

# Lecture 1 - 4<sup>th</sup> Sept 2024

Statistics: the science of understanding data and making decisions in the face of variability and uncertainty.

Probability: A branch of mathematics concerned with describing and modeling uncertain events.

## Preliminaries:

- Experiment: the process of obtaining an observed result of some phenomenon.
- Trial: the performance of an experiment.
- Outcome: the result of a single trial (attempt) of an experiment.
- Event: one or more outcomes of an experiment.
- Probability: the measure of how likely an event is.

Sample Space: the set of ALL possible distinct outcomes in a random experiment, denoted by  $S$ .

↳ note: one and only one of the outcomes occurs in any single trial of the experiment.

↳ example:

- Roll a six-sided die:  $S = \{1, 2, 3, 4, 5, 6\}$
- Flip a coin:  $S = \{\text{heads, tails}\}$
- Waiting time for a bus:  $S = \{t \in \mathbb{R}, 0 \leq t \leq 10\}$
- A sample space is finite if it consists of a

finite number of outcomes, say  $\{a_1, a_2, \dots, a_n\}$

- A sample space can be a set of countable infinite outcomes, say  $\{a_1, a_2, \dots\}$ , where the outcomes can be put into a one-on-one correspondence with positive integers.

↳ Example: suppose our experiment consists of tossing a coin until it lands on heads.

↳ the possible outcomes in the sample space are:

H, TH, TTH, TTTH, TTTTH, ...

∴ the outcomes are countable but infinite.

Discrete Sample Space: when a sample space is either finite or countably infinite. Else, it's non-discrete.

↳ Example:

- Roll a six-sided die: discrete
- Flip a coin: discrete
- Waiting time for a bus: non-discrete

Event: a subset of sample space S.

↳ we say A is an event if  $A \subset S$ .

↳ note: A and S are both sets!

Example: consider tossing two coins.

a) what is the event of obtaining at least one head?

$$S = \{ HH, HT, TH, TT \}$$

$$\therefore A = \{ HH, HT, TH \}$$

b) What is the event of obtaining at least one head and at least one tail?

$$A = \{ HT, TH \}.$$

**Elementary / Simple Event:** an event that contains only one outcome of the experiment. Eg,  $A = \{ \alpha, \beta \}$ .

**Compound Event:** an event made up of two or more simple events. Eg:  $A = \{ \alpha_1, \alpha_2 \}$ .

**Set Notations and Terminology:**

- A set is a collection of elements. Eg,  $A = \{ \text{apple, orange} \}$ .
- The notation  $\alpha \in A$  or  $\alpha \notin A$  will mean that  $\alpha$  is or is not an element of  $A$ .
- The empty / null set is  $\emptyset$ .
- Union:  $A \cup B = \{ \alpha \mid \alpha \in A \text{ or } \alpha \in B \}$
- Intersection:  $A \cap B = \{ \alpha \mid \alpha \in A \text{ and } \alpha \in B \}$
- The notation  $B \subset A$  means that  $B$  is a subset of  $A$ .
- Complement:  $A^c = \{ \alpha \mid \alpha \in S, \alpha \notin A \} = \bar{A}$
- Two sets  $A$  and  $B$  are disjoint if  $A \cap B = \emptyset$ .

- The set  $A \cap B^c$  is "A but not B"
- The set  $A^c \cap B^c$  is "neither A nor B"

- If  $A_1, \dots, A_n$  is a finite collection of sets,

$\hookrightarrow A_1 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i$  is "in all  $A_i$ ;"

$\hookrightarrow A_1 \cup \dots \cup A_n = \bigcup_{i=1}^n A_i$  is "in at least one  $A_i$ ;"

*p very similar to disjoint!*

**Mutually Exclusiveness:** when two events A and B if  $A \cap B = \emptyset$ . If events are exclusive, they have no outcomes in common.

$\hookrightarrow$  Example: tossing two coins. A is the event "at least one head", and B is the event "both tails".

$$\hookrightarrow S = \{HH, HT, TH, TT\}.$$

$$A = \{HH, HT, TH\}, B = \{TT\}.$$

$$A \cap B = \{HH, HT, TH\} \cap \{TT\} = \emptyset !$$

$\therefore$ , A and B are mutually exclusive.

$\hookrightarrow$  Events  $A_1, A_2, A_3, \dots$  are said to be mutually exclusive if they are pairwise mutually exclusive. That is,

$$A_i \cap A_j = \emptyset, \text{ wherever } i \neq j.$$

Probability Modeling : assigning a probability  $P(A)$  to each event  $A$ . This probability measures how likely it is that  $A$  will happen when the experiment is conducted.

- ↳ think of  $P(A)$  as a (set) function, whose domain is a collection of sets (events), and the range of which is a subset of real numbers.
- ↳ not all set functions are appropriate for assigning probabilities to events.

• Let  $S = \{\alpha_1, \alpha_2, \alpha_3, \dots\}$  be a discrete sample space. We assign to each elementary event  $A_i = \{\alpha_i\}$  for  $i = 1, 2, \dots$ , a number  $P(A_i) = P(\{\alpha_i\})$ , such that

$$\bullet 0 \leq P(A_i) \text{ and } \sum_{\text{all } i} P(A_i) = 1$$

• We call  $P(A)$  the probability of  $A$ , and call the set of probabilities  $\{P(A_i), i=1, 2, \dots\}$  the probability distribution on  $S$ .

#### ↳ General Properties :

- The null event (empty set) has a probability of 0. ∴  $P(\emptyset) = 0$ .
- If  $A$  and  $B$  are two mutually exclusive events, then  $P(A \cup B) = P(A) + P(B)$
- If  $A_1, A_2, \dots, A_k$  is a finite collection of pairwise mutually exclusive events, then

$$P(A_1 \cup A_2 \cup \dots \cup A_R) = P(A_1) + P(A_2) + \dots + P(A_R)$$

$$\cdot P(\bar{A}) = 1 - P(A)$$

↳ Example: Suppose a six-sided die is rolled once.  
If A is the event that an even number is obtained.  
What is  $P(A)$ ?

- $S: \{1, 2, 3, 4, 5, 6\}$ . Since each number is equally likely,  $P(\{i\}) = \frac{1}{6}$  for  $i = 1, 2, 3, 4, 5, 6$ .

$$\begin{aligned} A &= \{2, 4, 6\}. \rightarrow P(A) = P(\{2\}) + P(\{4\}) + P(\{6\}) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}. \end{aligned}$$

↳ Example: Toss a coin twice. What's the probability of getting exactly one head?

