $X$

**Solution.**

(a) $E[X] = (0)(0.2) + (1)(0.5) + (2)(0.3) = 1.1$

(b) $E[X^2] = (0)^2(0.2) + (1)^2(0.5) + (2)^2(0.3) = 1.7$

(c) $\mathrm{Var}(X) = E[(X - 1.1)^2] = (-1.1)^2(0.2) + (-0.1)^2(0.5) + (0.9)^2(0.3) = 0.49$. Also, using $\mathrm{Var}(X) = E[X^2] - (E[X])^2$ we can calculate $\mathrm{Var}(X) = (1.7) - (1.1)^2 = 0.49$.

Then we obtain the SD of $\sqrt{0.49} = 0.7$.

**Expectation for two variables.** Suppose that we have two random variables $X$ and $Y$, and that $p(x, y)$ is their joint frequency function. Then the expectation of function $g(X, Y)$ of the two random variables $X$ and $Y$ is defined by

$$E[g(X, Y)] = \sum_{i,j} g(a_i, b_j)p(a_i, b_j),$$

where the sum is over all the possible values of $(X, Y)$.

**Properties of expectation.** One can think of the expectation $E(X)$ as "an operation on a random variable $X$" which returns the average value for $X$.

(a) Let $a$ be a constant, and let $X$ be a random variable having the frequency function $p(x)$. Then we can show that

$$E[a+X] = \sum_x (a+x)p(x) = a + \sum_x x\,p(x) = a + E[X].$$

(b) Let $a$ and $b$ be scalars, and let $X$ and $Y$ be random variables having the joint frequency function $p(x,y)$ and the respective marginal density functions $p_X(x)$ and $p_Y(y)$.

$$\begin{aligned} E[aX+bY] &= \sum_{x,y} (ax+by)p(x,y) \\ &= a\sum_x x\,p_X(x) + b\sum_y y\,p_Y(y) = aE[X] + bE[Y]. \end{aligned}$$

**Linearity property of expectation.** Let $a$ and $b_1, \ldots, b_n$ be scalars, and let $X_1, \ldots, X_n$ be random variables. By applying the properties of expectation repeatedly, we can obtain

$$\begin{aligned} E\left[a + \sum_{i=1}^n b_i X_i\right] &= a + E\left[\sum_{i=1}^n b_i X_i\right] = a + b_1 E[X_1] + E\left[\sum_{i=2}^n b_i X_i\right] \\ &= \cdots = a + \sum_{i=1}^n b_i E[X_i]. \end{aligned}$$

It is also useful to observe the above property as that of "linear operator."

**Expectation for independent random variables.** Suppose that $X$ and $Y$ are independent random variables. Then the joint frequency function $p(x,y)$ of $X$ and $Y$ can be expressed as

$$p(x,y) = p_X(x)p_Y(y).$$

And the expectation of the function of the form $g_1(X) \times g_2(Y)$ is given by

$$\begin{aligned} E[g_1(X)g_2(Y)] &= \sum_{x,y} g_1(x)g_2(y)\,p(x,y) \\ &= \left[\sum_x g_1(x)p_X(x)\right] \times \left[\sum_y g_2(y)p_Y(y)\right] = E[g_1(X)] \times E[g_2(Y)]. \end{aligned}$$

**Covariance.** Suppose that we have two random variables $X$ and $Y$. Then the **covariance** of two random variables $X$ and $Y$ can be defined as

$$\mathrm{Cov}(X,Y) := E((X-\mu_x)(Y-\mu_y)) = E(XY) - E(X) \times E(Y),$$

where $\mu_x = E(X)$ and $\mu_y = E(Y)$. Then the **correlation coefficient**

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

measures the strength of the dependence of the two random variables.

**Properties of variance and covariance.**

(a) If $X$ and $Y$ are independent, then $\mathrm{Cov}(X, Y) = 0$ by observing that $E[XY] = E[X] \cdot E[Y]$.

