香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

DDA2001: Introduction to Data Science

Lecture 4: Elementary Probability Theory (continued)

Zicheng Wang

Recap: 1 - Definition of Probability

What is Probability?

- An experiment that can result in different outcomes, even though it is repeated in the same manner every time, is called a **random experiment**.

- Probability is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.

What is Probability?

- Random Experiment:

- Consider one possible outcome： ω
- The outcome ω happens with probability $P(\omega)$

- **It means:**
  - If we repeat such experiment **N** times
  - We observe **n** observations that the outcome is ω.
  - Then if N goes to infinity, $\boldsymbol{n}/\mathbf{N}$ will approach $P(\omega)$.

Terminologies

- **Random Experiment**: a repeatable procedure
- **Sample space**: set of all possible outcomes Ω.
- **Event**: a subset of the sample space.
- **Probability function**, $P(\omega)$: gives the probability for each outcome $\omega \in \Omega$
  - Probability is between 0 and 1
  - Total probability of all possible outcomes is 1.
  - If $A=\{\omega_1, \omega_2, \omega_3, \ldots\}$, $P(A)=P(\omega_1)+P(\omega_2)+P(\omega_3)+\cdots$

Sample space - examples

- Consider the random experiment in which items are selected from a batch consisting of three items {a,b,c}

- Case 1: select two items **without replacement**

- Case 2: select two items **with replacement**

Sample space – examples

- Consider the random experiment in which items are selected from a batch consisting of three items {a,b,c}

- Case 1: select two items **without replacement**

  Sample space $\{ab, ac, ba, bc, ca, cb\}$

- Case 2: select two items **with replacement**

Sample space - examples

- Consider the random experiment in which items are selected from a batch consisting of three items {a,b,c}

- Case 1: select two items **without replacement**

  Sample space $\{ab, ac, ba, bc, ca, cb\}$

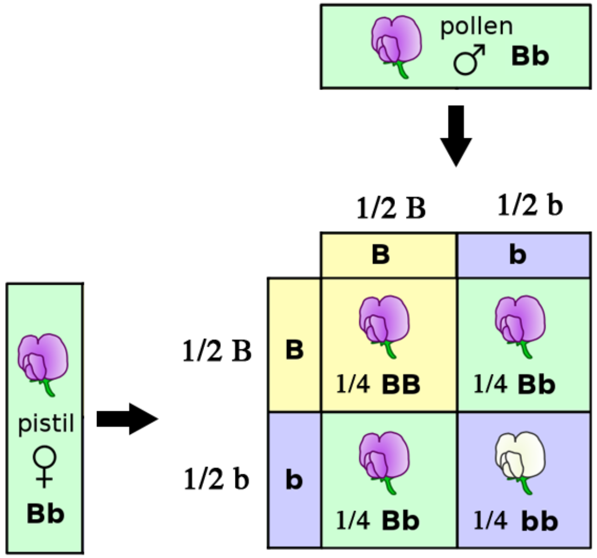
- Case 2: select two items **with replacement**

  Sample space $\{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$

Sample space – examples

Recall the heredity example in Lecture 2.

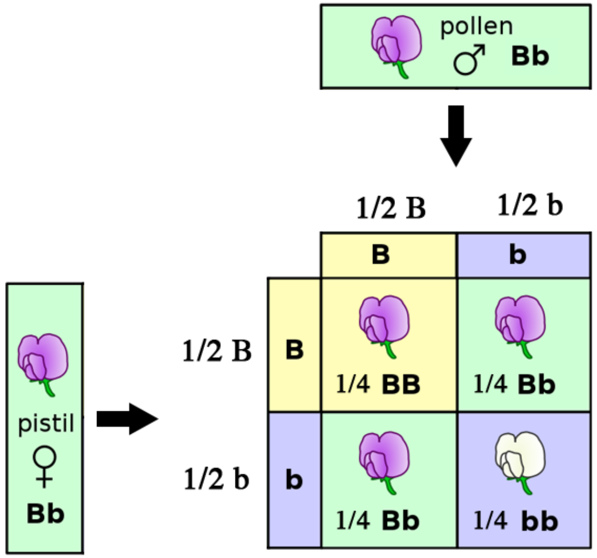
Describe the set of possible outcomes when a pistil of Bb and a pollen of Bb crosses.



Sample space – examples

Recall the heredity example in Lecture 2.

Describe the set of possible outcomes when a pistil of Bb and a pollen of Bb crosses.



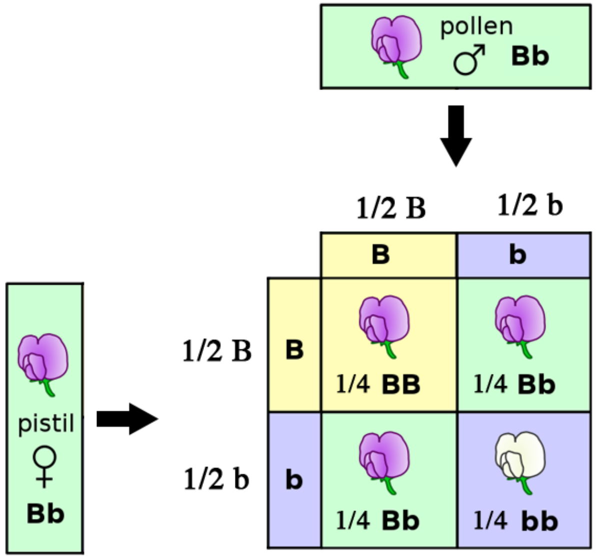
Case 1: Bb and bB are equivalent

Sample space: $\{BB,\ Bb,\ bb\}$

Sample space – examples

Recall the heredity example in Lecture 2.

Describe the set of possible outcomes when a pistil of Bb and a pollen of Bb crosses.



Case 2: Bb and bB are not equivalent

Sample space: $\{BB,\ Bb,\ bB, bb\}$

Events

- Events are sets:
  - ✓ Can describe in words
  - ✓ Can describe in notation
- Experiment: toss a coin 2 times.
- Event -- You get 1 or more heads

  = {HH, HT, TH}

Set operations

- Events are sets, so we can use set operations

  - ✓ Unions
  - ✓ Intersections
  - ✓ Complements

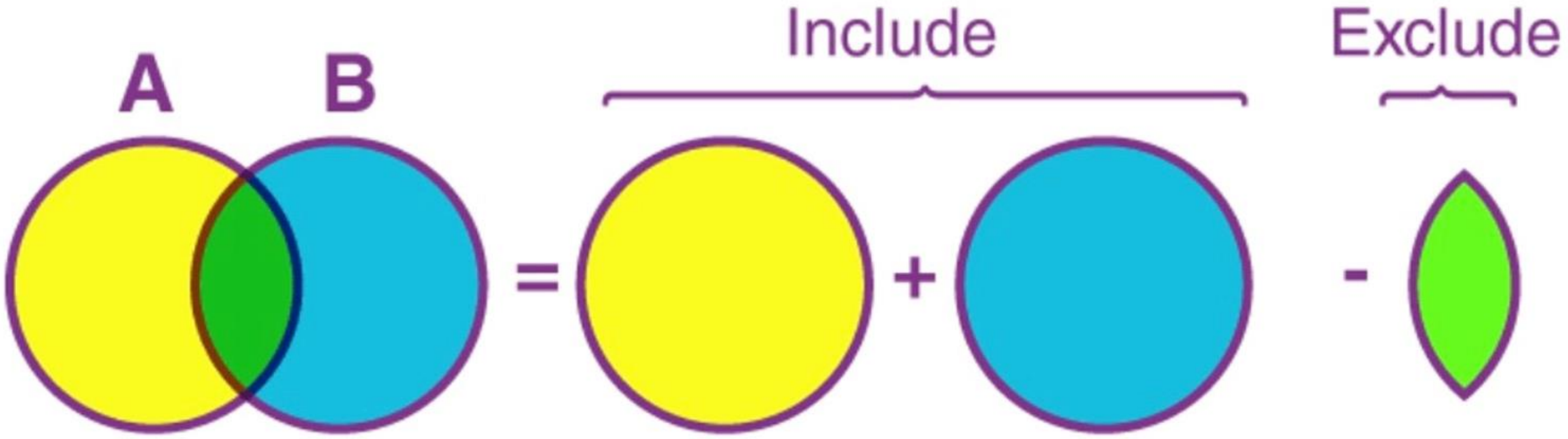
Set operations

- We denote the union as (A or B) in words, and $\mathrm{A} \cup \mathrm{B}$ in notation
- We denote the intersection as (A and B) in words, and $\mathrm{A} \cap \mathrm{B}$ in notation
- Two events A and B, such that $\mathrm{A} \cap \mathrm{B} = \emptyset$ are said to be **mutually exclusive**.
- We denote the complement as (not A) in words, and A' or A^c in notation