- $S: \{HH, HT, TH, TT\}$ . Each option is uniformly likely.
- $A = \{HT, TH\} \rightarrow P(A) = P(\{HT\}) + P(\{TH\}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ .

**Random Selection:** when choosing an object from a finite set where each object has an equal chance of being selected.

**Uniform Probability Model:** consider a discrete finite sample space  $S = \{a_1, a_2, \dots, a_n\}$ , where each

simple event has a probability of  $\frac{1}{N}$ . This is a uniform distribution over the set  $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ .

For a compound event A, we can calculate  $P(A)$  by:

$$P(A) = \frac{n(A)}{N},$$

where  $n(A)$  is the number of outcomes in A.

The formula  $P(A) = \frac{n(A)}{N}$  is classical probability.

(trivial) example: draw one card at random from a well-shuffled deck without jokers. What is the probability that the card is a spade?

$$S = \{A\heartsuit, 2\heartsuit, \dots\}.$$

$$A_s = \{A\spadesuit, 2\spadesuit, \dots, K\spadesuit\}.$$

$$P(A_s) = \frac{n(A_s)}{N} = \frac{13}{52} = \frac{1}{4}.$$

Example: suppose that P is a probability, and A and B are events such that A and B are mutually exclusive, and  $P(\bar{A}) = 0.8$  and  $P(\bar{B}) = 0.7$ . Is each statement true or false?

A) A and B are each non-empty

Since  $P(\bar{A}) = 0.8$  and  $P(\bar{B}) = 0.7$ , we know that

$$P(A) = 0.2 \text{ and } P(B) = 0.3.$$

we also know that  $P(\emptyset) = 0$ . Since  $0.2 \neq 0 \neq 0.3$ , we know A and B cannot be empty.  
 $\therefore$  the statement is true.

B)  $P(A \cap B) = 0$

Since we're told that A and B are mutually exclusive, we know that  $A \cap B = \emptyset$ .

we also know that  $P(\emptyset) = 0$ , so the statement is true!

C)  $P(A \cup B) = 0.4$

we know that  $P(A \cup B) = P(A) + P(B)$

"the probability of  
A or B"

$$P(A) + P(B) = 0.2 + 0.3 = 0.5 \neq 0.4.$$

$\therefore$  the statement is false!

Example: suppose two six-sided die are rolled where the outcomes are of the form (die 1, die 2). If A is the event that the first die is even, and B is the event that the sum of both rolls is 6, compute the event  $C = \bar{A} \cap B$

$$\hookrightarrow A = \{(2,1), (2,2), \dots, (4,1), (4,2), \dots, (6,1), (6,2), \dots, (6,6)\}.$$

$$B = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$$

$\bar{A}$  means the event that A does not happen, so in other words,  $\bar{A}$  is the event that the first die is odd.

$$\therefore \bar{A} = \{(1,1), \dots, (1,6), (3,1), \dots, (3,6), (5,1), \dots, (5,6)\}.$$

$\bar{A} \cap B$  is A AND B.

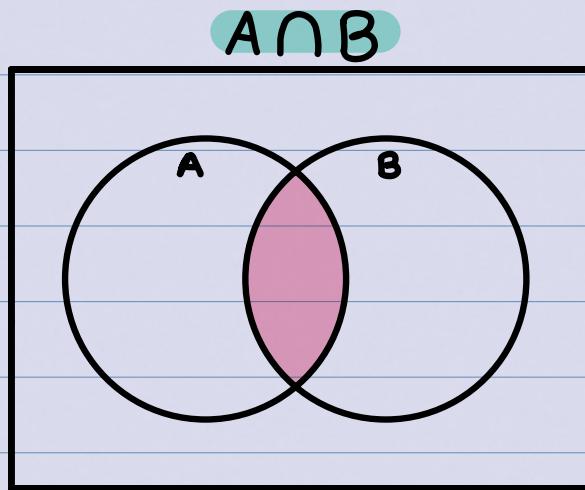
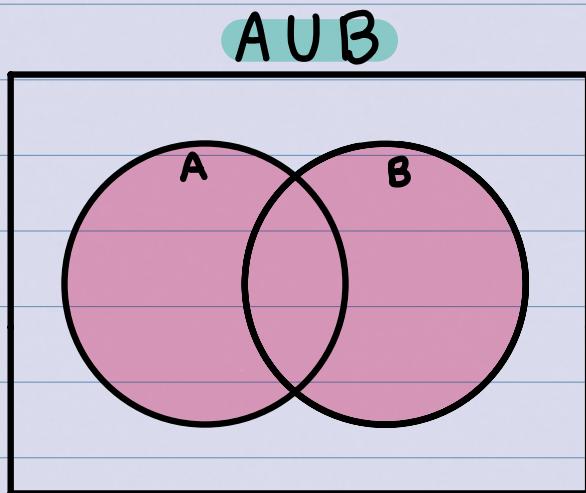
$$\hookrightarrow \therefore C = \{(1,5), (3,3), (5,1)\}.$$

Venn Diagram: a tool to illustrate the relationship among sets (or events).

↳ made up of:

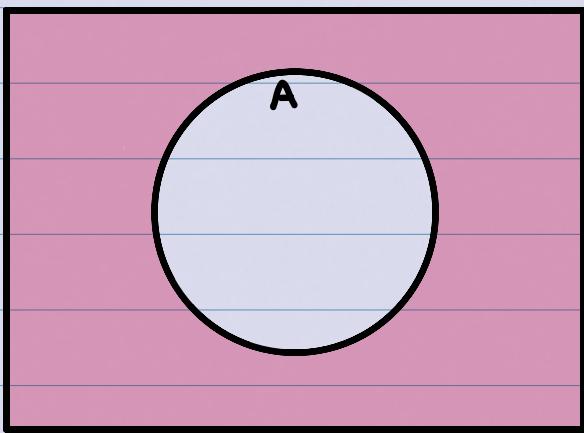
- a rectangle to represent the sample space S.
- Circles within the rectangle to illustrate the events.

## The Union and Intersection of Two Events



## The Compliment of an Event:

$\bar{A}$



"S but not A"

De Morgan's Laws:

$$\cdot \overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\cdot \overline{A \cap B} = \bar{A} \cup \bar{B}$$

Probability Rules

- Probability of the Complement of an Event
    - $P(A) = 1 - P(\bar{A})$
- Proof:  $A \cup \bar{A} = S$ , so  $P(A \cup \bar{A}) = P(A) + P(\bar{A})$   
since  $P(A \cup \bar{A}) = P(S) = 1$ ,  $1 = P(A) + P(\bar{A})$   
bc mutually exclusive!

Example: an experiment involves tossing a coin four times. The event A is "at least one head". Find the probability of A, ie, find  $P(A)$ .