(b) In contrast to the expectation, the variance is *not* a linear operator. For two random variables $X$ and $Y$, we have

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y). \tag{2.2}$$

Moreover, if $X$ and $Y$ are independent, by observing that $\mathrm{Cov}(X, Y) = 0$ in (2.2), we obtain $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. In general, we have

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n).$$

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

**Variance and covariance under linear transformation.** Let $a$ and $b$ be scalars (that is, real-valued constants), and let $X$ be a random variable. Then the variance of $aX + b$ is given by

$$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X).$$

Now let $a_1$, $a_2$, $b_1$ and $b_2$ be scalars, and let $X$ and $Y$ be random variables. Then similarly the covariance of $a_1X + b_1$ and $a_2Y + b_2$ can be given by

$$\mathrm{Cov}(a_1X + b_1, a_2Y + b_2) = a_1a_2\mathrm{Cov}(X, Y)$$

**Example 8.** The joint frequency function $p(x, y)$ of two discrete random variables, $X$ and $Y$, is given by

| | | $X$ | | |
|---|---|---|---|---|
| | | $-1$ | $0$ | $1$ |
| $Y$ | $-1$ | $0$ | $1/2$ | $0$ |
| | $1$ | $1/4$ | $0$ | $1/4$ |

(a) Find the marginal frequency function for $X$ and $Y$.

(b) Find $E[X]$ and $E[Y]$.

(c) Find $\mathrm{Cov}(X, Y)$.

(d) Are $X$ and $Y$ independent?

**Solution.**

(a) $p_X(-1) = \frac{1}{4}$; $p_X(0) = \frac{1}{2}$; $p_X(1) = \frac{1}{4}$.
$p_Y(-1) = \frac{1}{2}$; $p_Y(1) = \frac{1}{2}$.

(b) $E[X] = (-1)\left(\frac{1}{4}\right) + (0)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{4}\right) = 0$
$E[Y] = (-1)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{2}\right) = 0$

(c) $E[XY] = (-1)(1)\left(\frac{1}{4}\right) + (0)(-1)\left(\frac{1}{2}\right) + (1)(1)\left(\frac{1}{4}\right) = 0$
Thus, we obtain $\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0$.

(d) No, $X$ and $Y$ are not independent, because $p(-1, -1) = 0 \neq p_X(-1)p_Y(-1) = \frac{1}{8}$.

**Expectation of continuous random variables.** Let $f(x)$ be the pdf of a continuous random variable $X$. Then we define the expectation of the random variable $X$ by

$$E[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx.$$

Furthermore, we can define the expectation $E[g(X)]$ of the function $g(X)$ of random variable by

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx,$$

if $g(x)$ is integrable with respect to $f(x)\, dx$, that is, $\int_{-\infty}^{\infty} |g(x)|\, f(x)\, dx < \infty$.

**Properties of expectation.** When $X$ is a discrete random variable, $E[X]$ is considered as a "linear operator." This remains true even if $X$ is a continuous random variable. Thus, we have the following properties (without proof):

(a) $E[a + X] = a + E[X]$.

(b) $E[aX + bY] = aE[X] + bE[Y]$.

(c) $E\left[a + \sum_{i=1}^{n} b_i X_i\right] = a + \sum_{i=1}^{n} b_i E[X_i]$.

(d) If $X$ and $Y$ are independent then

$$E[g_1(X)g_2(Y)] = E[g_1(X)] \times E[g_2(Y)].$$

**Example 9.** Let $X$ be a uniform random variable on $[a, b]$. Compute $E[X]$ and $E[X^2]$, and find $\mathrm{Var}(X)$.

**Solution.**

$$\begin{aligned}
E[X] &= \int_a^b x \left(\frac{1}{b-a}\right) dx = \left(\frac{1}{b-a}\right)\left[\frac{x^2}{2}\right]_a^b \\
&= \frac{a+b}{2} \\
E[X^2] &= \int_a^b x^2 \left(\frac{1}{b-a}\right) dx = \left(\frac{1}{b-a}\right)\left[\frac{x^3}{3}\right]_a^b \\
&= \frac{a^2+ab+b^2}{3} \\
\mathrm{Var}(X) &= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}
\end{aligned}$$

**Moment generating function.** Let $X$ be a random variable. Then the *moment generating function* (*mgf*) of $X$ is given by

$$M(t) = E\left[e^{tX}\right].$$

The mgf $M(t)$ is a function of $t$ defined on some open interval $(c_0, c_1)$ with $c_0 < 0 < c_1$. In particular when $X$ is a continuous random variable having the pdf $f(x)$, the mgf $M(t)$ can be expressed as

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx.$$

**Properties of moment generating function.** (a) The most significant property of moment generating function is that "the moment generating function uniquely determines the distribution."

