

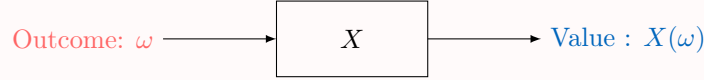
DISCRETE RANDOM VARIABLES

A RANDOM VARIABLES



A.1 DEFINITIONS

Definition Random Variable

A **random variable**, denoted X , is a function that assigns a numerical value to each outcome ω in a random experiment. We write this value as $X(\omega)$.



The **possible values** of X are the real numbers that X can take.

Ex: Let X be the number of heads when tossing 2 fair coins:  (red coin) and  (blue coin). Find $X(H, T)$.
Answer: The outcome (H, T) means the red coin shows heads (H) and the blue coin shows tails (T). Since X counts heads, there's 1 head. Thus, $X(H, T) = 1$.



Definition Discrete Random Variable

A random variable is **discrete** if its set of possible values is finite or countably infinite. This means we can list all possible values.

Definition Events Involving a Random Variable

For a random variable X :

- $(X = x)$: The set of outcomes where X takes the value x .
- $(X \leq x)$: The set of outcomes where X is less than or equal to x .
- $(X \geq x)$: The set of outcomes where X is greater than or equal to x .

Ex: Let X be the number of heads when tossing 2 coins:  and . List the outcomes for $(X = 0)$, $(X = 1)$, $(X = 2)$, $(X \leq 1)$, and $(X \geq 1)$.

Answer:

- $(X = 0) = \{(T, T)\}$ (no heads).
- $(X = 1) = \{(T, H), (H, T)\}$ (one head).
- $(X = 2) = \{(H, H)\}$ (two heads).
- $(X \leq 1) = (X = 0) \cup (X = 1) = \{(T, T), (T, H), (H, T)\}$ (at most one head).
- $(X \geq 1) = (X = 1) \cup (X = 2) = \{(T, H), (H, T), (H, H)\}$ (at least one head).

A.2 PROBABILITY DISTRIBUTION



Definition Probability Distribution

The **probability distribution** of a random variable X lists the probability $P(X = x_i)$ for each possible value x_1, x_2, \dots, x_n . It can be shown as a table or formula.

Proposition Characteristic of a Probability Distribution

For a random variable X with possible values x_1, x_2, \dots, x_n , we have

- $0 \leq P(X = x_i) \leq 1$ for all $i = 1, \dots, n$,
- $\sum_{i=1}^n P(X = x_i) = P(X = x_1) + P(X = x_2) + \dots + P(X = x_n) = 1$.

Ex: Let X be the number of heads when tossing 2 fair coins:  and .

1. List the possible values of X .
2. Find the probability distribution.
3. Create the probability table.
4. Draw the probability distribution graph.

Answer:

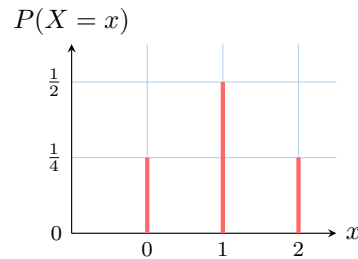
1. Possible values: 0 (no heads), 1 (one head), 2 (two heads).
2. Probability distribution:

- $P(X = 0) = P(\{(T, T)\}) = \frac{1}{4},$
- $P(X = 1) = P(\{(T, H), (H, T)\}) = \frac{2}{4} = \frac{1}{2},$
- $P(X = 2) = P(\{(H, H)\}) = \frac{1}{4}.$

3. Probability table:

x	0	1	2
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

4. Graph:



A.3 EXISTENCE OF A RANDOM VARIABLE WITH A GIVEN PROBABILITY DISTRIBUTION

Usually, defining a random variable begins by establishing:

1. a sample space, that is, the set of all possible outcomes,
2. a probability associated with this sample space,
3. a function X that assigns a number to each outcome in the sample space.

This is quite a lengthy task. However, often, we prefer to directly define a random variable X with a given probability distribution, relying on the context of the situation being studied. For example, imagine we survey a class of 30 students about their siblings and obtain these results: 10 students have 0 siblings, 12 have 1 sibling, 5 have 2 siblings, and 3 have 3 siblings. We can then define the random variable X as the number of siblings of a randomly chosen student, with this probability distribution:

x	0	1	2	3
$P(X = x)$	$\frac{10}{30}$	$\frac{12}{30}$	$\frac{5}{30}$	$\frac{3}{30}$

The theorem below shows that it is always possible to construct a sample space, a probability, and a function X to obtain a random variable with this probability distribution.

Theorem Existence of a Random Variable with a Given Probability Distribution

Suppose you have possible values x_1, x_2, \dots, x_n and probabilities p_1, p_2, \dots, p_n .
If:

- $0 \leq p_i \leq 1$ for each $i = 1, 2, \dots, n,$
- $\sum_{i=1}^n p_i = p_1 + p_2 + \dots + p_n = 1,$

then there exists a random variable X with the probability distribution $P(X = x_i) = p_i$ for each $i = 1, 2, \dots, n.$

Method Defining a Random Variable X with a Valid Probability Distribution

In practice, we often define a random variable X directly by specifying its probability distribution. The key is to ensure that this distribution is valid, meaning it satisfies the conditions for a probability distribution: all probabilities must be non-negative and sum to 1.