- the only case where there is NOT at least one head is  $\bar{A} = \{\text{TTTT}\}$ .

flipping a coin 4 times would result in  $2^4$  possible

outcomes, and  $\bar{A}$  is only 1 of these.

$$\therefore P(\bar{A}) = \frac{1}{16}, \text{ so } P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{16} = \frac{15}{16}.$$

probability of complement!

Example: two ordinary dice are rolled. Find the probability that at least one lands on a 6.

$$S = \{(1,1), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,1), \dots, (6,6)\}.$$

$A$  = the event of at least one six.

$\therefore \bar{A}$  is the event that neither die is a six.

classical probability of each die to NOT be a six:  $\frac{n(x)}{N}$ .

where  $n(x)$  is the number of outcomes that satisfy the roll not being a 6 - ie;  $\{1, 2, 3, 4, 5\} \Rightarrow 5$ , and  $N$  is the total number of outcomes in the set - ie;  $\{1, 2, 3, 4, 5, 6\} \Rightarrow 6$ .

$$\therefore \frac{n(x)}{N} = \frac{5}{6} !$$

since both dice are independent, the probability of both NOT being a six (aka  $\bar{A}$ ) is  $\frac{5}{6} \cdot \frac{5}{6}$ .

$$\therefore P(\bar{A}) = \frac{25}{36}, \text{ so } P(A) = 1 - \frac{25}{36} = \frac{11}{36} !$$

Rules for Union:

• Union of two events:  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Example: At a university, 70% of the students own a car, 60% of students live on campus, and 50% both live on campus and own a car. If a student is chosen at random, what is the probability that the student does not own a car and does not live on campus?

- Let A be the event that a student has a car.
- Let B be the event that a student lives on campus.

we want:  $P(\bar{A} \cap \bar{B})$ .

↳ we know that  $\bar{A} \cap \bar{B} = \overline{A \cup B}$ .

$$\hookrightarrow P(\bar{A} \cap \bar{B}) = 1 - P(\overline{A \cup B}) = 1 - P(A \cup B)$$

from the union of two events,  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$$\begin{aligned}\therefore P(\bar{A} \cap \bar{B}) &= 1 - (P(A) + P(B) - P(A \cap B)) \\ &= 1 + P(A \cap B) - P(A) - P(B) \\ &= 1 + 0.5 - 0.7 - 0.6\end{aligned}$$

$$\hookrightarrow \therefore P(\bar{A} \cap \bar{B}) = 0.2 = 20\%.$$

Union of Three Events: for any 3 events A, B, and C:



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

**Conditional Probability:** determining the probability of some event A, while knowing that some related event B has occurred.

**Example:** consider that a fair die has been rolled and you are asked to give the probability that it was a 5.

$$\hookrightarrow A = \{5\}, S = \{1, 2, 3, 4, 5, 6\}.$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{1}{6}.$$

However, if you're told that the dice has landed on an odd number, the odds change:

$$\hookrightarrow A = \{5\}, S = \{1, 3, 5\}.$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{1}{3}.$$

**Example:** a box contains 100 microchips, some produced by Factory 1 and the rest by Factory 2. Among these microchips, some are defective, while others are good. Let A be the event of "obtaining a defective microchip" and B be the event of "the microchip was produced in Factory 2."

the number of microchips in each category:

	B	$\bar{B}$	totals
A	15	5	20
$\bar{A}$	45	35	80
totals	60	40	100

a) Find the probability of obtaining a defective microchip.

the total of A is 20, and there are 100 total microchips.

$$\hookrightarrow P(A) = \frac{n(A)}{N} = \frac{20}{100} = \frac{1}{5} \quad (20\%)$$

b) Given that a microchip is from factory 1, find the probability that it is defective.

the sample space S shrinks from all 100 microchips to only the 60 made by factory 1 (event B).

$$\hookrightarrow \therefore P(A) = \frac{n(A)}{N} = \frac{15}{60} = \frac{1}{4} \quad (25\%)$$

Conditional Probability mathematical definition:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (P(B) > 0)$$

↪ "the probability of A given B"

$$\hookrightarrow 0 \leq P(A|B) \leq 1$$

- $P(A|B) = 1 - P(\bar{A}|B)$

Pretty much the same as  
the regular union of two events

- $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2 | B)$

**Multiplication Theorem of Probability:** a way to compute the joint occurrence of A and B:

$$\hookrightarrow P(A \cap B) = P(B) \cdot P(A|B) = P(A) \cdot P(B|A)$$

**Law of Total Probability:** if  $B_1, B_2, \dots, B_k$  is a collection of mutually exclusive events and  $B_1 \cup B_2 \cup \dots \cup B_k = S$ , then for any event A,

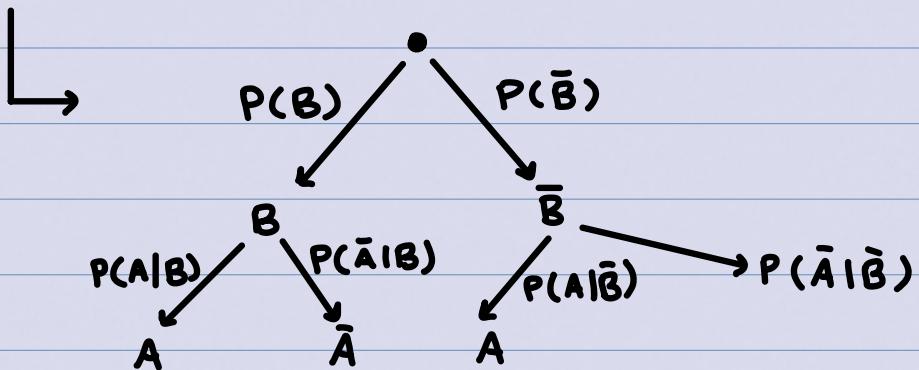
$$P(A) = \sum_{i=1}^k P(B_i) \cdot P(A|B_i)$$

Note: for a special case  $k=2$ :

only for  
 $k=2$  !!