(b) Let $a$ and $b$ be constants, and let $M_X(t)$ be the mgf of a random variable $X$. Then the mgf of the random variable $Y = a + bX$ can be given as follows.

$$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t(a+bX)}\right] = e^{at} E\left[e^{(bt)X}\right] = e^{at} M_X(bt).$$

(c) Let $X$ and $Y$ be independent random variables having the respective mgf's $M_X(t)$ and $M_Y(t)$. We can obtain the mgf $M_Z(t)$ of the sum $Z = X + Y$ of random variables as follows.

$$\begin{aligned} M_Z(t) &= E\left[e^{tZ}\right] = E\left[e^{t(X+Y)}\right] = E\left[e^{tX} e^{tY}\right] \\ &= E\left[e^{tX}\right] \cdot E\left[e^{tY}\right] = M_X(t) \cdot M_Y(t). \end{aligned}$$

**Properties of moment generating function, continued.** (d) When $t = 0$, it clearly follows that $M(0) = 1$. Now by differentiating $M(t)$ $r$ times, we obtain

$$M^{(r)}(t) = \frac{d^r}{dt^r} E\left[e^{tX}\right] = E\left[\frac{d^r}{dt^r} e^{tX}\right] = E\left[X^r e^{tX}\right].$$

In particular when $t = 0$, $M^{(r)}(0)$ *generates* the $r$-th moment of $X$ as follows.

$$M^{(r)}(0) = E[X^r], \quad r = 1, 2, 3, \ldots$$

**Example 10.** Find the mgf $M(t)$ for a uniform random variable $X$ on $[a, b]$. Then show that (a) $\lim_{t\to 0} M(t) = 1$ and (b) $\lim_{t\to 0} M'(t) = \frac{a+b}{2}$.

**Solution.**

$$M(t) = \int_a^b e^{tx}\left(\frac{1}{b-a}\right)dx = \begin{cases} \int_a^b \left(\frac{1}{b-a}\right)dx = 1 & \text{if } t = 0; \\ \left(\frac{1}{b-a}\right)\left[\frac{e^{tx}}{t}\right]_{x=a}^{x=b} = \frac{e^{tb}-e^{ta}}{t(b-a)} & \text{if } t \neq 0. \end{cases}$$

$$\lim_{t\to 0} M(t) = \lim_{t\to 0} \frac{e^{tb}-e^{ta}}{t(b-a)} = \lim_{t\to 0} \frac{be^{tb}-ae^{ta}}{b-a} = 1 \text{ by l'Hospital's rule.}$$

$$\lim_{t\to 0} M'(t) = \lim_{t\to 0} \frac{(be^{tb}-ae^{ta})t - (e^{tb}-e^{ta})}{t^2(b-a)}$$

$$= \lim_{t\to 0} \frac{b^2e^{tb}-a^2e^{ta}}{2(b-a)} = \frac{a+b}{2} \text{ again, by l'Hospital's rule.}$$

# Homework

**Problem 1.** Let $p(k)$, $k = -1, 0, 1$, be the frequency function for random variable $X$. Suppose that $p(0) = \frac{1}{4}$, and that $p(-1)$ and $p(1)$ are unknown.

(a) Show that $E[X^2]$ does not depend on the unknown values $p(-1)$ and $p(1)$.

(b) If $E[X] = \frac{1}{4}$, then find the values $p(-1)$ and $p(1)$.

**Problem 2.** The joint frequency function $p(x, y)$ of two discrete random variables, $X$ and $Y$, is given by

| | | $X$ | | |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| $Y$ | 1 | $c$ | $3c$ | $2c$ |
| | 2 | $c$ | $c$ | $2c$ |

(a) Find the constant $c$.

(b) Find $E[X]$ and $E[XY]$.

(c) Are $X$ and $Y$ independent?

**Problem 3.** A pair $(X, Y)$ of discrete random variables has the joint frequency function

$$p(x, y) = \frac{xy}{18}, \ x = 1, 2, 3 \text{ and } y = 1, 2.$$

(a) Find $P(X + Y = 3)$.

(b) Find the marginal frequency function for $X$ and $Y$.

(c) Find $E[Y]$ and $\mathrm{Var}(Y)$.

(d) Are $X$ and $Y$ independent? Justify your answer.

**Problem 4.** (a) Determine the constant $c$ so that $p(x)$ is a frequency function if $p(x) = cx$, $x = 1, 2, 3, 4, 5, 6$.