Set operations

- The commutative laws: $A \cup B = B \cup A$, $A \cap B = B \cap A$
- The associative laws: $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$
- The distributive laws: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- De Morgan's laws: $(A \cup B)' = A' \cap B'$, $(A \cap B)' = A' \cup B'$

Set operations

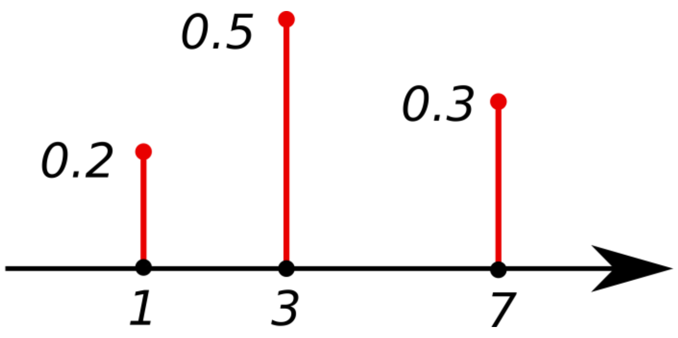


$$P(A \text{ or } B) = P(A) + P(B) - P(\ A \text{ and } B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Probability function

- Discrete:
  - ✓ Probability mass function.
  - ✓ P(ω): gives the probability for **each** outcome ω ∈ S

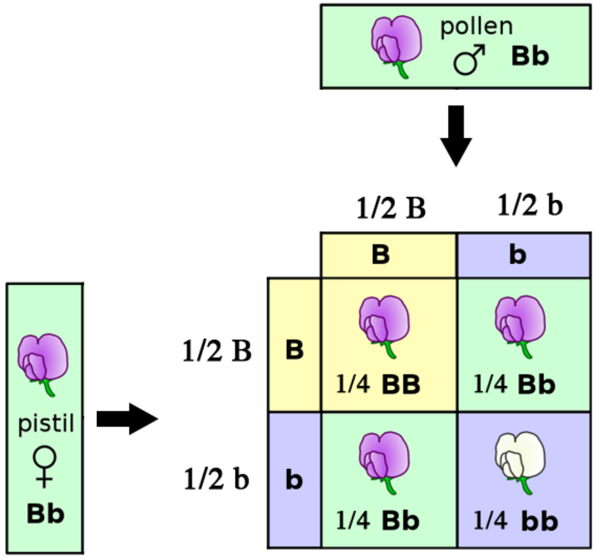


Continuous case will be defined later.

Probability Function- examples

Recall the heredity example in Lecture 2.

Write down P(ω) for each ω ∈ S



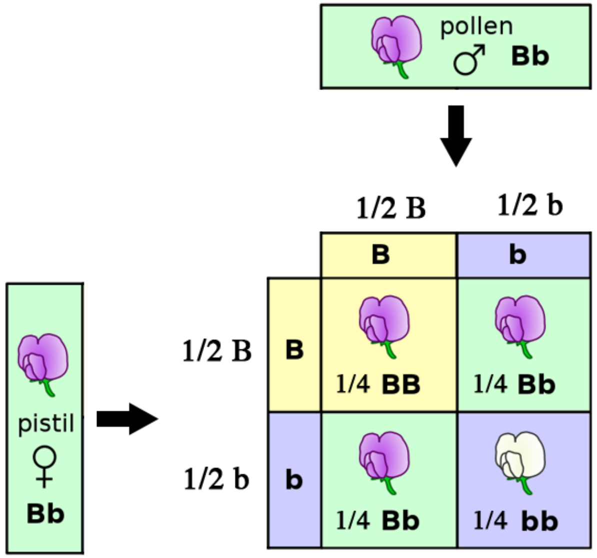
Case 1: Bb and bB are equivalent

Sample space: $\{BB,\ Bb,\ bb\}$

Probability Function- examples

Recall the heredity example in Lecture 2.

Write down P(ω) for each ω ∈ S



Case 1: Bb and bB are equivalent

Sample space: $\{BB,\ Bb,\ bb\}$

$\mathrm{P}(BB)$=1/4
$\mathrm{P}(Bb)$=1/2
$\mathrm{P}(bb)$=1/4

Independence

- Two events: A and B
- Does knowing something about A tell us whether B happens (and vice versa)?

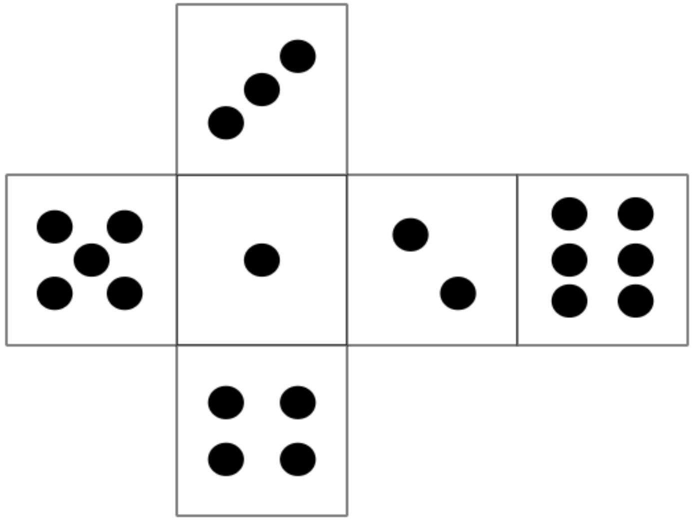
Independence

- Two events: A and B

- Does knowing something about A tell us whether B happens (and vice versa)?

- **Independent**
  - A: the first number is 1;
  - B: the second number is 1

  P(A and B) = P(A)·P(B)



Two events A, and B are independent if and only if $P(A \cap B) = P(A) \cdot P(B)$

Does knowing something about A tell us whether B happens (and vice versa)?

- **Dependent**
  - A: the first number is 1;
  - B: the sum of two numbers is larger than 2

P(A) = 1/6
P(B) = 35/36
P(A and B) = 5/36
$P(A \text{ and } B) \neq P(A) \cdot P(B)$

SAMPLE SPACE FOR A PAIR OF DICE

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

- Independence: $P(A \text{ and } B) = P(A)P(B)$
- Mutually exclusive: $P(A \text{ or } B) = P(A)+P(B)$

2. (Discrete) Random Variable and Probability Distributions.

In applications we are interested in **quantitative** properties of experimental results.

Example 1

- Toss a coin three times and count the number of heads.
- The sample space is

$$S = \{(t,t,t),(t,t,h),(t,h,t),(h,t,t),(t,h,h),(h,t,h),(h,h,t),(h,h,h)\}$$

- Lex X count the number of heads. Thus if $s = (t,t,h)$ occurs then $X(s) = 1$.
- X is called a random variable as it takes a numerical value that depends on the outcome of an experiment.

Examples 2:

Roll a fair die twice, and let the random variable **X** denote the summation of the two numbers.

- What is the **range** (possible values) of the random variable? (think about the sample space)

- How can you describe the event that the summation is larger than 10 using the random variable X?

Examples 2:

Roll a fair die twice, and let the random variable **X** denote the summation of the two numbers.

- What is the **range** (possible values) of the random variable? (think about the sample space)

$$\{2,3,4,5,6,7,8,9,10,11,12,13,14,15,16\}$$

- How can you describe the event that the summation is larger than 10 using the random variable X?

Examples 2:

Roll a fair die twice, and let the random variable **X** denote the summation of the two numbers.

- What is the **range** (possible values) of the random variable? (think about the sample space)

$$\{2,3,4,5,6,7,8,9,10,11,12\}$$

- How can you describe the event that the summation is larger than 10 using the random variable X?