$$P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) \cdot P(\bar{B})$$

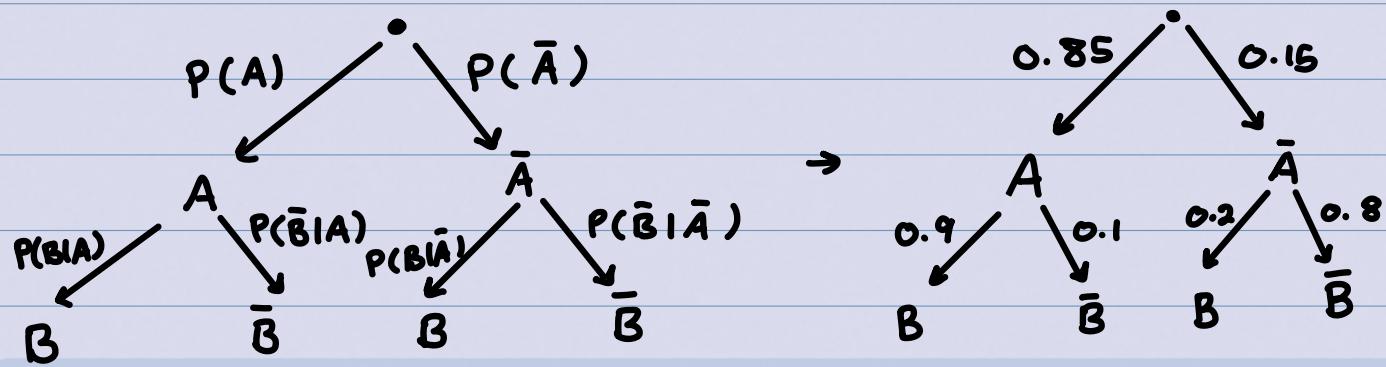
for event A to occur, it must happen with either B or  $\bar{B}$ .  $\therefore$  only  $A \cap B$  or  $A \cap \bar{B}$  can occur.



Example: without water, a plant will die with a

probability of 0.8, and with water, it will die with a probability of 0.1. I'll remember to water the plant with a probability of 0.85. Represent this info using a tree diagram.

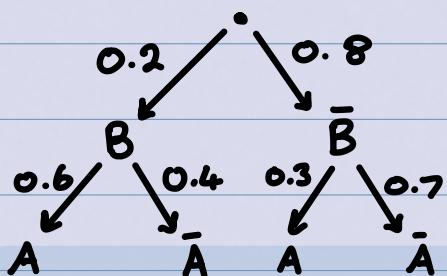
↳ let A be the event that I remember to water the plant, and B be the event that the plant survives.



Example: in a typical year, 20% of the days have a high temperature greater than  $22^{\circ}\text{C}$ . 40% of these days, there is no rain. During the rest of the year, when the temperature is  $\leq 22^{\circ}\text{C}$ , 70% of the days have no rain.

a) represent this information in a tree diagram.

let A be the event that it rains, and let B be the event that the temperature peaks above  $22^{\circ}\text{C}$ .



b) solve for the proportion of days in the year which have rain and a temperature  $\leq 22^{\circ}\text{C}$ .

↳ aka, find  $P(A \cap \bar{B})$

$$\text{we know } P(A|\bar{B}) = \frac{P(A \cap \bar{B})}{P(\bar{B})} = \frac{P(A \cap \bar{B})}{1 - P(B)}$$

$$\begin{aligned} \therefore P(A \cap \bar{B}) &= P(A|\bar{B}) \cdot (1 - P(B)) \\ &= 0.3 \cdot (1 - 0.2) = 0.3 \cdot 0.8 = 0.24 \\ &= 24\% . \end{aligned}$$

Baye's Rule:  $P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$ ,

and more generally,  $P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{i=1}^k P(B_i)P(A|B_i)}$  ↗ for each  $j=1, 2, \dots, k$ .

Example: revisiting the plant problem from before, if the plant is alive, what's the probability that I remembered to water it?

↗ same as previous time

↳ let A be the event that I remember to water the plant, and B be the event that the plant survives.

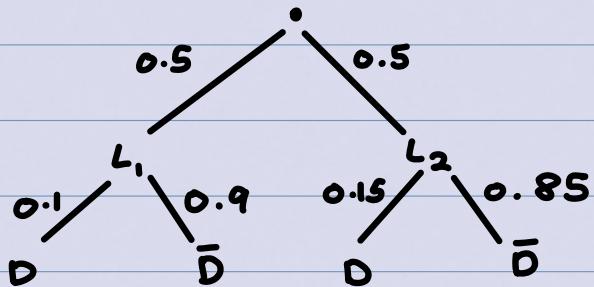
$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} = \frac{(0.9)(0.85)}{(0.9)(0.85) + (0.2)(0.15)} \\ &= 0.9623 = 96.23\% ! \end{aligned}$$

Example: electric motors from two assembly lines are

pooled in a common stockroom, with equal numbers of motors from each line. Motors are periodically sampled for testing. It's known that 10% of the motors from line 1 are defective, while 15% of the motors from line two are defective.

a) if a motor is randomly selected, what is the probability that it is defective?

Let D be the event that a motor is defective.



$$\begin{aligned}
 P(D) &= P(D \cap L_1) + P(D \cap L_2) \\
 &= 0.5 \cdot 0.1 + 0.5 \cdot 0.15 \\
 &= 0.125
 \end{aligned}$$

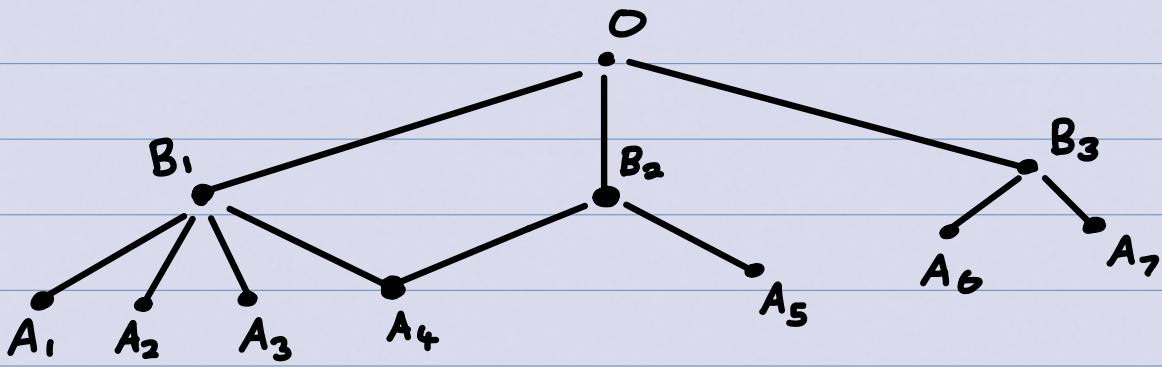
↳ 12.5% chance.

b) if the randomly selected motor is defective, what is the probability that it came from line 1?

$$P(L_1 | D) = \frac{P(D | L_1) P(L_1)}{P(D)} = \frac{0.1 \cdot 0.5}{0.125} = 0.4$$

= 40% chance.