Ex: We survey a class of 30 students about their siblings and obtain these results: 10 students have 0 siblings, 12 have 1 sibling, 5 have 2 siblings, and 3 have 3 siblings. We define a random variable X as the number of siblings of a randomly chosen student, with this probability distribution:

x	0	1	2	3
$P(X = x)$	$\frac{10}{30}$	$\frac{12}{30}$	$\frac{5}{30}$	$\frac{3}{30}$

Determine if this probability distribution is valid.

Answer:

- $P(X = x) \geq 0$ for all $x = 0, 1, 2, 3$ (true: $\frac{10}{30}$, $\frac{12}{30}$, $\frac{5}{30}$, and $\frac{3}{30}$ are all non-negative),
- $P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) = \frac{10}{30} + \frac{12}{30} + \frac{5}{30} + \frac{3}{30} = \frac{30}{30} = 1$ (true: the sum equals 1).

Since both conditions are satisfied, the probability distribution is valid.

B MEASURES OF CENTER AND SPREAD

B.1 EXPECTATION

The **expected value** of a random variable X is the "average you'd expect if you repeated the experiment many times". It's found by taking all possible values, multiplying each by its probability, and adding them up — essentially a weighted average where the probabilities act as the weights.

Definition Expected Value

For a random variable X with possible values x_1, x_2, \dots, x_n , the **expected value**, $E(X)$, also called the **mean**, is:

$$\begin{aligned} E(X) &= \sum_{i=1}^n x_i P(X = x_i) \\ &= x_1 P(X = x_1) + x_2 P(X = x_2) + \dots + x_n P(X = x_n) \end{aligned}$$

Ex: You toss 2 fair coins, and X is the number of heads. The probability distribution is:

x	0	1	2
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Find the expected value of X .

Answer: Calculate $E(X)$ using the formula:

$$\begin{aligned} E(X) &= 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} \\ &= \frac{1}{2} + \frac{2}{4} \\ &= 1 \end{aligned}$$

So, on average, you expect 1 head when tossing 2 coins.

Proposition Linearity of Expectation

For any random variable X and constants a and b , the expected value of a linear transformation of X is:

$$E(aX + b) = aE(X) + b$$

This property is derived from two simpler rules:

- $E(aX) = aE(X)$ (The expectation of a scaled variable is the scaled expectation).
- $E(X + b) = E(X) + b$ (The expectation of a shifted variable is the shifted expectation).

B.2 VARIANCE AND STANDARD DEVIATION

The **variance** measures how spread out the values of a random variable are from its expected value. The **standard deviation** is the square root of the variance, giving a sense of typical deviation in the same units as X .

Definition Variance and Standard Deviation

The **variance**, denoted $V(X)$, is:

$$\begin{aligned} V(X) &= \sum_{i=1}^n (x_i - E(X))^2 P(X = x_i) \\ &= (x_1 - E(X))^2 P(X = x_1) + (x_2 - E(X))^2 P(X = x_2) + \cdots + (x_n - E(X))^2 P(X = x_n) \end{aligned}$$

The **standard deviation**, denoted $\sigma(X)$, is $\sigma(X) = \sqrt{V(X)}$.

Ex: You toss 2 fair coins, and X is the number of heads. The probability table is:

x	0	1	2
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Given $E(X) = 1$, find the variance.

Answer: Calculate $V(X)$:

$$\begin{aligned} V(X) &= (0 - 1)^2 \times \frac{1}{4} + (1 - 1)^2 \times \frac{1}{2} + (2 - 1)^2 \times \frac{1}{4} \\ &= 1 \times \frac{1}{4} + 0 \times \frac{1}{2} + 1 \times \frac{1}{4} \\ &= \frac{1}{4} + 0 + \frac{1}{4} \\ &= \frac{1}{2} \end{aligned}$$

The variance is $\frac{1}{2}$.

Proposition Computational Formula for Variance

A more convenient formula for computation is:

$$V(X) = E(X^2) - [E(X)]^2$$


C CLASSICAL DISTRIBUTIONS

C.1 UNIFORM DISTRIBUTION

Definition Uniform Distribution

A random variable X follows a **uniform distribution** if each possible value has the same probability:

$$P(X = x) = \frac{1}{\text{Number of possible values}}, \quad \text{for any possible value } x$$

Ex: Let X be the result of rolling a fair die: .

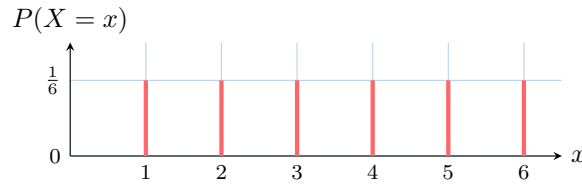
1. List the possible values of X .
2. Create the probability table.
3. Draw the probability distribution graph.

