1. Discrete Random Variables, PMF and CDF

average.

2. Expected Values and Variance deviation from average.

3. Discrete Distribution: Bernoulli, Binomial and Geometric

Discrete Random Variables, PMF and CDF

Random Variable

Motivating example: A fair dice is rolled 4 times.

$$\Omega = \text{set of all 4-tuples } (x_1, x_2, x_3, x_4) \text{ with } x_i \in \{1, 2, 3, 4, 5, 6\}.$$

Consider the following functions X and Y:

- X = sum of the rolls. E.g. $X((1, 2, 5, 6)) = 1 + 2 + 5 + 6 = 14$.
- Y = maximum among the four numbers. E.g. $Y((1, 2, 5, 6)) = 6$.

These functions are called random variables on Ω.

Calculus:

$$f(x) = x^2$$

$$f : x \longmapsto x^2$$

Probability:

$$X : \Omega \longrightarrow \mathbb{R}$$

$$X : \omega \longmapsto X(\omega) \text{ is a number.}$$

A **random variable** on Ω is a function X that assigns a real number $X(\omega)$ to every outcome ω.

Random variables provide an efficient and intuitive way to specify events.

E.g. Using the random variable X above,

$$X = 5 \iff E = \underbrace{\{(1,1,1,2),(1,1,2,1),(1,2,1,1),(2,1,1,1)\}} \subseteq \Omega$$

$$\mathbb{P}(X=5) = \mathbb{P}(E) = \frac{|E|}{|\Omega|}.$$

Example 1

A fair coin is tossed three times.

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

Consider the random varibles X and Y defined by

- $X =$ number of heads that occur
- $Y =$ number of tails that occur

- $\mathbb{P}(X = 3) = \mathbb{P}(\{HHH\}) = \frac{1}{8}$.
- $\mathbb{P}(X \leq 1) = \mathbb{P}(\{HTT, THT, TTH, TTT\}) = \frac{4}{8}$.
- $\mathbb{P}(X \in \{0, 3\}) = \mathbb{P}(\{HHH, TTT\}) = \frac{2}{8}$.
- $\mathbb{P}(X > Y) = \mathbb{P}(\{HHH, HHT, HTH, THH\}) = \frac{4}{8}$.

A **discrete random variable** is a random variable whose set of possible values is finite or countably infinite.

enumerate possibles
one by one infinitely.

A dice is thrown repeatedly.

Consider the following random variables

- X: number of 6's among the first 10 throws
- Y: number of throws until the first 6 is thrown

Set of possible values of X: $\{0, 1, 2, \ldots, 10\}$ (finite set)

Set of possible values of Y: $\{1, 2, \ldots\}$ (countably infinite set)

$Y = 1 \Leftrightarrow$ get "6" at 1st throw.

$Y = 2 \Leftrightarrow$ " " " 2nd throw but not 1st throw.

PMF

Let X be a discrete random variable. The **probability mass function (PMF)** of X is defined as

$$p_X(x) = \mathbb{P}(X = x)$$

x is a possible value

for all real numbers x.

Note:

- $p_X(x) = 0 \Leftrightarrow x$ is not a possible value of X.
- the PMF of X uniquely determines the probabilities of all events involving X.
- sometimes we just write $p(x)$ instead of $p_X(x)$.

$$|\Omega| = 2 \times 2 \times 2 = 2^3 = 8$$

Example 2

A fair coin is tossed 3 times. Let X = number of heads that occur. The PMF is given by

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Eg $x=2$ have 2 heads.

$$E = \{HHT, HTH, THH\}$$

$$\mathbb{P}(X=2) = \frac{3}{8}$$

If X is a discrete random variable with PMF $p(x)$, then the **Cumulative Density Function (CDF)** of X is defined by

$$F(x) = \mathbb{P}(X \le x) = \sum_{t \le x} p(t), \quad -\infty < x < \infty$$

where the sum runs over all numbers $t \le x$.