Example: a man starts at point O on this map. He chooses a path at random and follows it to point B<sub>1</sub>, B<sub>2</sub>, or B<sub>3</sub>. From that point, he chooses a new path at random and follows it to one of the points A, ... A<sub>7</sub>.



a) what is the probability that the man arrives at point  $A_4$ ?

$$P(A_4) = P(A_4 | B_1) + P(A_4 | B_2) = \frac{1}{4} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12} + \frac{1}{6}$$

$$= \frac{1}{4} = 0.25$$

$= 25\%$  !

b) suppose that the man arrives at point  $A_4$ . What is the probability that he passed through  $B_1$ ?

$$P(B_1 | A_4) = \frac{P(A_4 | B_1)}{P(A_4)} = \frac{\frac{1}{4} \cdot \frac{1}{3}}{\frac{1}{4}} = \frac{\frac{1}{12}}{\frac{1}{4}} = \frac{4}{12} = \frac{1}{3}$$

$\therefore$ , there was an 33.33% chance he came from  $B_1$ .

**Independent Event:** events A and B are called independent if and only if  $P(A \cap B) = P(A)P(B)$ . Otherwise, A and B are called dependent events.

↳ an equivalent formulation can be given in terms of conditional probability :

→ If A and B are events such that  $P(A) > 0$  and  $P(B) > 0$ , then A and B are called independent if and only if either of the following hold:

$$P(A|B) = P(A), \text{ or } P(B|A) = P(B)$$

Note: independence  $\neq$  mutually exclusive!

unless  $P(A) = 0$ !

General rule:  $P(A \cap B) = P(B|A) \cdot P(A)$

Independence of Compliments: Suppose events A and B are independent. Then the following events are also independent:

- $\bar{A}$  and  $\bar{B}$
- $\bar{A}$  and B
- A and  $\bar{B}$

Independence of n events: the k events  $A_1, A_2, \dots, A_k$  are said to be independent, or mutually independent, if and only if:

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \dots P(A_{i_k})$$

forall possible distinct  $i_1, i_2, \dots, i_k$  chosen from 1, 2, ..., n.

→ Note: for a sequence of events to be independent, all of their subsets must also be independent.

Example: suppose a fair die is rolled twice. Let A be the event that the first roll is a 6, and let B be the event

that the second roll is a 6. Show that A and B are independent events.

↳ we want to show that  $P(A \cap B) = P(A) \cdot P(B)$

$$P(\{6,6\}) = P(A \cap B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

$$P(A) = \frac{n(A)}{N} = \frac{1}{6} = \frac{n(B)}{N} = P(B). \therefore P(A) = P(B) = \frac{1}{6}.$$

$$P(A) \cdot P(B) = \frac{1}{36} = P(A \cap B).$$

↳ ∴, the events are independent.

## Counting Techniques

often, manually counting many outcomes manually can be very time-consuming and tedious.

1) Addition Rule : "OR" interpreted as addition

↳ Eg: suppose job 1 can be done in p ways and job 2 can be done in q ways. ∴, we can do either job 1 or job 2 in  $p+q$  ways.

2) Multiplication Rule: "AND" interpreted as multiplication.

↳ Eg: suppose job 1 can be done in p ways and for each of these ways, we can do job 2 in q ways. ∴, we can do both Job 1 and Job 2 in  $p \times q$  ways.

Example: how many are there to answer a 20

## question true/false test?

$$Q_1 \cdot Q_2 \cdot \dots \cdot Q_{20} = 2^{20} \text{ ways.}$$

**Sampling with Replacement:** what we get on the first selection does not affect what we get on second selection.

↳ Eg: every time an object is selected, it's put back into the pool of possible objects.

**Sampling without Replacement:** what we get on the first selection will affect what we get on the second selection.

↳ Eg: once an object is selected, it stays out of the pool of possible objects.

**Example:** A bag contains 3 blue marbles and 5 red marbles. If you pick two marbles from the bag,

a) what is the probability of picking 2 blue marbles if the selection is with replacement?

let A be the event of picking two marbles.

↳  $P(A) = P(\text{blue on first pick}) + P(\text{blue on second pick})$

$$\hookrightarrow P(A) = \frac{3}{8} \cdot \frac{3}{8} = \frac{9}{64} = 14.06\%$$

b) what about without replacement?

$$P(A) = \frac{3}{8} \cdot \frac{2}{7} = \frac{6}{56} = \frac{3}{28} = 10.71\%$$

Example: a personal identification number (PIN) of length 4 is formed by randomly selecting 4 digits from the set of digits  $\{0, 1, 2, \dots, 9\}$ . If selection is done with replacement, find the probability that:

a) The PIN is even:

for the PIN to be even, the last digit must be a multiple of 2.  $\therefore A = \{0, 2, 4, 6, 8\}$ .

$$\therefore P(A) = \frac{n(A)}{N} = \frac{5}{10} = \frac{1}{2} \rightarrow 50\%$$

b) the PIN contains at least one 0?

$$\therefore P(\text{at least one } 0) = 1 - P(\text{no } 0s \text{ in PIN})$$

$\therefore$  there are 9 ways to fill each of the digits, as we want the probability of no zeros.

$$\therefore P(\text{no } 0s) = 9 \cdot 9 \cdot 9 \cdot 9 = 9^4 \quad \begin{matrix} \xrightarrow{\text{multiplication counting rule}} \\ \hookrightarrow \text{not } 0 \text{ AND not } 0 \text{ AND...} \end{matrix}$$

total sample space is  $10^4$  (including 0s).

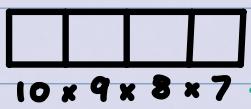
$$\therefore P(\text{at least one zero}) = 1 - \frac{n(\text{no } 0s)}{n(S)} = 1 - \frac{9^4}{10^4}$$

c) Redo a and b but without replacement:

i) even PIN:

A: event of getting an even number.  $\{0, 2, 4, 6, 8\}$ .

S: the sample space



we have less options  
for each digit, bc  
no replacement!

