

# **Section 3: Random variables – discrete**

STA 35C – Statistical Data Science III

**Instructor:** Akira Horiguchi

Fall Quarter 2025 (Sep 24 – Dec 12)  
MWF, 12:10 PM – 1:00 PM, Olson 158  
University of California, Davis

# Overview

Based on Chapter 3 of textbook: <https://www.probabilitycourse.com/>

- Contains problems with solutions, and problems without solutions.

- 1 Basic concepts
- 2 Independent random variables
- 3 Special distributions
- 4 Cumulative distribution function
- 5 Expected value
- 6 Functions of random variables
- 7 Variance

# **Basic concepts**

## Random variables

We usually focus on some *numerical aspects* of a random experiment.

- For example, in a soccer game we may be interested in the number of goals, shots, shots on goal, corners kicks, fouls, etc.
- On any given day at UCD, we may be interested in the number of Cheeto sightings.
- These are examples of *random variables*.

## Definition 1: Random variable

A *random variable*  $X: \Omega \rightarrow \mathbb{R}$  is a function from the sample space  $\Omega$  to the real numbers.

- E.g., toss a coin three times. Sample space is

$$\Omega = \{TTT, TTH, THT, THH, HTT, HTH, HHT, HHH\}.$$

We can define a random variable  $X$  whose value is the number of observed heads.

- Usually denote random variables by capital letters such as  $X$ ,  $Y$ , and  $Z$ .
- The *range* of a random variable  $X$  is the set of possible values for  $X$ . For example:
  - ▶ I toss a coin 100 times. Let  $X$  be the number of heads I observe.
  - ▶ I toss a coin until the first heads appears. Let  $Y$  be the total number of coin tosses.
  - ▶ The random variable  $T$  is defined as the time (in hours) from now until the next earthquake occurs in a certain city.

# Discrete random variables

A random variable  $X$  is *discrete* if its range is countable.

- Recurring examples:
  1. number of heads after two coin flips,
  2. number of coin flips needed before a heads turns up.
- Here probabilities can be assigned to each realizable value.
  1. For  $\{0, 1, 2\}$  (finite), we can assign probabilities  $1/4$ ,  $1/2$ , and  $1/4$ .
  2. For  $\mathbb{N}$  (countably infinite), we can assign probabilities  $(1/2)^k$  to each  $k \in \mathbb{N}$ .
- For a discrete r.v.  $X$  with range  $\{x_1, x_2, x_3, \dots\}$ , the function  $f_X(\cdot)$  defined as

$$f_X(x_k) = P(X = x_k), \quad \text{for } k = 1, 2, 3, \dots,$$

is called the *probability mass function (PMF)* of  $X$ .

1.  $f_X(0) = 1/4$ ,  $f_X(1) = 1/2$ , and  $f_X(2) = 1/4$ .
2.  $f_X(k) = (1/2)^k$  for each  $k \in \mathbb{N}$ .

Here  $f_X(a)$  is “*the probability that  $X$  equals  $a$ .*”

## Discrete random variables: PMF

PMF of the number of heads after two flips of a fair coin.