$\{\omega: X(\omega) > 10\}$ or $\{X > 10\}$

Examples 3:

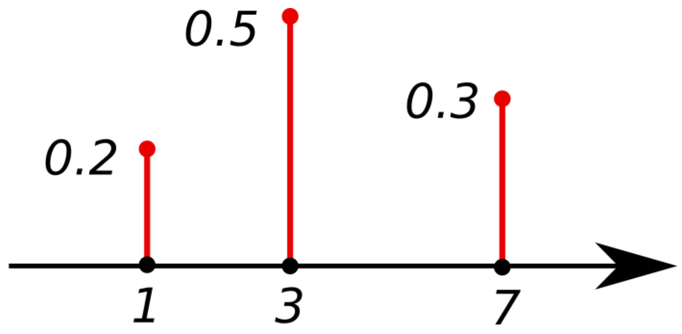
A group of 10,000 people are tested for a gene called Ifi202 that has been found to increase the risk for lupus. The random variable **X** is the number of people who carry the gene.

- What is the range (possible values) of the random variable? (think about the sample space)

- How can you describe the event that more than half of the people carry the gene using the random variable X?

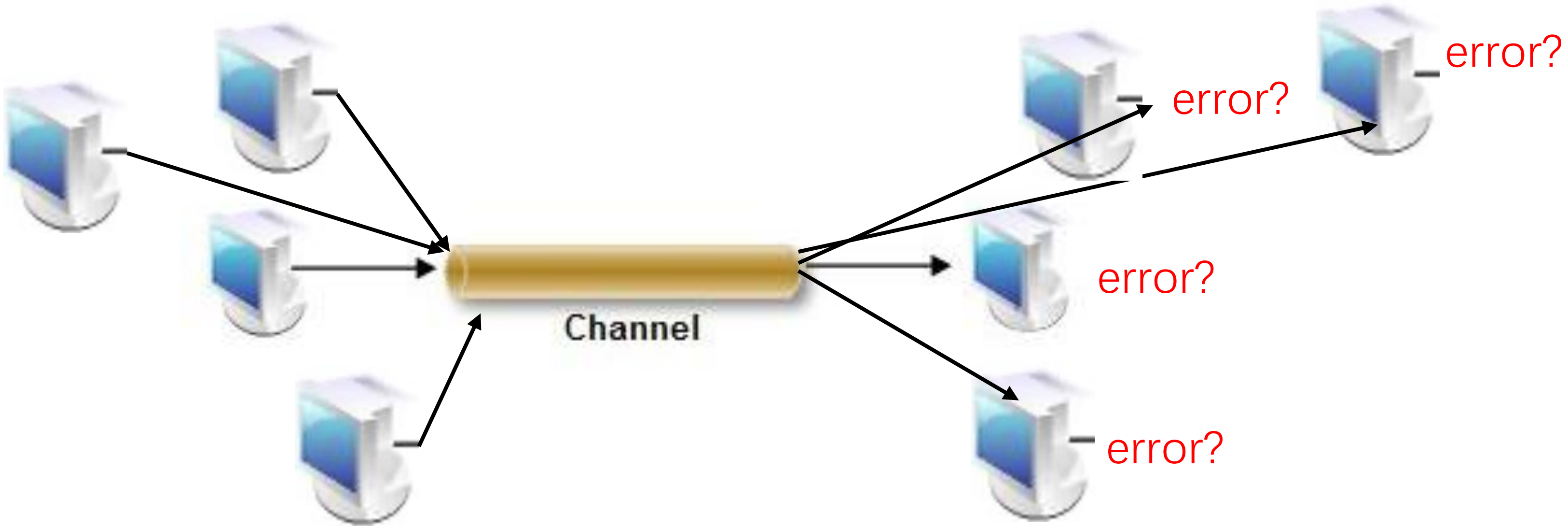
Probability Distributions

- The probability distribution of a random variable X is a description of the probabilities associated with the possible values of X.
- For discrete random variable, the distribution is just a list of values, e.g., {0.2,0.5,0.3}.



Probability Distributions - Example

- There is a chance that a bit transmitted through a digital transmission channel is received in error.
- Let X equals the number of bits in error in the next four bits transmitted.
- Q: What is the possible values for X?



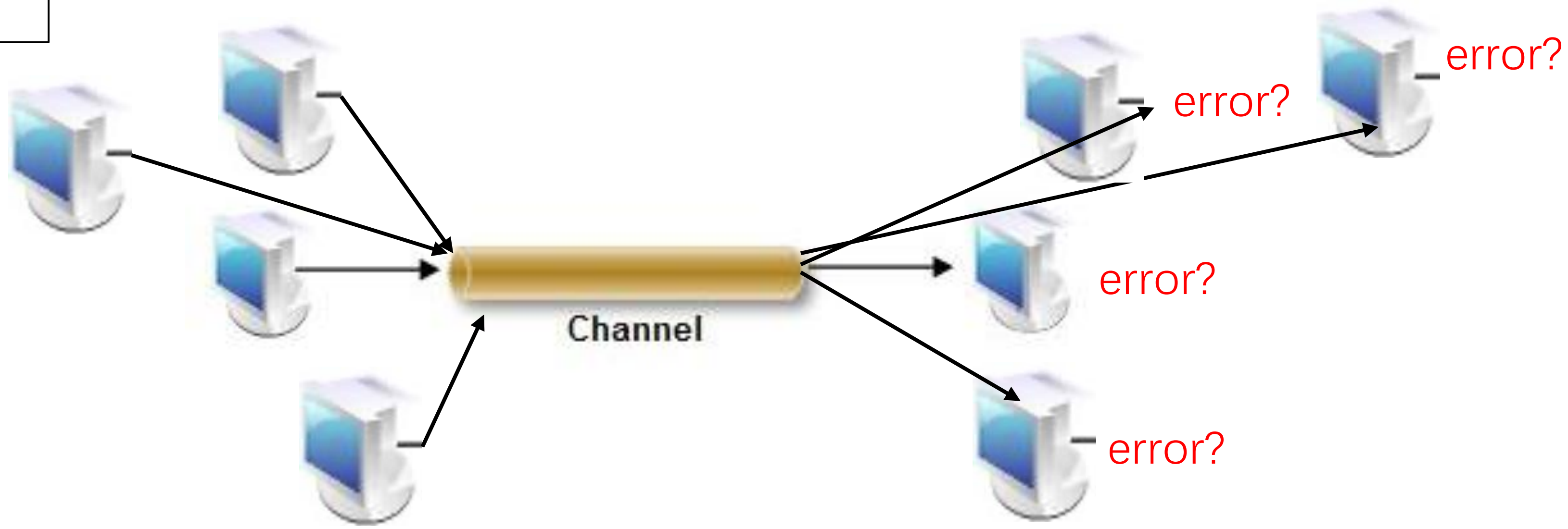
Probability Distributions - Example

- The expert gives the following probability distribution on X.

$$\boxed{P(X = 0)} = 0.6561 \qquad P(X = 1) = 0.2916 \qquad P(X = 2) = 0.0486$$

$$P(X = 3) = 0.0036 \qquad P(X = 4) = 0.0001$$

Probability that no error received.