A:  $n(A) = 9 \times 8 \times 7 \times 5$   
 $9 \times 8 \times 7 \times 5 \rightarrow \text{bc } \text{len}(\{0, 2, 4, 6, 8\}) = 5$ .

$$\hookrightarrow P(A) = \frac{9 \times 8 \times 7 \times 5}{10 \times 9 \times 8 \times 7} = \frac{5}{10} = \frac{1}{2} = 50\%$$

ii) at least one zero:

$\curvearrowleft$  no 0 in PIN

$$P(B) = 1 - P(\bar{B}) = 1 - \frac{9 \times 8 \times 7 \times 6}{10 \times 9 \times 8 \times 7} = 1 - \frac{6}{10} = \frac{2}{5} = 40\%$$

Example: a fair die is tossed 3 times. What is the probability that exactly one of the tosses produces a number greater than 4?

let A be the event that the first toss is greater than 4, and the second and third tosses are  $\leq 4$ .

$$\hookrightarrow \{5, 6\}, \{1, 2, 3, 4\}, \{1, 2, 3, 4\}.$$

"AND"  
 $\therefore n(A) = 2 \times 4 \times 4 = 32$ .

let B be the event where the 2<sup>nd</sup> toss is  $> 4$ , and let C be the event where it's the 3<sup>rd</sup> toss.  $\rightarrow n(A) = n(B) = n(C)$

$\therefore$  total number of ways to get exactly one number greater than 4 is  $3(2 \times 4 \times 4) = 96$ .

$$\therefore P(\text{exactly one toss} > 4) = \frac{96}{6^3} = 44.44\%.$$

## Counting Permutations:

- Permutation: an ordered arrangement of a set of objects.

Example: how many different ordered arrangements of the letters a, b, and c are possible?

↳ abc, acb, bac, bca, cab, cba.

∴ 6 arrangements!

This will very quickly become very tedious and slow.

- The number of permutations of n distinguishable objects is:

$$n \times (n-1) \times \dots \times 1 = n!$$

- Given n distinct objects, a permutation of length r is an ordered subset of r objects.

- The number of permutations of length r taken from n objects is denoted as  $n^{(r)}$ , where:

$$n^{(r)} = n \times (n-1) \times \dots \times (n-r+1) = \frac{n!}{(n-r)!}$$

↑ "n to r factors"

Example: five separate awards are to be presented to some students from a class of size 30. How many different outcomes are possible if

a) a student can receive any number of awards?

$$\frac{30}{1^{\text{st}}} \times \frac{30}{2^{\text{nd}}} \times \frac{30}{3^{\text{rd}}} \times \frac{30}{4^{\text{th}}} \times \frac{30}{5^{\text{th}}} = 30^5.$$

$$\therefore n(A) = 30^5$$

b) each student can receive, at most, one award?

$$\frac{30}{1^{\text{st}}} \times \frac{29}{2^{\text{nd}}} \times \frac{28}{3^{\text{rd}}} \times \frac{27}{4^{\text{th}}} \times \frac{26}{5^{\text{th}}} = 30 \cdot 29 \cdot 28 \cdot 27 \cdot 26$$

$$\hookrightarrow n(A) = 30^{(5)} = \frac{30!}{25!}$$

Sometimes the outcomes in the sample space are subsets of a fixed size:

Example: suppose we have three books:  $b_1$ ,  $b_2$ , and  $b_3$ . We choose two of the books to read. In how many ways can the two books be read if:

a) the order that the books are read in matters:

$(b_1, b_2)$ ,  $(b_2, b_1)$

$(b_1, b_3)$ ,  $(b_3, b_1)$  → 6 ways!

$(b_2, b_3)$ ,  $(b_3, b_2)$

b) the order that the books are read in does not matter?

$$\{b_1, b_2\}$$

$$\{b_1, b_3\}$$

$$\{b_2, b_3\}$$

→ 3 ways!

Given  $n$  distinct objects, a combination of size  $r$  is an unordered subset of  $r$  objects chosen from  $n$  distinct objects.

↳ the number of combinations (or subsets) of size  $r$  chosen from  $n$  distinct objects is given by:

$$\binom{n}{r} = \frac{n^{(r)}}{r!} = \frac{n!}{(n-r)r!}$$

↳ "n choose r"

Example: we randomly select a subset of 3 digits from  $\{0, 1, \dots, 9\}$ . All the digits in each outcome are unique, but the order of the elements in a subset is not relevant.

Find the probability that:

a) All the digits in the selected subset are even:

$$S = \{\{0, 1, 2\}, \{0, 1, 3\}, \dots, \{7, 8, 9\}\}$$

$$n(S) = \binom{10}{3}.$$

↗ length of  $S$

the set of even digits is  $\{0, 2, 4, 6, 8\}$ .

let  $A$  be the event that all 3 digits are even.

$$\hookrightarrow n(A) = \binom{5}{3}$$

$$\hookrightarrow P(A) = \frac{n(A)}{n(S)} = \frac{\binom{5}{3}}{\binom{10}{3}} = \frac{1}{12} = 8.3\dot{3}\%$$

b) At least one of the digits in the selected subset is less than or equal to 5.

$$P(\text{at least one digit} \leq 5)$$

$$= 1 - P(\text{no digit} \leq 5)$$

$$= 1 - P(\text{all digits} > 5)$$

the set of digits  $> 5$ :  $\{6, 7, 8, 9\}$ .

$$= 1 - \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{29}{30} = 96.6\dot{6}\%$$

Example: there are 30 geese of which 6 were tagged. Later, 5 of the geese are randomly captured.

a) How many samples of 5 are possible?  $\binom{30}{5}$

b) How many samples of 5, which include 2 of the tagged geese, are possible?

$$\hookrightarrow \binom{6}{2} \cdot \binom{24}{3}$$

c) if the five captured geese represent a simple random sample drawn from the 30

geese, find the probability that:

i) two of the 5 captured geese are tagged.

$$P(A) = \frac{\binom{6}{2} \binom{24}{3}}{\binom{30}{5}} = 0.2130 = 21.3\%$$

ii) none of the 5 captured geese are tagged.

$$P(B) = \frac{\binom{24}{5}}{\binom{30}{5}} = 0.2983 = 29.83\%$$

### → Key Properties of Permutations and Combinations:

1)  $n^{(r)} = n \times (n-1)^{(r-1)}$  for  $r \geq 1$ .

2)  $\binom{n}{r} = \frac{n^{(r)}}{r!}$

3) Symmetry Property:  $\binom{n}{r} = \binom{n}{n-r}$  for  $r \geq 0$ .

4) since  $0! = 1$ ,  $\binom{n}{0} = \binom{n}{n} = 1$ .

5)  $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$

### 6) Binomial Theorem:

$$(1+x)^n = \binom{0}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \dots + \binom{n}{n}x^n \text{ for } x \in \mathbb{R}.$$

### Summary of Counting Techniques:

• Addition rule : OR (+)

• Multiplication rule : AND (×)

- Factorial:  $n!$  is the number of arrangements of  $n$  distinct objects where order matters.
- Permutation:  $n^{(k)}$  is the number of ways to choose  $k$  objects from  $n$  distinct objects where order matters.
- Combination:  $\binom{n}{k}$  is the number of ways to choose  $k$  objects from  $n$  objects where order does NOT matter.