Example 3

A fair coin is tossed 3 times. Let X = number of heads that occur.

x	0	1	2	3
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

PMF

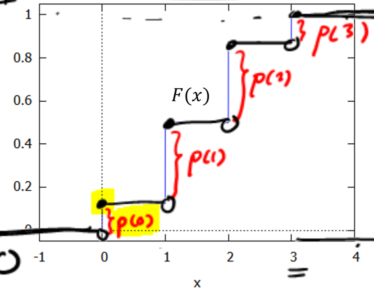
Let F be the CDF of X. Then

$F(c) = \sum_{t \le c} p(t) = 0$

$\boxed{c < 0}$ $F(-0.01) = 0$

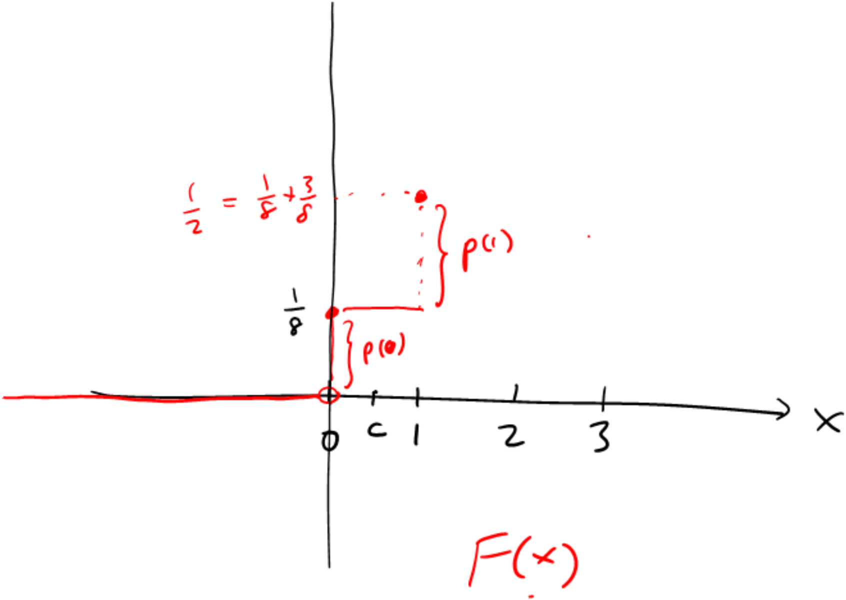
- $F(-1) = \sum_{t \le -1} p(t) = 0.$ ✓
- $F(0) = \sum_{t \le 0} p(t) = p(0) = \frac{1}{8}.$ ✓
- $F(1) = \sum_{t \le 1} p(t) = p(0) + p(1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}.$ ✓
- $F(2) = \sum_{t \le 2} p(t) = p(0) + p(1) + p(2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}.$ ✓
- $F(3) = \sum_{t \le 3} p(t) = p(0) + p(1) + p(2) + p(3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$ ✓

$$F(x) = \begin{cases} 0, x < 0 \\ 1/8, 0 \le x < 1 \\ 1/2, 1 \le x < 2 \\ 7/8, 2 \le x < 3 \\ 1, x \ge 3. \end{cases}$$



$p(0)+p(1)+p(2)+p(3)=1$

Note: At every x with $p(x) > 0$ there is a jump by $p(x)$.



1/2 = 1/8 + 3/8
p(1)
1/8
p(0)
0
c
1
2
3
x
F(x)

The CDF has the following properties:

- $F(x)$ is a non-decreasing function of x, for $-\infty < x < \infty$.
- $F(x)$ ranges from 0 to 1.
- If a is the minimum possible value of X, then $F(a) = p_X(a)$. If $c < a$ then $F(c) = 0$.
- If b is the maximum possible value of X, then $F(b) = 1$.
- Also called the distribution function.

weighted average.

Expected Values and Variance

If a fair coin is tossed 1000 times, we expect around 500 heads.
If a dice is rolled 6000 times, around 1000 sixes are expected.

Both statements can be expressed in terms of random variables:

- Let X be the number of heads among 1000 tosses. Then $\mathbb{E}[X] = 500$ (expected value of X)
- Let Y be the number of sixes among 6000 throws. Then $\mathbb{E}[Y] = 1000$

The definition of expected values formalizes this.

Expected Value of Randon Variable

The **expected value** (or **mean**) of a discrete random variable X with PMF $p(x)$ is

$$\mathbb{E}[X] = \sum_x xp(x)$$

where the sum runs over all numbers x with $p(x) > 0$.

Intuitive interpretation: $\mathbb{E}[X]$ is the sum of all possible values of X, weighted by their probabilities.

Remark: If c is a constant, then $\mathbb{E}[c] = c$.