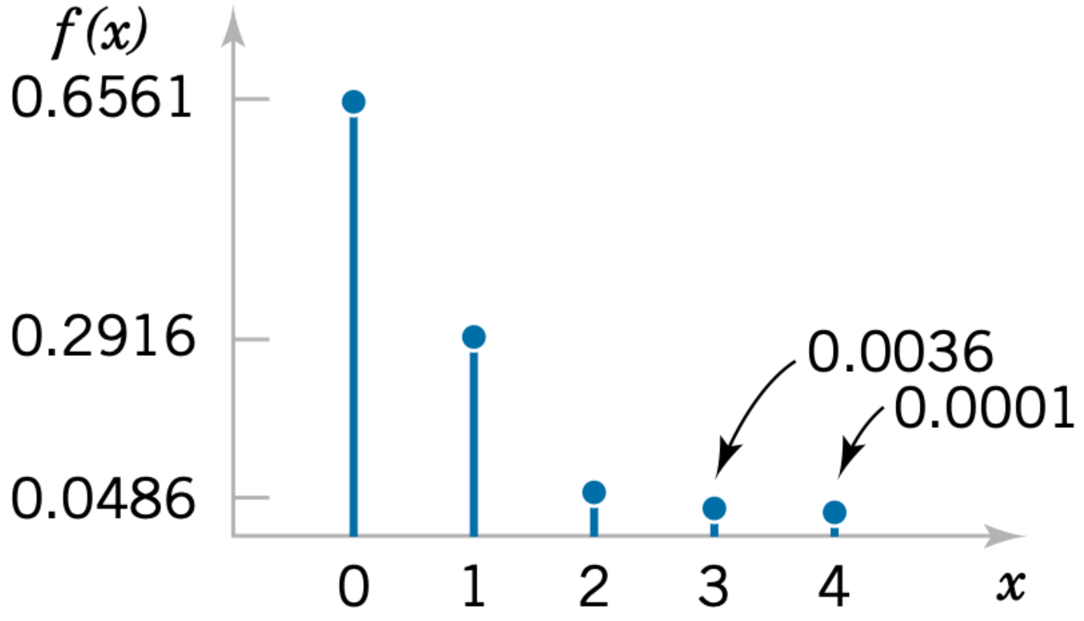
Probability Distributions - Example

- The expert gives the following probability distribution on X.

$\boxed{P(X=0)} = 0.6561 \qquad P(X=1) = 0.2916 \qquad P(X=2) = 0.0486$

$P(X=3) = 0.0036 \qquad P(X=4) = 0.0001$

Probability that no error received.



Probability distribution for bits in error

Formal definition

- For a discrete random variable X with possible values $x_1, x_2, \ldots, x_n$. A probability mass function $f(\cdot)$ is a function such that:

✓ $f(x_i) \geq 0$ for all $x_1, x_2, \ldots, x_n$.

✓ $\sum_{i=1}^{n} f(x_i) = 1$

✓ $f(x_i) = \boxed{P(X = x_i)}$ for all $x_1, x_2, \ldots, x_n$.

Probability that the random variable takes value x_i.

Q: What is the probability mass function f for the previous digital transmission example?

Exercise

- The sample space of a random experiment is {a,b,c,d,e,f}, and each outcome is equally likely. A random variable X is defined as follows:

outcome	a	b	c	d	e	f
x	0	0	1.5	1.5	2	3

- Q: What is the probability mass function f of X?
- Use the probability mass function to determine:

(a) $P(X = 1.5)$
(b) $P(0.5 < X < 2.7)$
(c) $P(X > 3)$
(d) $P(0 \leq X < 2)$
(e) $P(X = 0 \quad \text{or} \quad X = 2)$

Exercise

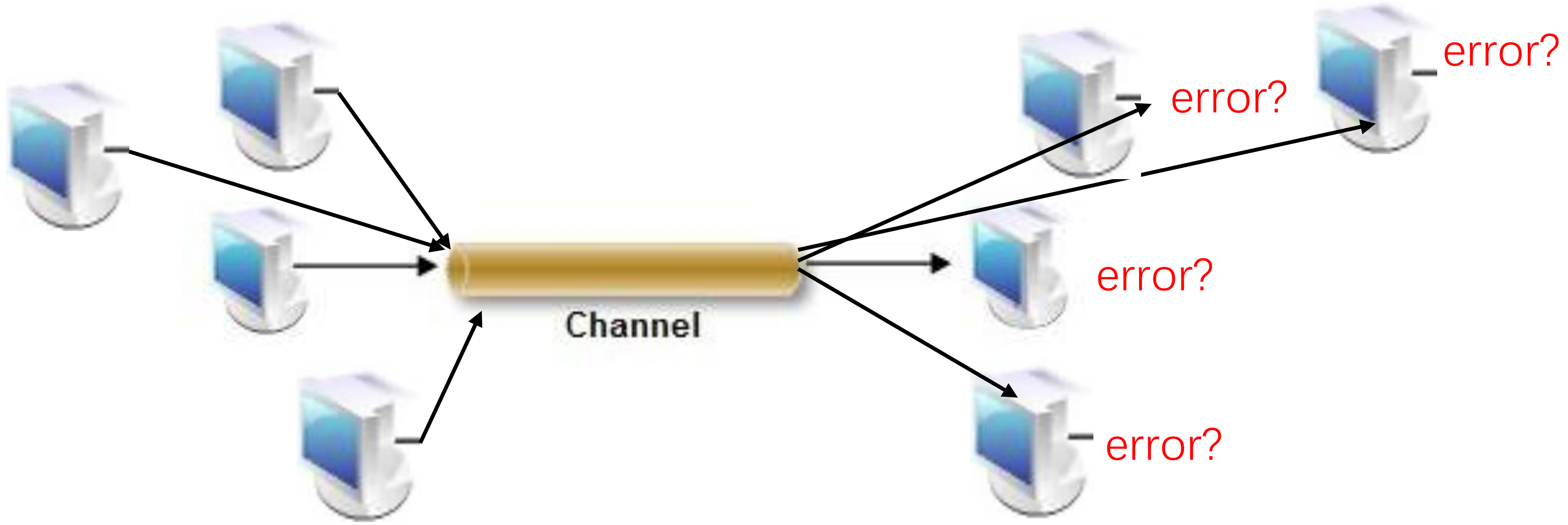
- Verify that the following function f is a probability mass function, and determine the requested probabilities.

$$f(x) = \frac{2x + 1}{25}, \quad x = 0, 1, 2, 3, 4$$

(a) $P(X = 4)$ (b) $P(X \leq 1)$
(c) $P(2 \leq X < 4)$ (d) $P(X > -10)$

Cumulative Distribution Functions (CDF)

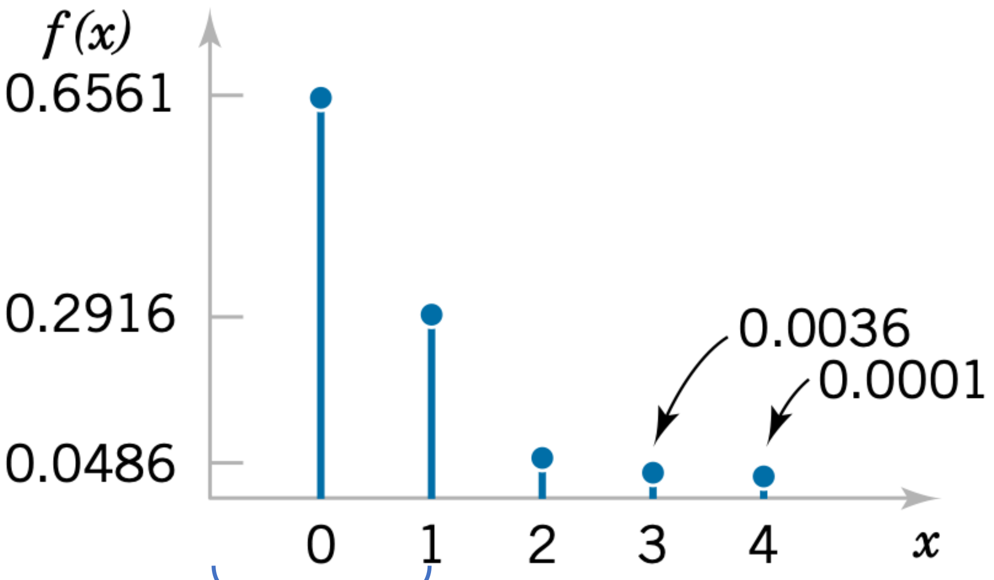
- Recall the digital transmission example.

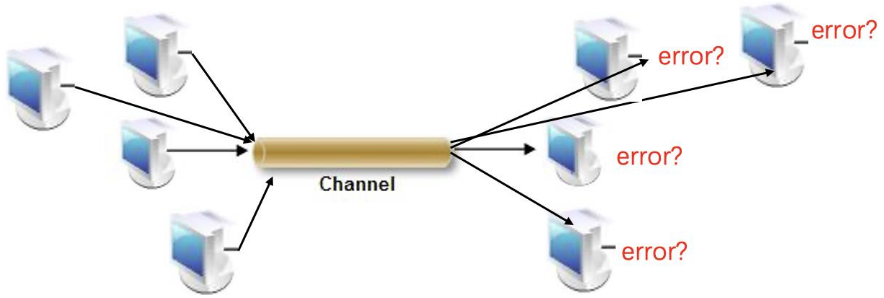


- What is the probability of having <u>three or fewer</u> bits in error?
- What about <u>two or fewer</u>?
- What about <u>four or fewer</u>? → Cumulative probabilities

Cumulative Distribution Functions (CDF)

- Cumulative Probabilities





probability of having one or fewer bits in error: 0.6561+0.2916

probability of having two or fewer bits in error: 0.6561+0.2916+0.0486

Cumulation!