Example: a person has 10 friends, of whom 6 will be invited to a party. How many choices are there if:

a) 2 of the friends will only attend together?

$$\hookrightarrow \binom{8}{4} \cdot \binom{2}{2}$$

b) 2 of the friends will not attend if the other is also attending?

$$\hookrightarrow \binom{8}{6} \cdot \binom{2}{0} + \binom{8}{5} \cdot \binom{2}{1}$$

OR (+)!  $\rightsquigarrow$  only one or the other can occur.

Example: 13 cards are picked at random from a standard deck of 52 cards without replacement. Find the probability that we pick:

a) at least one ace:

$\hookrightarrow A =$  the event we pick at least one Ace, so

$\bar{A} =$  the event we pick no aces

$$\hookrightarrow P(A) = 1 - P(\bar{A}) = 1 - \frac{\binom{48}{13}}{\binom{52}{13}}$$

b) 6 spades, 4 hearts, 2 diamonds, and one club:

$$\hookrightarrow \frac{\binom{13}{6} \cdot \binom{13}{4} \cdot \binom{13}{2} \cdot \binom{13}{1}}{\binom{52}{13}}$$

Example: consider drawing 3 numbers at random with replacement from the digits  $\{0, 1, \dots, 9\}$ . What is the probability that there is a repeated number among the 3?

$$S = 10 \cdot 10 \cdot 10 = 10^3$$

$A$  = event of repeated number

$\bar{A}$  = event of no repeated number

$\hookrightarrow 10 \cdot 9 \cdot 8$  options

$$\therefore P(A) = 1 - P(\bar{A}) = 1 - \frac{n(\bar{A})}{n(S)} = 1 - \frac{\frac{10^3}{10^3}}{10^3} = \frac{7}{25}.$$

Example: Melissa selects 7 numbers between 1 and 50, and then a computer does the same. What is the probability that at least 5 of Melissa's numbers match the computer's?

$$\hookrightarrow S = \binom{50}{7}, \text{ and Melissa must match 5, 6, or 7 of the numbers: } (\binom{7}{5})(\binom{43}{2}) + (\binom{7}{6})(\binom{43}{1}) + (\binom{7}{7})(\binom{43}{0})$$

$$\therefore P = \frac{(\binom{7}{5})(\binom{43}{2}) + (\binom{7}{6})(\binom{43}{1}) + (\binom{7}{7})(\binom{43}{0})}{\binom{50}{7}}$$

Example: a box contains 4 coins - 3 fair coins and 1 biased coin which lands on heads 80% of the time. A coin is randomly picked and tossed 6 times. It lands on heads 5 times. What is the probability the coin is fair?

let  $H_5$  be the event of getting 5 heads in 6 tosses.  
 let  $F$  be the event the coin is fair, and let  $B$   
 be the event the coin is biased.

$$P(F | H_5) = \frac{P(F \cap H_5)}{P(H_5)} = \frac{P(H_5 | F) P(F)}{P(H_5 | F) P(F) + P(H_5 | B) P(B)}$$

$$= \frac{6 \times 0.5^5 \times 0.5 \times \frac{3}{4}}{6 \times 0.5^5 \times 0.5 \times \frac{3}{4} + 6 \times 0.8^5 \times 0.2 \times \frac{1}{4}} = 0.417$$

Example: there are 4 passengers on a 5-floor elevator. What is the probability that:

a) the passengers all get off on different floors?

↳ 5 people getting off on 4 different floors  
 is  $5^{(4)}$ , out of total possibilities  $S = 5^4$ .

$$\therefore \frac{5^{(4)}}{5^4} = \frac{24}{125} = 19.2\%$$

b) 2 passengers get off on floor two and 2  
 get off on floor 3?

$$\frac{\binom{4}{2} \binom{2}{2}}{5^4}$$

c) 2 passengers get off one a floor and 2  
 passengers get off on a different floor?

$$\frac{\binom{4}{2} \binom{2}{2} \times \binom{5}{2}}{5^4}$$

**Random Variable:** a function, denoted by  $X$ , that assigns a real number  $x = X(\alpha)$  to each possible outcome  $\alpha$  in a sample space  $S$ .

aka, we say  $X$  is a random variable if  $X: S \rightarrow \mathbb{R}$ .

- Random variables are denoted by capital letters such as  $X$ ,  $Y$ , and  $Z$ .
- Their possible values/realizations are denoted by lowercase letters such as  $x$ ,  $y$ , and  $z$ .

**Example:** Consider an experiment where a fair coin is tossed 3 times.

$$S = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}, \text{TTH}, \text{THT}, \text{HTT}, \text{TTT}\}$$

For each outcome  $\alpha$  in  $S$ , the value of the function  $X(\alpha)$  is the number of heads corresponding to  $\alpha$ .

↳ Eg: if  $\alpha = \text{THH}$ , then  $X(\alpha) = 2$ .

∴ Value of $X$	Definition of the event
$X = 0$	$\{\text{TTT}\}$
$X = 1$	$\{\text{TTH}, \text{THT}, \text{HTT}\}$
$X = 2$	$\{\text{HHT}, \text{HTH}, \text{THH}\}$
$X = 3$	$\{\text{HHH}\}$

↳ basically, if in the naturals

**Discrete Variable:** if a random variable's range is a discrete subset of  $\mathbb{R}$  (a finite set  $\{x_1, \dots, x_n\}$  or a countably infinite set  $\{x_1, x_2, \dots\}$ ).

eg: number of coin flips, number of people, ...

eg: monthly rainfall, time, ...

**Continuous Variable:** if a random variable's range is an interval that is a subset of  $\mathbb{R}$ .

also referred to as discrete probability density function (discrete pdf)

**Probability Mass Function (pmf / pf):** the probability function of a discrete random variable  $X$  is:

$$\rightarrow P\{\omega \in S : X(\omega) = x\}$$

$$f_x(x) = P(X=x)$$

**Probability distribution of  $X$ :** the set of pairs  $\{(x, f(x)) : x \in A\}$ .

**Example:** a fair coin is tossed 3 times.  $X(a)$  is the number of heads corresponding to  $a$ .  
Find the probability distribution of  $a$ .

$$\hookrightarrow f(0) = P(X=0) = P(\{\text{TTT}\}) = 1/8$$

$$f(1) = P(X=1) = P(\{\text{TTH}, \text{HTH}, \text{HTT}\}) = 3/8$$

$$f(2) = P(X=2) = P(\{\text{HHT}, \text{HTH}, \text{THH}\}) = 3/8$$

$$f(3) = P(X=3) = P(\{\text{HHH}\}) = 1/8$$

$$\hookrightarrow f(0) + f(1) + f(2) + f(3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{8}{8} = 1 !$$

A function  $f(x)$  is a discrete pf if and only if it satisfies both of the following properties:

$$1) 0 \leq f(x) \leq 1 \text{ for all } x \in \mathbb{R}$$

$$2) \sum_{\text{all } x} f(x) = 1$$

Probability Histogram: the graph of a probability mass function of a discrete random variable.