(b) Similarly find $c$ if $p(x) = c\left(\frac{2}{3}\right)^x$, $x = 1, 2, 3, 4, \ldots$.

**Problem 5.** Let $X$ and $Y$ be independent random variables.

(a) Show that $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$.

(b) If $E[X] = 1, E[Y] = 2$ and $\mathrm{Var}(X) = 4, \mathrm{Var}(Y) = 9$ then find the mean and the variance of $Z = 3X - 2Y$.

**Problem 6.** Suppose that $X$ has the density function $f(x) = cx^2$ for $0 \le x \le 1$ and $f(x) = 0$ otherwise.

(a) Find $c$.

(b) Find the cdf.

(c) What is $P(.1 \le X < .5)$?

**Problem 7.** Let $X$ be a random variable with the pdf

$$f(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1; \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $E[X]$.

(b) Find $E[X^2]$ and $\mathrm{Var}(X)$.

**Problem 8.** Let $F(x) = 1 - \exp(-\alpha x^\beta)$ for $x \ge 0$, $\alpha > 0$ and $\beta > 0$, and $F(x) = 0$ for $x < 0$. Show that $F$ is a cdf, and find the corresponding density.

**Problem 9.** Suppose that random variables $X$ and $Y$ have the joint density

$$f(x, y) = \begin{cases} \frac{6}{7}(x + y)^2 & \text{if } 0 \le x \le 1 \text{ and } 0 \le y \le 1; \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the marginal density functions of $X$ and $Y$.

(b) Find the conditional density functions of $X$ given $Y = y$ and of $Y$ given $X = x$.

(c) Find $E[X]$ and $E[Y]$.

**Problem 10.** Discuss the following properties.

(a) Prove the property of variance: $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

(b) Suppose that $E[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$. Then $Z = (X - \mu)/\sigma$ is said to be standardized, in which $\sigma$ is often referred as "standard error." Show that $E[Z] = 0$ and $\mathrm{Var}(Z) = 1$.

**Problem 11.** Recall that the expectation is an "linear operator." By using the definitions of variance and covariance in terms of the expectation, show that

(a) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$, and

(b) $\mathrm{Var}(X-Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{Cov}(X,Y)$.

**Problem 12.** If $X$ and $Y$ are independent random variables with equal variances, find $\mathrm{Cov}(X+Y, X-Y)$.

**Problem 13.** Suppose that $X$ and $Y$ are independent random variables with their variances $\sigma_x^2 = \mathrm{Var}(X)$ and $\sigma_y^2 = \mathrm{Var}(Y)$. Let $Z = Y - X$.

(a) Find expressions for the covariance of $X$ and $Z$ in terms of $\sigma_x$ and $\sigma_y$.

(b) Find the correlation coefficient $\rho$ of $X$ and $Z$ in terms of $\sigma_x$ and $\sigma_y$.

**Problem 14.** Let $X$ and $Y$ be independent random variables, and let $\alpha$ and $\beta$ be scalars. Find an expression for the mgf $M_Z(t)$ of $Z = \alpha X + \beta Y$ in terms of the mgf's $M_X(t)$ and $M_Y(t)$ of $X$ and $Y$.

# Optional problems

**Problem 15.** The Cauchy distribution function is defined by

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x), \quad -\infty < x < \infty.$$

(a) Show that this is a cdf.

(b) Find the density function.

(c) Suppose that a random variable $X$ has the Cauchy distribution. Can you find $E[X]$? Why or why not?

**Problem 16.** Let $f(x) = (1+\alpha x)/2$ for $-1 \le x \le 1$ and $f(x) = 0$ otherwise, where $-1 \le \alpha \le 1$.

(a) Show that $f$ is a density.

(b) Find the cdf $F(x)$, and sketch the graph of $F(x)$ for $-1 \le \alpha < 0$, $\alpha = 0$, and $0 < \alpha \le 1$.

(c) Find the formulas for the quantile function $Q(u)$.
Hint: Consider the following three cases separately: $-1 \le \alpha < 0$, $\alpha = 0$, and $0 < \alpha \le 1$.

# Answers to exercises

**Problem 1.**

(a) $E[X^2] = p(-1) + p(1) = 1 - p(0) = \frac{3}{4}$

(b) $E[X] = -p(-1) + p(1) = \frac{1}{4}$. Together with $p(-1) + p(1) = \frac{3}{4}$, we obtain $p(-1) = \frac{1}{4}$ and $p(1) = \frac{1}{2}$.