Example 4

A fair coin is tossed 3 times. Let X = number of heads that occur.

x	0	1	2	3
$p(x)$	$\boxed{\frac{1}{8}}$	$\boxed{\frac{3}{8}}$	$\boxed{\frac{3}{8}}$	$\boxed{\frac{1}{8}}$

$$\mathbb{E}[X] = \sum_x xp(x) = 0 \cdot \boxed{\frac{1}{8}} + 1 \cdot \boxed{\frac{3}{8}} + 2 \cdot \boxed{\frac{3}{8}} + 3 \cdot \boxed{\frac{1}{8}} = \frac{12}{8} = \frac{3}{2}.$$

$\square$

Example 5

Two balls are randomly selected without replacement from an urn containing 5 balls numbered 1 through 5. Let X denote the larger number among the two balls selected. Find $\mathbb{E}[X]$.

$$\Omega = \{ \text{possible } \underline{\text{2-combinations}} \text{ of } \{1,2,3,4,5\} \}$$

$$|\Omega| = \binom{5}{2} = 10.$$

Ω: $\{1,2\}$ ✓, $\{1,4\}$, $\{3,4\}$, $\{2,5\}$. ....

$$X(\{1,2\}) = 2 \checkmark$$
$$X(\{3,4\}) = 4 \checkmark$$
$$X(\{1,4\}) = 4 \checkmark$$

etc.

PMF

x	2	3	4	5
$p(x)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$

✓

$1, 2, \ldots, x-1$

$$\mathbb{P}(X=x) = \frac{|\{\ldots, x\}|}{10} = \frac{x-1}{10}$$

$$\mathbb{E}[X] = \sum x p(x) = 2 \cdot \frac{1}{10} + 3 \cdot \frac{2}{10} + 4 \cdot \frac{3}{10} + 5 \cdot \frac{4}{10} = 4 \ \#.$$

Expected Value of Function of Random Variable

Let X be a discrete random variable with PMF $p(x)$, and $g(X)$ be a function of X (e.g $g(X) = X^2$, $g(X) = e^X$ etc.) Then

$$\mathbb{E}[g(X)] = \sum_x g(x)p(x).$$

If $g(x) = X$ then $\mathbb{E}[g(x)] = \mathbb{E}[X]$

Example 6

A fair coin is tossed 3 times. Let X = number of heads that occur.

x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$g(X) = X^2$$

$$\mathbb{E}[X^2] = \sum_x x^2 p_X(x) = 0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = \frac{24}{8} = 3.$$

$\mathbb{E}[X] = \frac{3}{2}$.

$\mathbb{E}[X^2] \neq (\mathbb{E}[X])^2$

Linearity of Expected Values

Theorem 7 (Linearity of Expected Values)

Let $X_1, \ldots, X_n$ be random variables such that $\mathbb{E}[X_i]$ exists for all $i = 1, \ldots, n$. Let $a_1, \ldots, a_n$ be real numbers (constants). Then

$$\mathbb{E}[a_1 X_1 + \cdots a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n].$$

Rules:

- constants can be pulled out of expected values
- expected value of a sum is the sum of expected values of the summands

Example 8

Suppose X, Y, Z are random variables with

$$\mathbb{E}[X] = -10, \quad \mathbb{E}[Y] = 20, \quad \mathbb{E}[Z] = 5000.$$

Then

$$\begin{aligned} \mathbb{E}[3X - 2Y + 5Z] &= 3\mathbb{E}[X] - 2\mathbb{E}[Y] + 5\mathbb{E}[Z] \\ &= 3(-10) - 2(20) + 5(5000) \\ &= 24930. \end{aligned}$$

□

Example 9

A firm purchases X number of computers each year, where X has the following probability distributions:

x	0	1	2	3
$f(x)$	$1/10$	$3/10$	$2/5$	$1/5$

If the cost of the computer is 1200 per unit and at the end of this year a rebate of $50X^2$ dollars will be issued, how much can this firm expect to spend on new computers during this year?

Give it a try!

$$\text{net Cost} = \underbrace{1200X - 50X^2}_{g(X)}$$

$$\mathbb{E}[1200X - 50X^2]$$

$$= 1200\,\underline{\underline{\mathbb{E}[X]}} - 50\,\underline{\underline{\mathbb{E}[X^2]}}$$

$$= 1200\left(0 \cdot \frac{1}{10} + 1 \cdot \frac{3}{10} + 2 \cdot \frac{2}{5} + 3 \cdot \frac{1}{5}\right)$$

$$- 50\left(0^2 \cdot \frac{1}{10} + 1^2 \cdot \frac{3}{10} + 2^2 \cdot \frac{2}{5} + 3^2 \cdot \frac{1}{5}\right)$$

#.