?

?

Cumulative Distribution Functions (CDF)

- The cumulative distribution function (CDF) of a discrete random variable X, denoted as $F(x)$ is

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

- $F(x)$ satisfies:

  ✓ $0 \leq F(x) \leq 1$

  ✓ If $x \leq y$, then $F(x) \leq F(y)$

Note: even if the random variable X can only take integer values, the cumulative distribution function can be defined at **non-integer** values, for example, $F(1.5) = P(X = 0) + P(X = 1) = 0.6561 + 0.2916$.

Cumulative Distribution Functions (CDF)

- The probability mass function provides probabilities.
- The cumulative distribution function can also provide probabilities, and it uniquely determines the probability mass function of a discrete random variable.

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

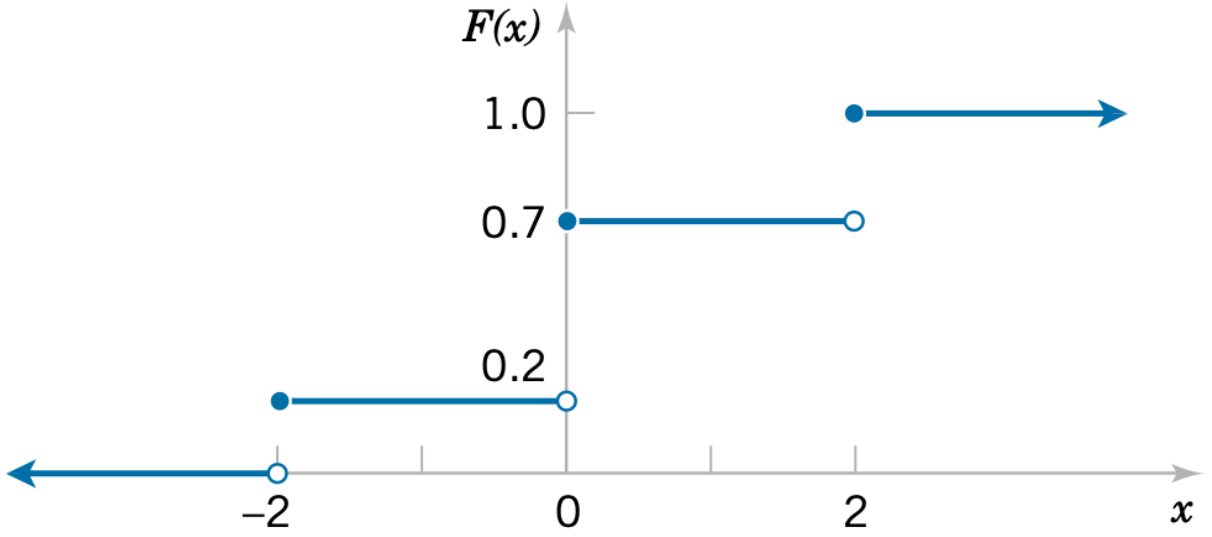
- Why?

Example

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

- Find the probability mass function of X form the following cumulative distribution function

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.7 & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

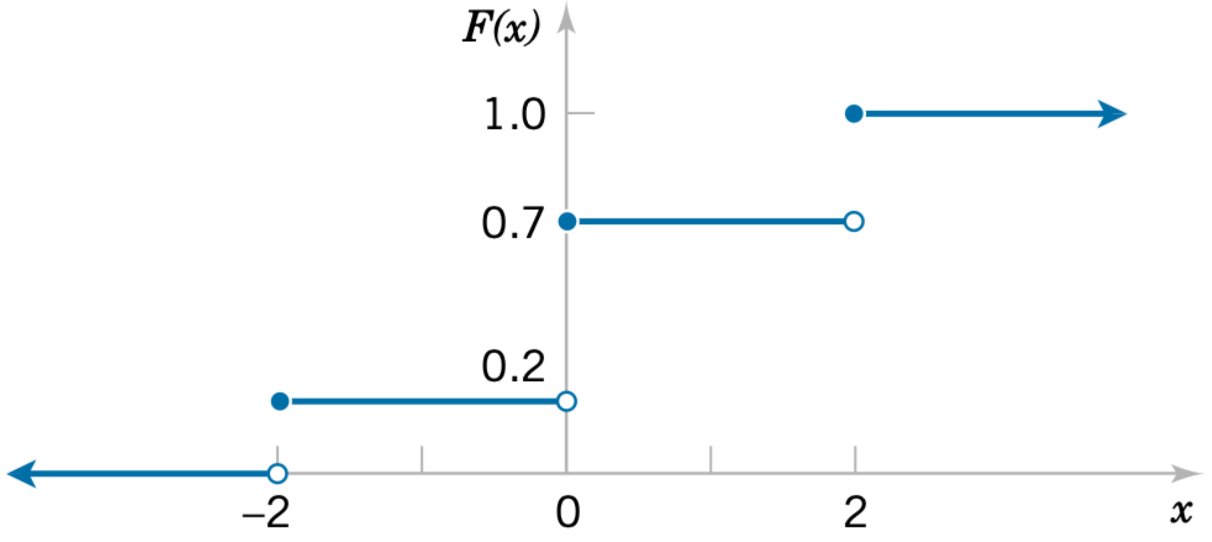


Example

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

- Find the probability mass function of X form the following cumulative distribution function

$$F(x) = \begin{cases} 0 & x < -2 \\ 0.2 & -2 \leq x < 0 \\ 0.7 & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$



$f(-2)=0.2$ $f(0)=0.5$ $f(2)=0.3$

3. Mean and Variance.

QUESTION

- Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Game 2:

- If Heads, you pay me $1000.1.
- If Tails, I pay you $1000.

QUESTION

- Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Game 2:

- If Heads, you pay me $1000.
- If Tails, I pay you $1000.1.

Return
Game 2 > Game 1

QUESTION

- Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

Game 2:

- If Heads, you pay me \$1000.1.
- If Tails, I pay you \$1000.

Return
Game 2 > Game 1

Risk
Game 2 > Game 1

QUESTION

- Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

Game 2:

- If Heads, you pay me $1000.1.
- If Tails, I pay you $1000.

Return
Game 2 > Game 1

Risk
Game 2 > Game 1

How can we quantify the return and risk?

Example: Investment

Maximize average return?

Example: Investment

Maximize average return?

No! Also reduce the chance of losing too much after one investment.

Example: Investment

Maximize average return?

No! Also reduce the chance $\underbrace{\text{of losing too much after one investment.}}_{\text{Risk}}$

Example: Investment

Maximize average return?

No! $\underbrace{\text{Also reduce the chance of losing too much after one investment.}}_{\textbf{Risk}}$

Real objective:

Example: Investment

Maximize average return?

No! Also reduce the chance of losing too much after one investment.

Risk

Real objective:

Maximize: Average Return - Risk

Example: Investment

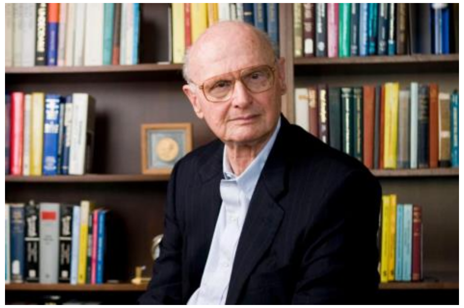
Maximize average return?

No! Also reduce the chance of losing too much after one investment.