**Problem 2.**

(a) $\sum_{x=1}^{3} \sum_{y=1}^{2} p(x,y) = 10c = 1$ implies that $c = \frac{1}{10}$.

(b) $E[X] = (1)\left(\frac{1}{5}\right) + (2)\left(\frac{2}{5}\right) + (3)\left(\frac{2}{5}\right) = \frac{11}{5}$
$E[Y] = (1)\left(\frac{3}{5}\right) + (2)\left(\frac{2}{5}\right) = \frac{7}{5}$
$E[XY] = (1)(1)\left(\frac{1}{10}\right) + (2)(1)\left(\frac{3}{10}\right) + (3)(1)\left(\frac{2}{10}\right) + (1)(2)\left(\frac{1}{10}\right) + (2)(2)\left(\frac{1}{10}\right) + (3)(2)\left(\frac{2}{10}\right) = \frac{31}{10}$

(c) $X$ and $Y$ are not independent because $p(1,1) = \frac{1}{10} \neq p_X(1)p_Y(1) = \frac{3}{25}$. Or, you can find it by calculating

$$\mathrm{Cov}(X,Y) = E[XY] - E[X]E[Y] = \frac{1}{50}$$

**Problem 3.**

(a) $P(X + Y = 3) = p(1,2) + p(2,1) = \frac{2}{9}$

(b) $p_X(y) = \frac{x}{6}$ for $x = 1, 2, 3$.
$p_Y(y) = \frac{y}{3}$ for $y = 1, 2$.

(c) $E[Y] = (1)\left(\frac{1}{3}\right) + (2)\left(\frac{2}{3}\right) = \frac{5}{3}$; $E[Y^2] = (1)^2\left(\frac{1}{3}\right) + (2)^2\left(\frac{2}{3}\right) = 3$
$\mathrm{Var}(Y) = E[Y^2] - (E[Y])^2 = \frac{2}{9}$

(d) Yes, since the joint frequency function satisfies $p(x,y) = p_X(x)p_Y(y)$ for all $x = 1, 2, 3$ and $y = 1, 2$.

**Problem 4.**

(a) $\sum_{x=1}^{6} p(x) = c \times \frac{6(6+1)}{2} = 21c = 1$ Thus, we obtain $c = \frac{1}{21}$

(b) $\sum_{x=1}^{\infty} p(x) = c \times \frac{(2/3)}{1-(2/3)} = 2c = 1$ Thus, we obtain $c = \frac{1}{2}$

**Problem 5.**

(a) $\mathrm{Var}(aX) = E[(aX - E[aX])^2] = E[(aX - aE[X])^2] = E[a^2(X - E[X])^2] = a^2E[(X - E[X])^2] = a^2\mathrm{Var}(X)$.

(b) $E[Z] = E[3X - 2Y] = (3)(1) - (2)(2) = -1$
$\mathrm{Var}(Z) = \mathrm{Var}(3X + (-2)Y) = \mathrm{Var}(3X) + \mathrm{Var}((-2)Y) = (3)^2(4) + (-2)^2(9) = 72$

**Problem 6.**

(a) $\displaystyle\int_0^1 cx^2\,dx = \frac{c}{3} = 1$. Thus, we have $c = 3$.

(b) $F(t) = \begin{cases} 0 & \text{if } x < 0; \\ t^3 & \text{if } 0 \le x \le 1; \\ 1 & \text{if } x > 1. \end{cases}$

(c) $P(.1 \le X < .5) = \displaystyle\int_{0.1}^{0.5} 3x^2\,dx = \left[x^3\right]_{0.1}^{0.5} = 0.124$

**Problem 7.**

(a) $E[X] = \displaystyle\int_0^1 x(2x)\,dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}$

(b) $E[X^2] = \displaystyle\int_0^1 x^2(2x)\,dx = \left[\frac{x^4}{2}\right]_0^1 = \frac{1}{2}$ Then we obtain $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$.

**Problem 8.** $F(x)$ is nonnegative and increasing for $x \ge 0$, and satisfies $F(0) = 0$, and $\lim_{x\to\infty} F(x) = 1$. Thus, $F(x)$ is cdf. The pdf is given by

$$f(x) = \frac{d}{dx}F(x) = \alpha\beta x^{\beta-1}\exp(-\alpha x^\beta)$$

for $x \ge 0$.