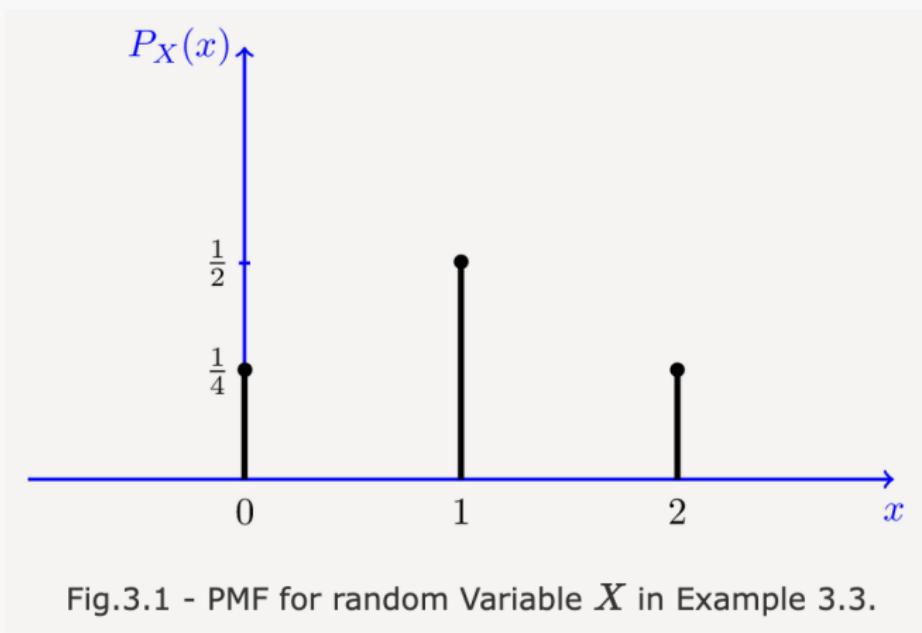


Fig.3.1 - PMF for random Variable  $X$  in Example 3.3.

The PMF of a discrete random variable is also called the r.v.'s *probability distribution*.

## Discrete random variables: PMF

A PMF is a probability measure, so it satisfies Definition 1 from Section 2.

- In particular, it satisfies countable additivity.
- This lets us deduce the probability  $P(X \in A)$  that a discrete r.v.  $X$  lies in an event  $A$ :

$$P(X \in A) = P\left(\bigcup_{a \in A} [X = a]\right) = \sum_{a \in A} f_X(a), \quad (1)$$

# **Independent random variables**

## Independent random variables

When dealing with more than one random variable, often need to consider the *dependence/correlation* between them.

- Concept of *independent random variables* is similar to that of independent events.
- Two random variables are independent if knowing the value of one does not change the probabilities for the other.

## Two independent random variables

### Definition 2: Two independent random variables

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

$$P(X = a, Y = b) = P(X = a) P(Y = b) \quad (2)$$

for all  $a, b$ .

If two random variables are independent, then we can write

$$\underbrace{P(X \in A, Y \in B)}_{\text{for all sets } A, B} = \underbrace{P(X \in A)}_{\text{for all sets } A} \underbrace{P(Y \in B)}_{\text{for all sets } B}$$

for all sets  $A, B$ . We can also write

$$P(Y = b | X = a) = P(Y = b)$$

for all  $a, b$ .

## Example

Toss a fair coin four times.

- Let  $X$  be the number of heads observed in the first and second coin flips.
- Let  $Y$  be the number of heads observed in the third and fourth coin flips.

Find  $P((X < 2) \text{ and } (Y > 1))$ ,

$$\begin{aligned} P(X < 2, Y > 1) &= P(X < 2) \cdot P(Y > 1) \\ &= P(X \in \{0, 1\}) \cdot P(Y \in \{2, 3\}) \\ &= \left[ 1 - P(X \in \{2, 3\}) \right] \underbrace{P(Y \in \{2, 3\})}_{\frac{1}{2} \cdot \frac{1}{2}} \\ &= \frac{3}{4} \cdot \frac{1}{4} = \boxed{\frac{3}{16}} \end{aligned}$$

## $\geq 2$ independent random variables

### Definition 3: $\geq 2$ independent random variables

Discrete random variables  $X_1, X_2, X_3, \dots, X_n$  are *independent* if

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = P(X_1 = x_1) P(X_2 = x_2) \cdots P(X_n = x_n) \quad (3)$$

for all  $x_1, x_2, \dots, x_n$ .

# **Special distributions**

## Uniform distribution

A random variable  $X$  with values in a finite set  $M$  is *uniformly* distributed if each element in  $M$  has the same probability:

$$P(X = k) = \frac{1}{|M|} \quad \text{for all } k \in M$$

- Such distributions occur when all possible outcomes are equally likely.
  - We write  $X \sim U(M)$  or  $X \sim \text{Unif}(M)$ .
  - Nine random draws in R: `sample(c(1,2,3,4,5,6), size=9, replace=T)`
- TRUE  
↓
- nine rolls of a fair six-sided die

## Bernoulli distribution

A random variable  $X$  is *Bernoulli* distributed with parameter  $p \in (0, 1)$ , if  $P(X = 1) = p$  and  $P(X = 0) = 1 - p$ .