Answer:

1. Possible values: 1, 2, 3, 4, 5, 6.
2. Probability table:

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

3. Graph:



Proposition Expectation and Variance of a Uniform Distribution

For a random variable X that follows a uniform distribution on the set of integers $\{1, 2, \dots, n\}$:

- The expected value is $E(X) = \frac{n+1}{2}$.
- The variance is $V(X) = \frac{n^2-1}{12}$.

Ex: Let X be the random variable for the score on a roll of a fair six-sided die. Find the mean and variance of X .

Answer: The random variable X follows an uniform distribution on $\{1, 2, 3, 4, 5, 6\}$.

- $E(X) = \frac{6+1}{2} = 3.5$
- $V(X) = \frac{6^2-1}{12} = \frac{35}{12} \approx 2.92$

C.2 BERNOULLI DISTRIBUTION

A **Bernoulli distribution** models an experiment with two outcomes: success (1) or failure (0), like flipping a coin where heads is 1 and tails is 0. The probability of success is p .

Definition Bernoulli Distribution

A random variable X follows a **Bernoulli distribution** if:

- Possible values are 0 and 1.
- $P(X = 1) = p$ and $P(X = 0) = 1 - p$.

We write $X \sim B(p)$.

Ex: A basketball player has an 80% chance of making a free throw. Let $X = 1$ if the shot is made, and $X = 0$ if it's missed.

1. Is X a Bernoulli random variable?
2. Find the probability of success.

Answer:

1. Yes, X has values 0 or 1, so it follows a Bernoulli distribution.
2. Probability of success: $P(X = 1) = 80\% = 0.8$.

Proposition Expectation and Variance of a Bernoulli Distribution

For a Bernoulli random variable X with a probability of success p , the following hold:

- The expected value is $E(X) = p$,
- The variance is $V(X) = p(1 - p)$,
- The standard deviation is $\sigma(X) = \sqrt{p(1 - p)}$.

C.3 BINOMIAL DISTRIBUTION

Suppose a basketball player takes n free throws, and we count the number of shots made. The probability of making a free throw is the same for each attempt, and each shot is independent of every other shot. This is an example of a binomial experiment.

Definition Binomial Experiment

A **binomial experiment** is a statistical experiment that consists of a sequence of repeated Bernoulli trials. It must satisfy the following four conditions:

1. **Fixed Number of Trials:** The experiment consists of a fixed number of trials, denoted by n .
2. **Independent Trials:** The outcome of each trial is independent of the outcomes of all other trials.
3. **Two Outcomes:** Each trial has only two possible outcomes, typically labeled "success" and "failure".
4. **Constant Probability:** The probability of success, denoted by p , is the same for each trial. The probability of failure is $1 - p$.

A random variable X that counts the number of successes in a binomial experiment is called a **binomial random variable**.

Proposition Distribution of a Binomial Random Variable

Let X be a binomial random variable with n independent trials and a probability of success p . The probability distribution of X is:

$$P(X = x) = \underbrace{\binom{n}{x}}_{\substack{\text{number of ways } x \text{ successes} \\ \text{can be ordered among the} \\ n \text{ trials}}} \underbrace{p^x(1-p)^{n-x}}_{\substack{\text{probability of} \\ \text{obtaining } x \text{ successes} \\ \text{and } n - x \text{ failures in a} \\ \text{particular order}}} \quad \text{where } x = 0, 1, 2, \dots, n.$$

This is called the **binomial distribution**, and we write $X \sim B(n, p)$.

Ex: A basketball player has an 80% chance of making a free throw and takes 5 shots. Let X be the number of shots made.

1. Is X a binomial random variable?
2. Find the probability of making 4 shots.

Answer:

1. Yes, X is a binomial random variable because it counts the number of successes (shots made) in 5 independent trials (free throws), each with a constant success probability of 0.8.
2. As $X \sim B(5, 0.8)$,

$$\begin{aligned} P(X = 4) &= \binom{5}{4} (0.8)^4 (1 - 0.8)^1 \\ &= 5 \times 0.4096 \times 0.2 \\ &= 0.4096 \end{aligned}$$

The probability of making 4 shots is 0.4096.

Proposition Expectation and Variance of a Binomial Random Variable

For $X \sim B(n, p)$:

- $E(X) = np$ (expected value),
- $V(X) = np(1 - p)$ (variance),
- $\sigma(X) = \sqrt{np(1 - p)}$ (standard deviation).

Ex: A basketball player has an 80% chance of making a free throw and takes 5 shots. Find the mean and standard deviation of the number of successful shots.

Answer: Let X be the number of successful shots. Since each shot is independent and has a success probability of 0.8, we have $X \sim B(5, 0.8)$.

$$E(X) = 5 \times 0.8 = 4,$$

$$V(X) = 5 \times 0.8 \times (1 - 0.8) = 5 \times 0.8 \times 0.2 = 0.8,$$

$$\sigma(X) = \sqrt{0.8} \approx 0.89.$$

Mean is 4 successful shots, standard deviation is about 0.89.