$\underbrace{\text{reduce the chance of losing too much after one investment}}_{\text{Risk}}$

Real objective:

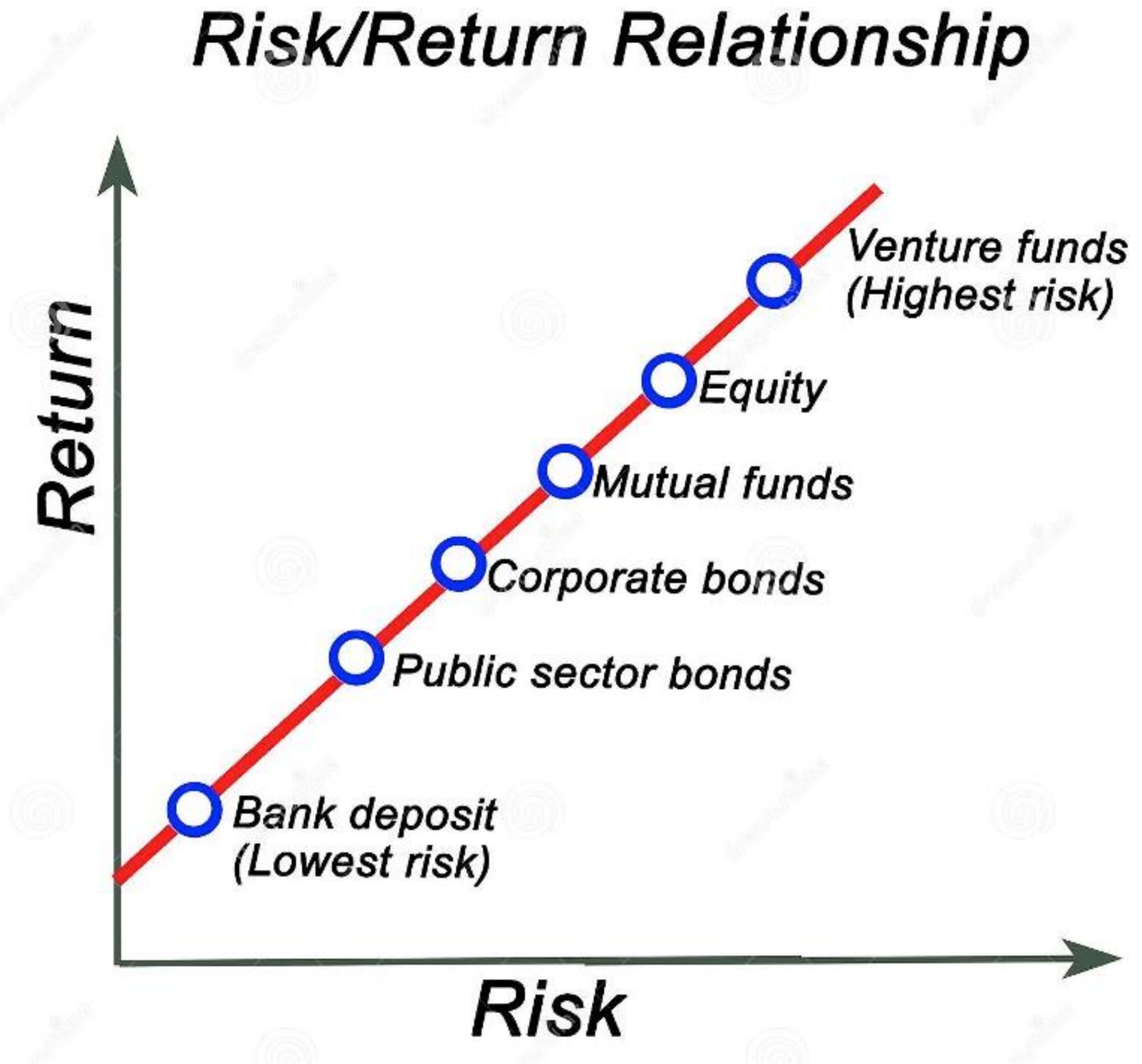
Maximize: Average Return - Risk

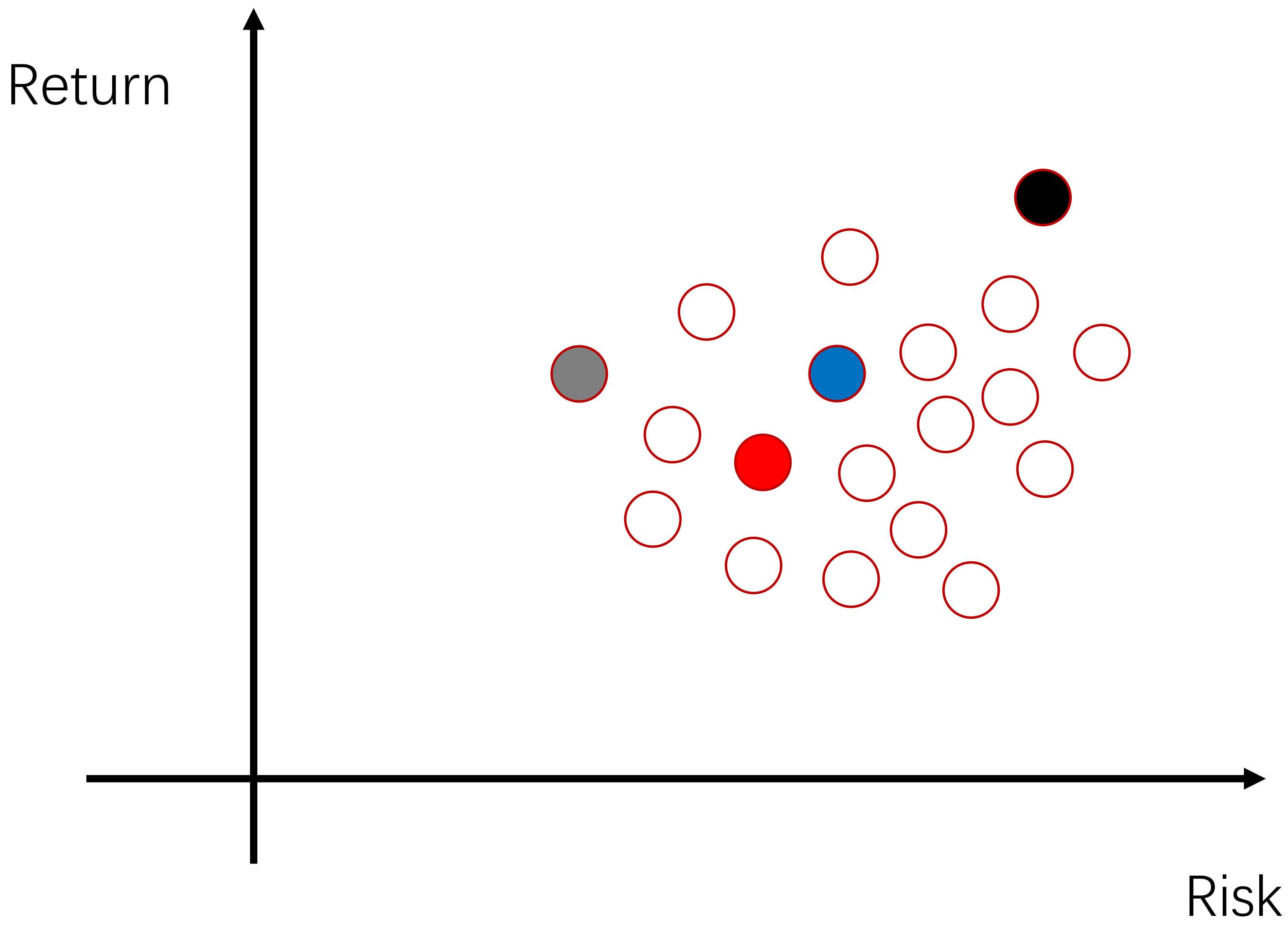


Markowitz Model

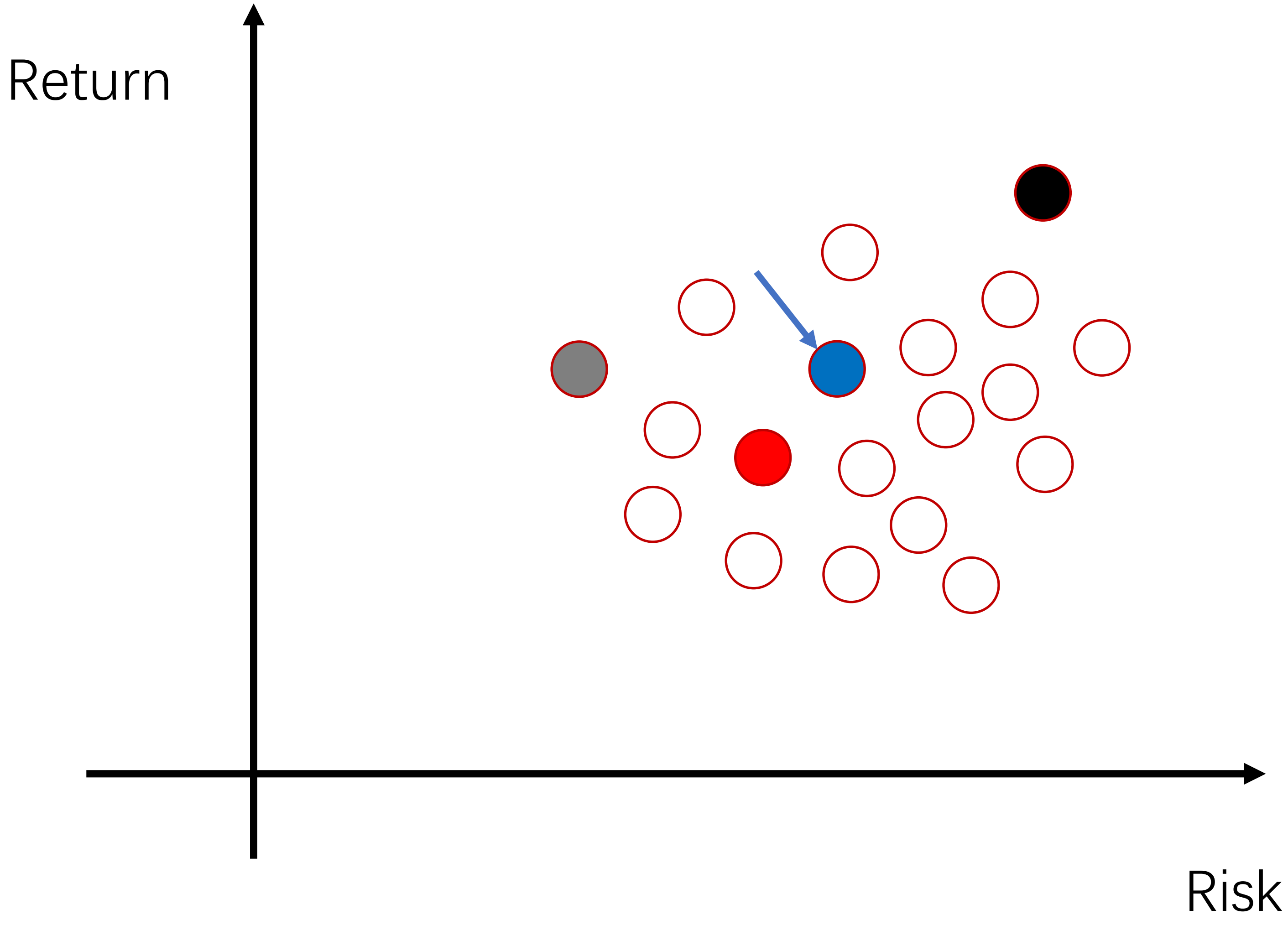
The first model you need to learn to work in finance.