$\mathbb{E}[g(X)]$.

Variance

The **variance** of a random variable X is defined as

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

x	0	1
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

x	0	0.5	1
$p(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Interpretation:

- $X - \mathbb{E}[X]$: deviation of X from its expected value
- $(X - \mathbb{E}[X])^2$: measures (squared) deviation from expected value
- $\mathbb{E}[(X - \mathbb{E}[X])^2]$ measure average (squared) deviation of X from its expected value. So variance measures how 'spread out' X is from its mean.

The **standard deviation** of a random variable X is defined as

$$\sigma_X = \sqrt{\mathrm{Var}(X)}.$$

Theorem 10 (Formula for Variance)

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

Proof. Write $\underline{\mu} = \underline{\mathbb{E}[X]}$. Note that μ is a constant.

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}[(X-\mu)^2] \\
&= \mathbb{E}[X^2 - 2\mu X + \mu^2] \\
&= \mathbb{E}[X^2] - 2\mu\underline{\mathbb{E}[X]} + \mathbb{E}[\mu^2] \quad \text{(by linearity of expected values)} \\
&= \mathbb{E}[X^2] - 2\mu^2 + \mu^2 \quad \text{(expected value of constant)} \\
&= \mathbb{E}[X^2] - \mu^2 \\
&= \underline{\mathbb{E}[X^2]} - (\mathbb{E}[X])^2.
\end{aligned}$$

□

Discrete Distribution: Bernoulli, Binomial and Geometric

Bernoulli distribution

We say that a random variable X has a **Bernoulli distribution**, denoted by $X \sim \mathit{Bernoulli}(p)$ if X only takes value 0 (failure) and 1 (success) with $\mathbb{P}(X = 1) = p$. That is, its PMF is given by

x	0	1
$p(x)$	$1-p$	p

$$p(x) = \begin{cases} 1 - p & \text{if } x = 0 \\ p & \text{if } x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 11 (Bernoulli)

If $X \sim \mathit{Bernoulli}(p)$, then

$$\mathbb{E}[X] = p, \quad \mathrm{Var}[X] = p(1 - p).$$

It follows that the standard deviation of X is $\sqrt{p(1-p)}$.

$$\mathbb{E}[X] = 0 \cdot (1-p) + 1 \cdot p$$

$$= p$$

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

$$= \left(\underline{0^2 (1-p)} + \underline{1^2 \cdot p} \right) - p^2$$

$$= p - p^2$$

$$= p(1-p).$$

Some applications of Bernoulli distribution

- Experiments with only two outcomes, e.g. $X = 1$ if coin toss is head and $X = 0$ for tail ✓
- Yes-no-questions, e.g., $X = 1$ if person voted for candidate A and $X = 0$ otherwise ✓
- True-false conditions, e.g., $X = 1$ if total of 4 dice rolls is ≥ 20 and $X = 0$ otherwise

$Y \geqslant 20.$

$$X = \begin{cases} 1 & \text{if } Y \geqslant 20 \\ 0 & \text{otherwise.} \end{cases}$$

Binomial distribution

A random variable X has a **Binominal distribution**, denoted by $X \sim \mathit{Binomial}(n, p)$ if X is a sum of n independent Bernoulli random variables $\mathit{Bernoulli}(p)$.

$X = \sum_{i=1}^{n} X_i$ $\qquad X_i \sim \mathrm{Bernoulli}(p)$

$\mathbb{E}[X_i] = p$.

Interpretation: X = number of successes among n independent experiments with success probability p.

Theorem 12 (Binomial dsitribution)

If $X \sim \mathit{Binomial}(n, p)$, then

probability have x number of successes (1) in n of the Bernoulli experiment.