- For when our random experiment has only two possible outcomes ("success" and "failure").
- Example: flip a coin with probability  $p$  of heads ("success"). Is it heads?
- We write  $X \sim Ber_p$  or  $X \sim Bern(p)$ .
- Nine random draws in R: `rbinom(n=9, size=1, prob=1/3)`

# Binomial distribution

A random variable  $X$  is **Binomial** distributed with parameters  $n \in \mathbb{N}$  and  $p \in (0, 1)$  if

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{for all } k = 0, \dots, n.$$

- We think of  $n$  as the number of experiments and  $p$  the success probability. In the above equation,  $k$  is the number of successes.
- For measuring the probability of the number of successes of  $n$  independent Bernoulli experiments with parameter  $p$ .
- Example: flip a coin  $n$  times, each flip with probability  $p$  of heads ("success"). How many heads?
- We write  $X \sim Bin_{n,p}$  or  $X \sim Bin(n, p)$ .
- A random draw in R: `rbinom(n=3, size=1, prob=0.25) |> sum()`

# **Cumulative distribution function**

The PMF is one way to describe the distribution of a discrete random variable.

- Pro: intuitive.
- Con: it cannot be defined for continuous random variables.

The *cumulative distribution function (CDF)* can characterize the distribution of *any kind* of random variable (discrete, continuous, mixed).

# Cumulative distribution function

The **CDF** of a random variable  $X$  is the function  $F_X: \mathbb{R} \rightarrow [0, 1]$  defined by

$$F_X(a) := P(X \leq a), \quad a \in \mathbb{R}. \quad (4)$$

This is “the probability that  $X$  is less than or equal to  $a$ .”

- Definition holds regardless of whether  $X$  is discrete, continuous, or mixed.
- In the discrete case – recall Eq. (1) – holds for any  $a \in \mathbb{R}$ ,

$$\begin{aligned} A &= (-\infty, a) \\ F_X(a) &= P(X \in A) = \sum_{s \leq a} f_X(s) \end{aligned}$$

$F_X(a) = \sum_{s \leq a} f_X(s).$

$F_X(a) = P(X \leq a) = P\left(\bigcup_{s \leq a} [x=s]\right)$

countable additivity.

■ For any  $a, b \in \mathbb{R}$  with  $b > a$  holds,

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

# Cumulative distribution function

From the definition of  $F_X$  in Eq. (4) come the following properties:

1.  $F_X$  is **right-continuous** and **monotonically increasing**,

2.  $\lim_{a \rightarrow -\infty} F_X(a) = 0$ ,

3.  $\lim_{a \rightarrow +\infty} F_X(a) = 1$ .

"non-decreasing"

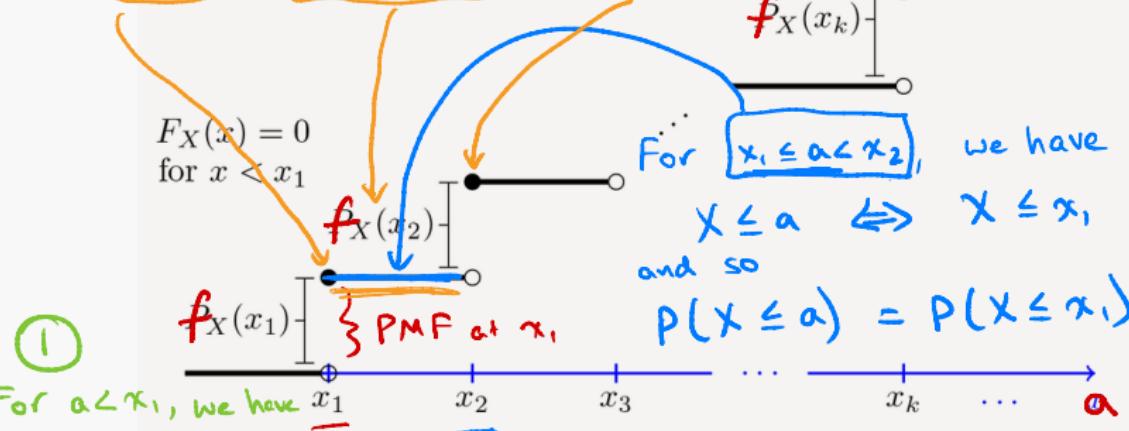
CDF at  $x_2$

(4)

$$P(X \leq x_1) + P(X = x_2) = P(X \leq x_2)$$

$$\lim_{x \rightarrow \infty} F_X(x) = 1$$

$$f_{X(x_k)}$$



$P(X \leq a) = 0$

(2)