Example: Investment



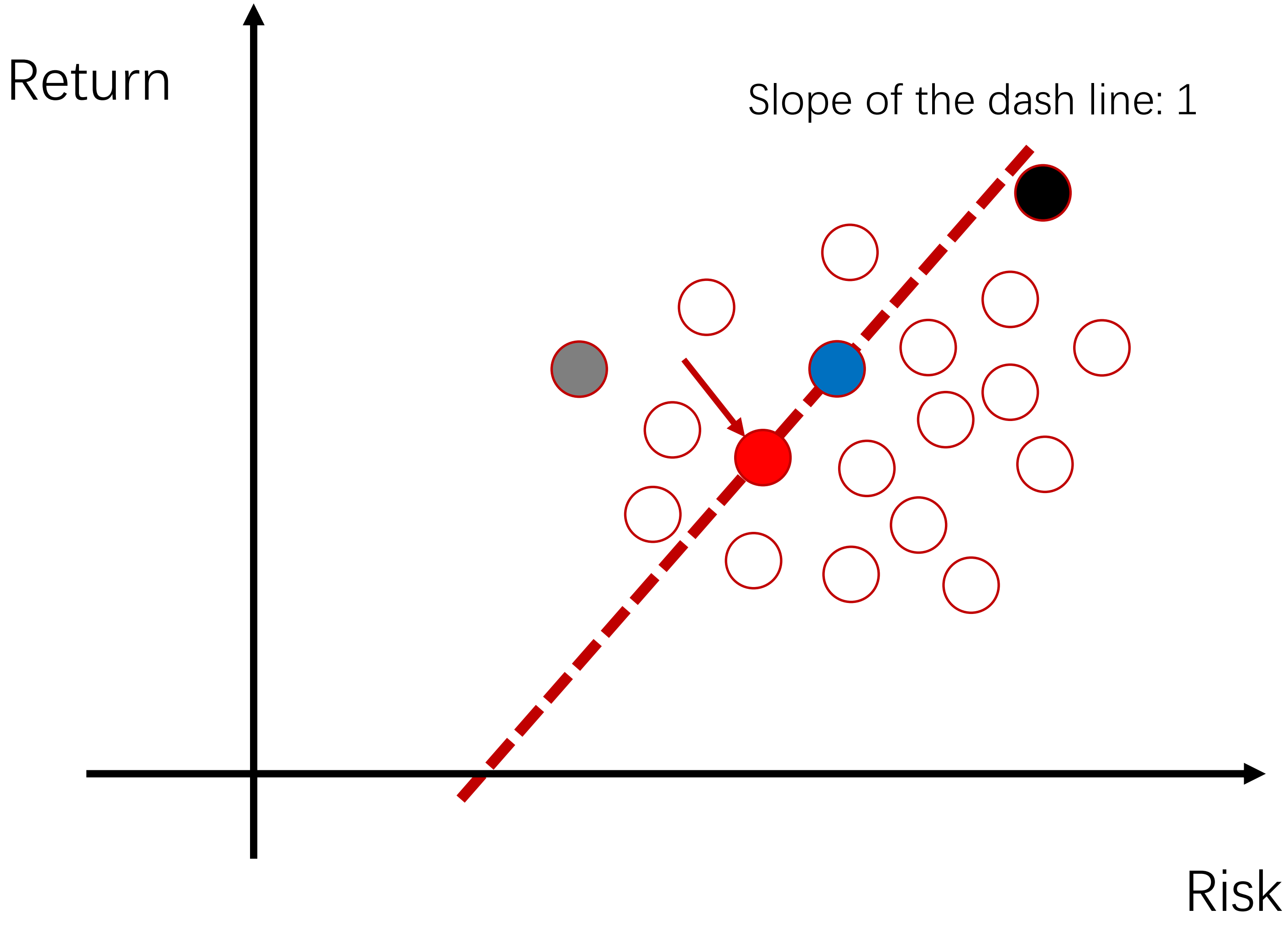


Each dot represents a stock



Objective: Return – Risk
Suppose you choose blue

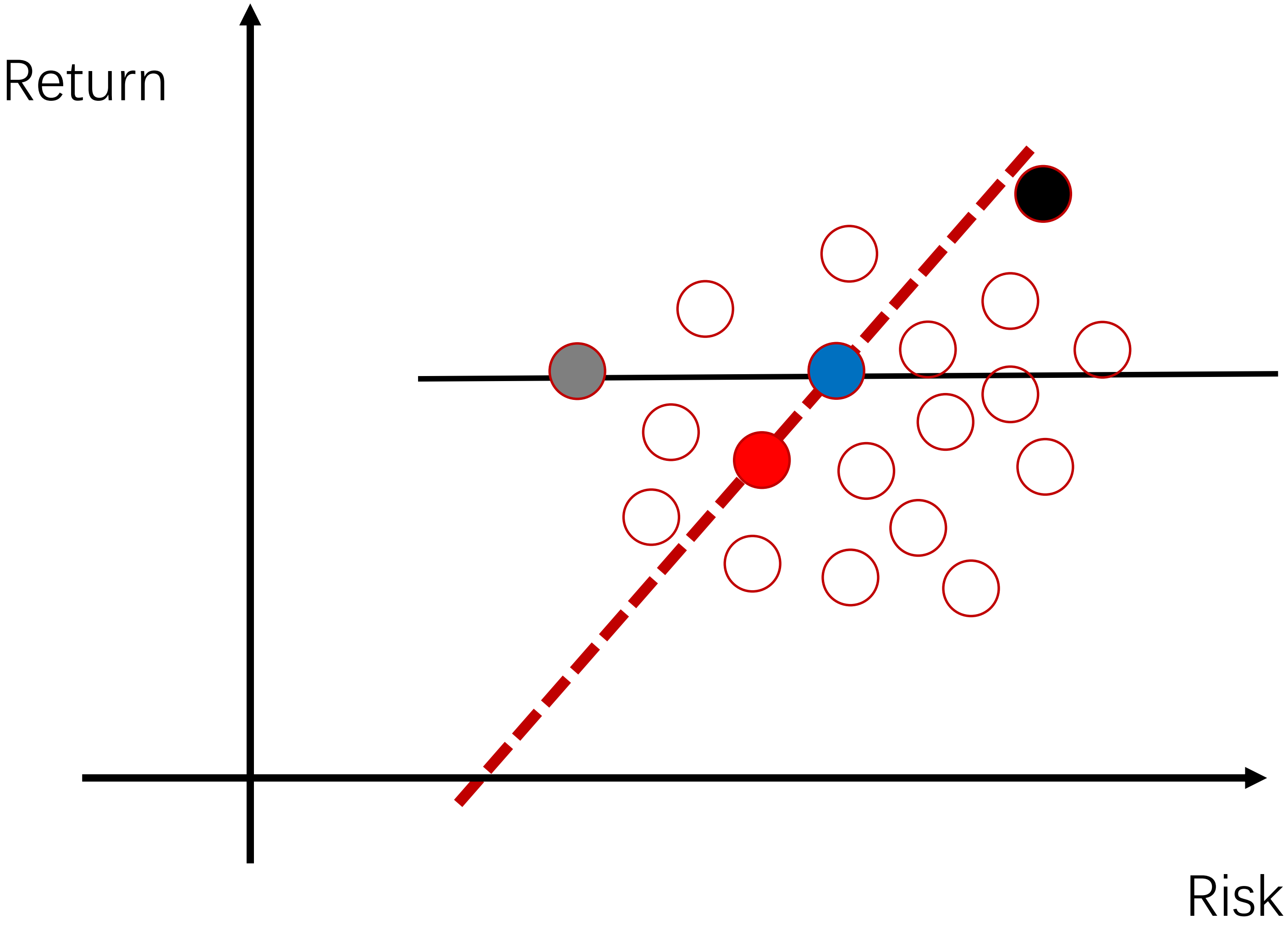
Each dot represents a stock



Objective: Return – Risk
Suppose you choose blue

Objective unchanged when choosing red

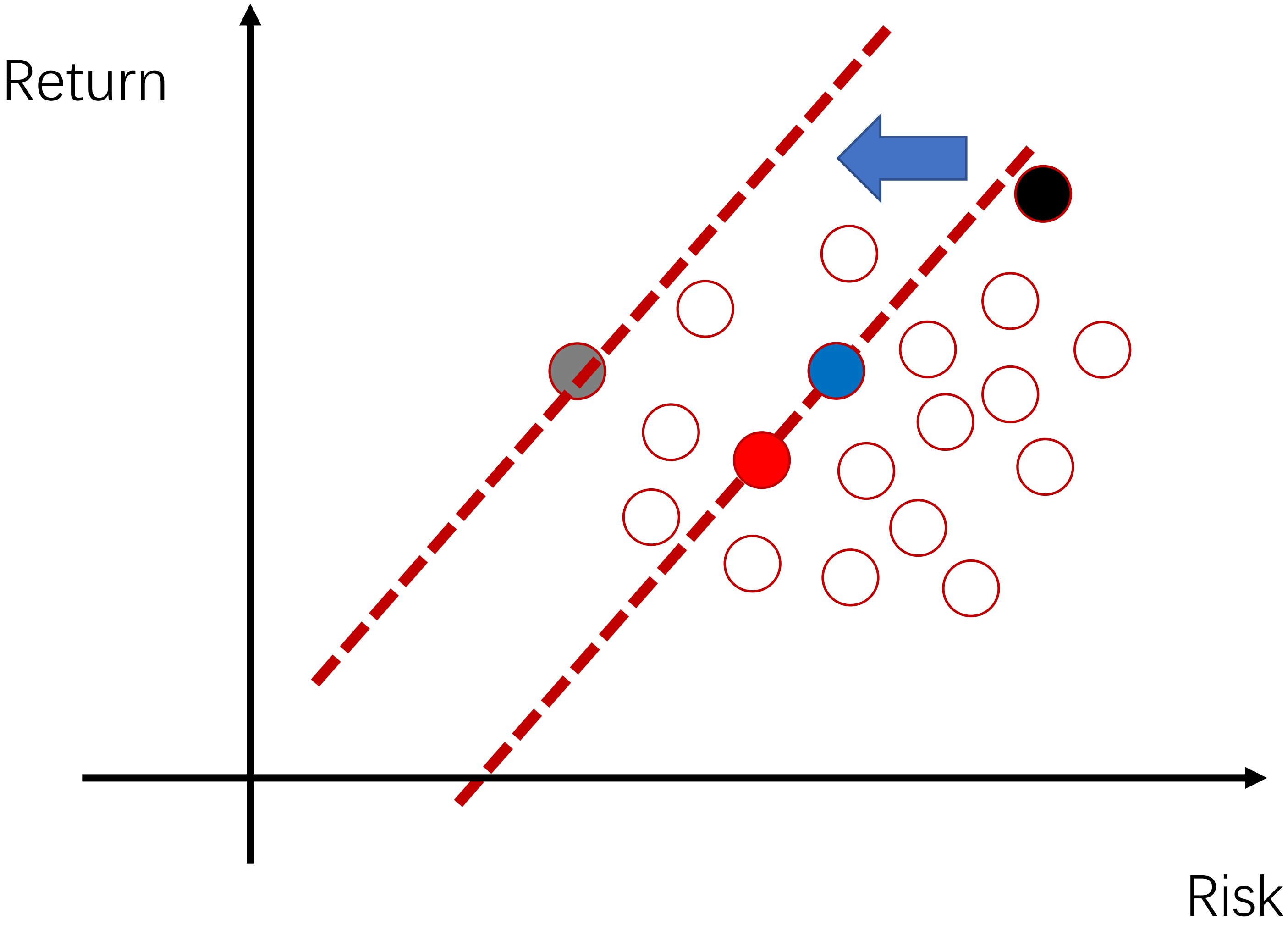
Each dot represents a stock



Objective: Return – Risk
Suppose you choose blue

When choosing grey:
same return but smaller risk

Each dot represents a stock



Objective: Return – Risk
Suppose you choose blue

When choosing grey:
same return but smaller risk

Best choice: move the line with slope 1 to left side until there is no dot at the left side of the line

Each dot represents a stock

How to calculate Return and Risk?

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

- Return

- Risk

How to calculate Return and Risk?

Game 1:

- If heads, You pay me $1.
- If Tails, I pay you $1.

- Return -> Mean

$$E[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x)$$

- Risk

How to calculate Return and Risk?

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

- Return -> Mean

$$\mathrm{E}[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$$

- Risk

How to calculate Return and Risk?

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

- Return -> Mean

$$E[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$$

- Risk -> How far is a random variable from its mean, on average?

$$|\boldsymbol{X} - \boldsymbol{E}[\boldsymbol{X}]|$$

How to calculate Return and Risk?

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

- Return -> Mean

$$\mathrm{E}[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$$

- Risk -> Variance

$$\mathrm{Var}[X] = \Sigma_x (x - \mathrm{E}[X])^2\, f(x)$$

How to calculate Return and Risk?

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

- Return -> Mean

$$\mathrm{E}[X] = \Sigma_x x\, P(X = x) = \Sigma_x x\, f(x) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0$$

- Risk -> Variance

$$\mathrm{Var}[X] = \Sigma_x (x - \mathrm{E}[X])^2\, f(x) = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1$$