**Problem 9.**

(a) $f_X(x) = \int_0^1 \frac{6}{7}(x+y)^2 dy = \frac{2}{7}(3x^2 + 3x + 1)$, $0 \le x \le 1$;
$f_Y(y) = \int_0^1 \frac{6}{7}(x+y)^2 dx = \frac{2}{7}(3y^2 + 3y + 1)$, $0 \le y \le 1$.

(b) $f_{X|Y}(x|y) = \frac{3(x+y)^2}{3y^2+3y+1}$, $0 \le x \le 1$;
$f_{Y|X}(y|x) = \frac{3(x+y)^2}{3x^2+3x+1}$, $0 \le y \le 1$;

(c) $E[X] = \int_0^1 (x)\left(\frac{2}{7}\right)(3x^2 + 3x + 1)dx = \frac{9}{14} = E[Y]$.

**Problem 10.**

(a) Since $E[aX+b] = aE[X]+b$, by applying the definition of variance we obtain

$$\begin{aligned}\mathrm{Var}(aX+b) &= E[(aX+b-E[aX+b])^2] = E[a^2(X-E[X])^2] \\ &= a^2E[(X-E[X])^2] = a^2\mathrm{Var}(X)\end{aligned}$$

(b) Since $Z = \left(\frac{1}{\sigma}\right)X - \left(\frac{\mu}{\sigma}\right)$ is a linear transformation, we obtain

$$E[Z] = \left(\frac{1}{\sigma}\right)E[X] - \left(\frac{\mu}{\sigma}\right) = 0$$

$$\mathrm{Var}(Z) = \left(\frac{1}{\sigma}\right)^2 \mathrm{Var}(X) = 1$$

**Problem 11.** Let $\mu_x = E[X]$ and $\mu_y = E[Y]$. Then we show (a) as follows:

$$\begin{aligned}\mathrm{Var}(X+Y) &= E[(X+Y-(\mu_x+\mu_y))^2] = E[((X-\mu_x)+(Y-\mu_y))^2] \\ &= E[(X-\mu_x)^2 + 2(X-\mu_x)(Y-\mu_y) + (Y-\mu_y)^2] \\ &= E[(X-\mu_x)^2] + 2E[(X-\mu_x)(Y-\mu_y)] + E[(Y-\mu_y)^2] \\ &= \mathrm{Var}(X) + 2\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)\end{aligned}$$

We can similarly verify (b).

**Problem 12.** Let $\mu_x = E[X]$ and $\mu_y = E[Y]$. By applying the definition of variance and covariance, we obtain

$$\begin{aligned}\mathrm{Cov}(X+Y, X-Y) &= E[(X+Y-(\mu_X+\mu_y))(X-Y-(\mu_x-\mu_y))] \\ &= E[(X-\mu_X)^2 - (Y-\mu_y)^2] \\ &= E[(X-\mu_X)^2] - E[(Y-\mu_y)^2] \\ &= \mathrm{Var}(X) - \mathrm{Var}(Y) = 0\end{aligned}$$

where the last equality comes from the assumption of equal variances.

**Problem 13.** (a) Let $\mu_x = E[X]$ and $\mu_y = E[Y]$. Since $\mathrm{Cov}(X,Y) = 0$, we obtain

$$\begin{aligned}\mathrm{Cov}(X,Z) &= E[(X-\mu_x)((Y-X)-(\mu_y-\mu_x))] \\ &= E[(X-\mu_x)(Y-\mu_y) - (X-\mu_x)^2] \\ &= \mathrm{Cov}(X,Y) - \mathrm{Var}(X) = -\sigma_x^2\end{aligned}$$

(b) We can calculate

$$\mathrm{Var}(Z) = \mathrm{Var}(Y-X) = \mathrm{Var}(Y) + \mathrm{Var}(-X) = \sigma_y^2 + \sigma_x^2$$

Then we obtain

$$\rho = \frac{\mathrm{Cov}(X,Z)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Z)}} = -\frac{\sigma_x}{\sqrt{\sigma_x^2+\sigma_y^2}}$$

**Problem 14.** By the property of expectation for two independent random variables, we obtain

$$\begin{aligned}M_Z(t) &= E[e^{t(\alpha X+\beta Y)}] \\ &= E[e^{(t\alpha)X}] \times E[e^{(t\beta)Y}] = M_X(\alpha t) \times M_Y(\beta t)\end{aligned}$$