$$P(X \leq x_1) = P(X = x_1) + P(X < x_1)$$

CDF at  $x_1$

PMF at  $x_1$

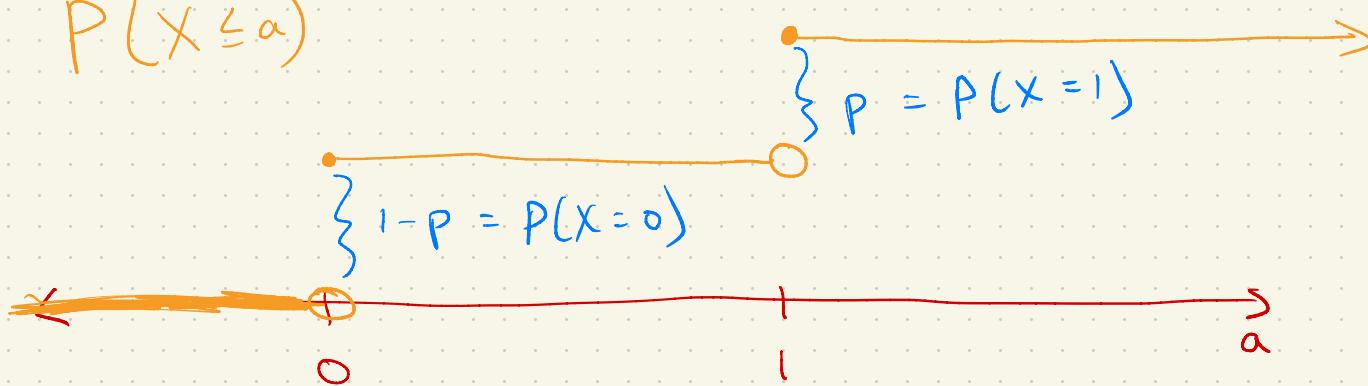
(16)

Fig. 3.4 - CDF of a discrete random variable.

## CDF of Bernoulli random variable with parameter $P$

- For  $a < 0$ , we have  $P(X \leq a) = 0$ .
- For  $\underline{a = 0}$ , we have  $P(X \leq a) = \frac{P(X=0)}{1-p} + \frac{P(X<0)}{0} = 1-p$
- For  $0 \leq a < 1$ , we have  $P(X \leq a) = P(X \leq 0) + P(0 < X \leq a)$
- For  $\underline{a = 1}$ , we have  $P(X \leq a) = \frac{P(X=1)}{p} + \frac{P(X<1)}{P(X \leq 0)} = 1$ .
- For  $\underline{a > 1}$ , we have  $P(X \leq a) = P(X \leq 1) = 1$ .

$$P(X \leq a)$$



## Example

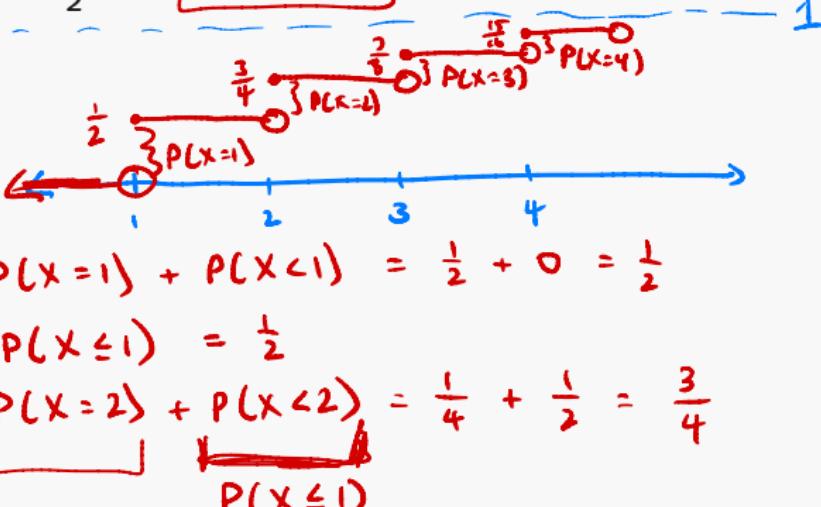
Suppose the PMF of a discrete random variable  $X$  is given by

$$P(X=k) = f_X(k) = \frac{1}{2^k} \quad \text{for } k = 1, 2, 3, \dots$$

$$F_X(k) = 1 - \frac{1}{2^k}$$

**CDF**

- Find and plot the CDF.
- Find  $P(2 < X \leq 5)$ .
- Find  $P(X > 4)$ .