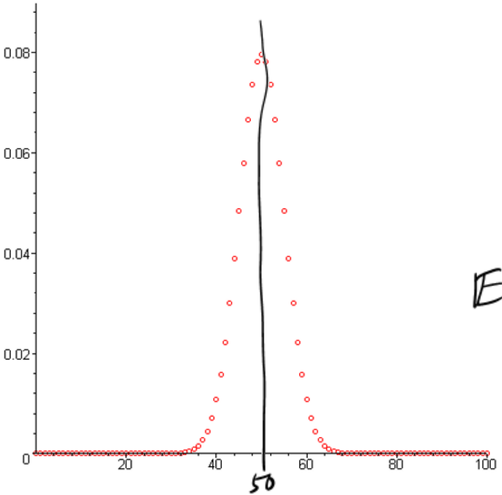
$$\text{PMF:}\quad p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n.$$

$$\mathbb{E}[X] = np, \quad \mathrm{Var}[X] = np(1-p).$$

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

$$= \sum_{i=1}^{n} p$$

$$= np.$$



PMF of Binomial(100,0.5)

n p

$$\mathbb{E}[X] = np$$
$$= 100 \times 0.5$$
$$= 50.$$

$$X = \sum_{i=1}^{10} X_i \qquad X_i \sim \text{Bernoulli}\left(\frac{1}{6}\right)$$

$$X_i = \begin{cases} 1 & \text{if 6 occurs} \\ 0 & \text{otherwise} \end{cases}$$

Example 13

A dice is rolled 10 times. let X be the number of 6's rolled. Then

$$X \sim Binomial\left(10, \frac{1}{6}\right)$$

$$\mathbb{P}(X = 2) = p(2) = \underbrace{\binom{10}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^8}_{\binom{n}{x} p^x (1-p)^{n-x}} \approx 0.29.$$

Example 14

In a production line, 10% of the items produced are defective. In a particular test five items are independently selected from the production line and are tested. Let X denote the number of defective items among the five items.

(i) Find the expected value and variance of X.

~~(ii)~~ What is probability that at most one item is defective?

Let $X_i = \begin{cases} 1 & \text{if } i\text{th item defective.} \\ 0 & \text{otherwise.} \end{cases}$

$X = \sum_{i=1}^{5} X_i \sim \text{Binomial}(5, 0.1)$

$$\mathbb{E}[X] = np$$
$$= 5 \times 0.1$$
$$= 0.5$$

$$\mathrm{Var}[X] = np(1-p)$$
$$= 5 \times 0.1 \times 0.9$$
$$= 0.45.$$

$$\mathbb{P}(X \leqslant 1) = \mathbb{P}(X=0) + \mathbb{P}(X=1)$$

$$= \binom{5}{0}(0.1)^0\, 0.9^5 + \binom{5}{1} 0.1^1\, 0.9^4$$

$$= 0.9185.$$

A random variable X has a **Geometric distribution**, denoted by $X \sim Geom(p)$, if X counts the number of experiments until the first success in a sequence of independent experiments with success probability p.

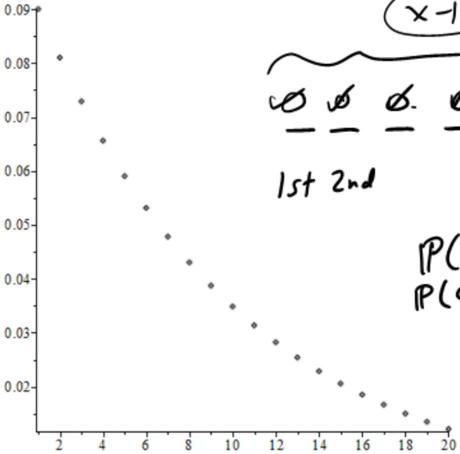
Bernoulli (p)

Theorem 15 (Geometric distribution)

If $X \sim Geom(p)$, then

$$\text{PMF:} \quad p(x) = \boxed{(1-p)^{x-1}} p, \quad x = 1, 2, \ldots,$$

$$\mathbb{E}[X] = \frac{1}{p}, \quad \text{Var}[X] = \frac{1-p}{p^2}.$$



$$\underbrace{0 \quad 0 \quad 0 \quad 0 \quad 0}_{x-1} \quad 1$$

1st 2nd ... xth.

$\mathbb{P}(1) = p$
$\mathbb{P}(0) = 1 - p.$
1 = success.
0 = failure.

PMF of Geom(0.1)

Example 16

A fair dice is rolled repeatedly. What is the probability that the 5th roll is the first roll for which a 1 or 6 occurs?

$$\text{Bernoulli}(p) \qquad p = \text{prob of getting } 1 \text{ or } 6$$

$$= \frac{2}{6} = \frac{1}{3}$$

$X = \#$ roll until we get 1 or 6.

$\sim \text{Geom}(p) \qquad p = \frac{1}{3}.$

$$\mathbb{P}(X=5) = (1-p)^4 p = \left(1-\frac{1}{3}\right)^4 \left(\frac{1}{3}\right) \approx 0.066$$