$$\alpha < 1 : P(X \leq \alpha) = 0$$

$$\alpha = 1 : P(X \leq 1) = P(X=1) + P(X < 1) = \frac{1}{2} + 0 = \frac{1}{2}$$

$$1 < \alpha < 2 : P(X \leq \alpha) = P(X \leq 1) = \frac{1}{2}$$

$$\alpha = 2 : P(X \leq 2) = P(X=2) + P(X < 2) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

$\underbrace{\hspace{1cm}}$   $\underbrace{\hspace{1cm}}$

$P(X \leq 1)$

$$2 < \alpha < 3 : P(X \leq \alpha) = P(X \leq 2) + \underbrace{P(2 < X < \alpha)}_{0} = \frac{3}{4} + 0 = \frac{3}{4}$$

$$\alpha = 3 : P(X \leq 3) = P(X=3) + \underbrace{P(X < 3)}_{P(X \leq 2)} = \frac{1}{8} + \frac{3}{4} = \frac{7}{8}$$

$\underbrace{\hspace{1cm}}$

$\frac{1}{2^3}$

## Example

Suppose the PMF of a discrete random variable  $X$  is given by

$$f_X(k) = \frac{1}{2^k} \quad \text{for } k = 1, 2, 3, \dots$$

1. Find and plot the CDF.
2. Find  $P(2 < X \leq 5)$ .
3. Find  $P(X > 4)$ .

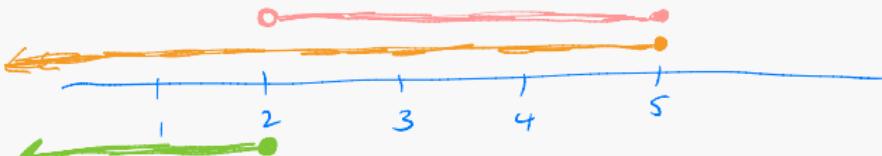
Closed form expression for CDF of  $X$  for  $k \in \mathbb{N}$ :

$$\begin{aligned} F_X(k) &= P(X \leq k) = 1 - P(X > k) \\ &= 1 - \sum_{a=k+1}^{\infty} P(X = a) \\ &= 1 - \sum_{a=k+1}^{\infty} \frac{1}{2^a} \\ &= \boxed{1 - \frac{1}{2^k}} \end{aligned}$$

## Example

Suppose the PMF of a discrete random variable  $X$  is given by

$$f_X(k) = \frac{1}{2^k} \quad \text{for } k = 1, 2, 3, \dots$$



1. Find and plot the CDF.

→ 2. Find  $P(2 < X \leq 5)$ .

→ 3. Find  $P(X > 4)$ .

$$\underline{P(2 < X \leq 5)} = \underline{P(X \leq 5)} - \underline{P(X \leq 2)}$$

$$= F_X(5) - F_X(2)$$

$$= \left[1 - \frac{1}{2^5}\right] - \left[1 - \frac{1}{2^2}\right] = \frac{1}{2^2} - \frac{1}{2^5}$$

---

$$\begin{aligned} P(X > 4) &= 1 - P(X \leq 4) \\ &= 1 - F_X(4) = 1 - \left[1 - \frac{1}{2^4}\right] = \frac{1}{2^4} \end{aligned}$$

# Expected value

# Introduction

Given some numbers, we often want a descriptive summary of these values.

- Their *average* is a single number that represents/describes the whole collection.
- How might we describe a “representative value” for a random variable?
- With a random variable, some values occur more often than other values.
- We might want to weight the values more if they occur more often.

Example: suppose we have a fair die. How might we summarize the outcomes of this die using a single number? What about for an unfair die?

## Definition

The *expected value* of a random variable is the weighted average of all of its *values*, where the *weights* are the probabilities that these values occur.

### Definition 2: Expected value $E(\cdot)$

Let  $X$  be a discrete random variable. Then the *expected value* of  $X$  is defined as

$$E[x] \quad E[X]$$

$$E(X) = \sum_{\text{all } k} P(X = k) \cdot k \quad (5)$$

Example: Let  $X \sim \text{Bernoulli}(p)$ . Find  $E(X)$ .

$$\begin{aligned} E[X] &= \sum_{k=0}^1 P(X=k) \cdot k \\ &= \underbrace{P(X=0) \cdot 0}_{0} + \underbrace{P(X=1) \cdot 1}_{P(X=1)} \\ &= P(X=1) \end{aligned}$$

End of 10/8  
lecture

If  $X$  is a random variable, then any function of  $X$  is also a random variable.

- For example, if  $Y = aX + b$ , we can talk about  $EY = E[aX + b]$ .

## Theorem 3.2: Expectation is linear

We have

- $E[aX + b] = aEX + b$ , for all  $a, b \in \mathbb{R}$ ;
- $E[X_1 + X_2 + \dots + X_n] = EX_1 + EX_2 + \dots + EX_n$ , for any set of random variables  $X_1, X_2, \dots, X_n$ .

Example: Let  $X \sim \text{Binomial}(n, p)$ . Find  $E(X)$ .

# **Functions of random variables**

# Introduction

If  $X$  is a r.v. and any *function  $Y = g(X)$  of  $X$*  is itself a random variable.

- Range of  $Y$  is

$$R_Y := \{g(a) | a \in R_X\}$$

where  $R_X$  is the range of  $X$ .

- PMF of  $Y$  is

$$f_Y(b) = P(Y = b) = P(g(X) = b) = \sum_{a: g(a)=b} f_X(a).$$

- Expected value of  $Y$  is

$$EY = \sum_{b \in R_Y} b f_Y(b).$$

In practice, usually easier to use the *law of the unconscious statistician (LOTUS)*:

$$EY = E[g(X)] = \sum_{a \in R_X} g(a) f_X(a).$$

## Example

Find  $E[\sin(X)]$ , where  $X$  is a discrete random variable with range

$$R_X = \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi\right\}$$

and PDF values

$$f_X(0) = f_X\left(\frac{\pi}{4}\right) = f_X\left(\frac{\pi}{2}\right) = f_X\left(\frac{3\pi}{4}\right) = f_X(\pi) = \frac{1}{5}.$$

# Variance

## Intuition

Often summarize a probability distribution by its center and spread.

- Center: expected value
- Spread: *variance*

Think of variance as “how much a random variable varies about its mean.”

## Definition

### Definition 3: Variance $\text{Var}(\cdot)$

Let  $X$  be a random variable with  $E(X^2) < \infty$ . Then the **variance** of  $X$  is defined as

$$\text{Var}(X) := E\left[\{X - E(X)\}^2\right]. \quad (6)$$

- A **large value** of  $\text{Var}(X)$  means that  $\{X - E(X)\}^2$  is often large, so  $X$  often takes values far from its mean.
  - ▶ Implies that the distribution is very spread out.
- A **small value** of  $\text{Var}(X)$  means that  $\{X - E(X)\}^2$  is often small, so  $X$  often takes values close to its mean.
  - ▶ Implies that the distribution is concentrated around its average.

## Standard deviation

$\text{Var}(X)$  has a different unit than  $X$ . E.g., if  $X$  is a stock price.

- Can instead measure spread using the square root of variance:

### Definition 4: Standard deviation $\text{Var}(\cdot)$

Let  $X$  be a r.v. with  $E(X^2) < \infty$ . Then the **standard deviation** of  $X$  is defined as

$$SD(X) := \sqrt{\text{Var}(X)} \tag{7}$$

- Despite having the same unit of  $X$ , the variance is easier to mathematically find the minimum of (i.e., take the derivative of).
- Usually we will describe a distribution's spread using the variance.

## Properties and calculation tools

From Definition 3, we can deduce the following properties:

- $\text{Var}(X) \geq 0$ .
- If  $\text{Var}(X) = 0$ , then  $X$  is constant.
- The variance of  $X$  can also be calculated as

$$\text{Var}(X) = E(X^2) - (E[X])^2. \quad (8)$$

### Properties of $\text{Var}(\cdot)$

Let  $c \in \mathbb{R}$  be a constant, and let  $X$  be a random variable with  $E(X^2) < \infty$ . Then

- $\text{Var}(c) = 0$ ;
- $\text{Var}(X + c) = \text{Var}(X)$ ;
- $\text{Var}(cX) = c^2 \text{Var}(X)$ ;

**Example:** consider  $c = 5$ ,  $\text{Var}(X) = 1$ .

## Properties and calculation tools

**Theorem:** variance of sum of independent random variables

If  $X_1, X_2, \dots, X_n$  are *independent* random variables, then

$$\text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n). \quad (9)$$

**Example:** if  $X \sim \text{Binomial}(n, p)$ , find  $\text{Var}(X)$ .