Example: the random variable  $x$  has a pmf:

$x$	0	1	2	3	4
$f(x)$	$0.1c$	$0.2c$	$0.5c$	$c$	$0.2c$

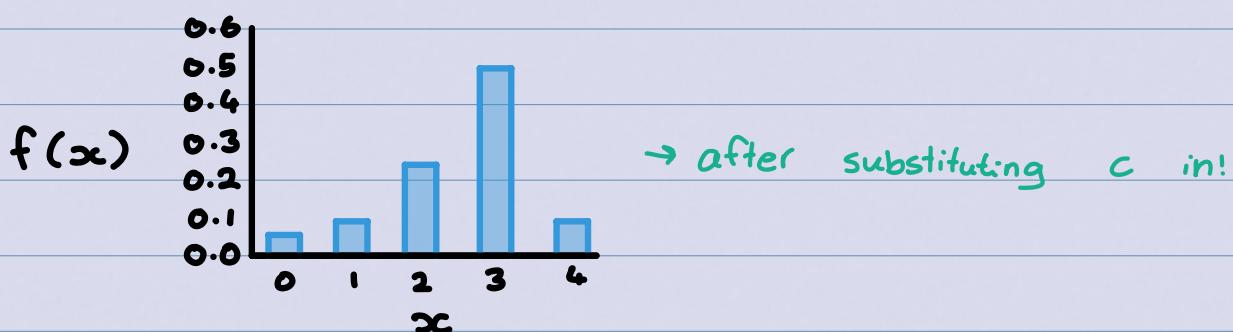
a) Find the value of  $c$ :

$$\hookrightarrow \sum f(x) = 1 \rightarrow 0.1c + 0.2c + 0.5c + c + 0.2c = 2c.$$
$$\therefore 2c = 1, \text{ so } c = \frac{1}{2} = 0.5.$$

b) Find  $P(x > 2)$

$$\begin{aligned} P(x > 2) &= P(x = 3) + P(x = 4) \\ &= c + 0.2c = 1.2c \\ &\Rightarrow 1.2c = 0.6 \end{aligned}$$

c) Plot the probability histogram of  $f(x)$



Cumulative Distribution Function (cdf): the cdf

of a random variable  $x$  is defined for any real  $x$  by:

$$F(x) = P(X \leq x)$$

If  $X$  is a discrete random variable with probability function  $f(x)$ , then:

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

Example: find the cdf of random variable  $x$  from the previous example:

$$F_x(x) = \begin{cases} 0 & x < 0 \\ 0.05 & 0 \leq x < 1 \\ 0.15 & 1 \leq x < 2 \\ 0.4 & 2 \leq x < 3 \\ 0.9 & 3 \leq x < 4 \\ 1 & 4 \leq x \end{cases}$$

- For discrete random variables, the CDF  $F(x)$  is represented as a step function.

(General) Example: Consider rolling two fair six-sided dice, and let the random variable  $X$  be the minimum of the rolls. What is  $f_x(2)$ ?

$$\hookrightarrow f_x(2) = P(X=2)$$

$$= P(\{(2,2), (2,3), (2,4), (2,5), (2,6), (3,2), (4,2), (5,2), (6,2)\})$$

$$\hookrightarrow P(X=2) = \frac{9}{36}.$$

Properties of Cumulative Distribution Functions:

↪ let  $F(\cdot)$  be a cdf. Then:

- 1)  $0 \leq F(x) \leq 1 \quad \forall x \in \mathbb{R}$
- 2)  $F(x) \leq F(y)$  for  $x < y \rightarrow F(x)$  is a non-decreasing function of  $x \quad \forall x \in \mathbb{R}$ .
- 3)  $\lim_{h \rightarrow 0^+} F(x+h) = F(x) \rightarrow F(x)$  is continuous on the right
- 4)  $\lim_{x \rightarrow -\infty} F(x) = 0$  and  $\lim_{x \rightarrow \infty} F(x) = 1$ .

Example: Students A, B, and C each independently answer a question. The probability of getting the correct answer is 0.9 for A, 0.7 for B, and 0.4 for C. Let  $X$  be the number of people who get the answer correct.

a) Compute the probability function of  $X$ :

Since the events are independent,  $P(ABC) = P(A)P(B)P(C)$

↓

$$f(0) = P(X=0) = P(\bar{A}\bar{B}\bar{C}) = 0.1 \times 0.3 \times 0.6 = 0.018$$

$$\begin{aligned} f(1) &= P(X=1) = P(A\bar{B}\bar{C}) + P(\bar{A}B\bar{C}) + P(\bar{A}\bar{B}C) \\ &= 0.9 \times 0.3 \times 0.6 + 0.1 \times 0.7 \times 0.6 + 0.1 \times 0.3 \times 0.4 \\ &= 0.216 \end{aligned}$$

$$f(2) = P(X=2) = P(AB\bar{C}) + P(A\bar{B}C) + P(\bar{A}BC)$$

$$= 0.9 \times 0.7 \times 0.6 + 0.9 \times 0.3 \times 0.4 + 0.1 \times 0.7 \times 0.4$$

$$= 0.514$$

$$f(3) = P(x=3) = P(ABC) = 0.9 \times 0.7 \times 0.4 = 0.252$$

$$\hookrightarrow f_x(x) = \begin{cases} 0.018 & x=0 \\ 0.216 & x=1 \\ 0.514 & x=2 \\ 0.252 & x=3 \end{cases}$$

b) compute the cdf of  $X$

$$\hookrightarrow F_x(x) = \begin{cases} 0.018 & x < 0 \\ 0.234 & 0 \leq x < 1 \\ 0.748 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

c) compute the probability that at least one person gets the answer correct.

$$\hookrightarrow P(x \geq 1) = P(x=1) + P(x=2) + P(x=3)$$

$$= 0.216 + 0.514 + 0.252 = 0.982.$$

$$\text{OR, } P(x \geq 1) = 1 - P(x=0) = 1 - 0.018 = 0.982.$$

$\cdot F(x)$  can be obtained from  $f(x)$ , and vice versa!



Example: the random variable  $x$  has a cdf given by:

$x$	1	2	3	4
$F(x)$	0.2	0.5	0.8	1

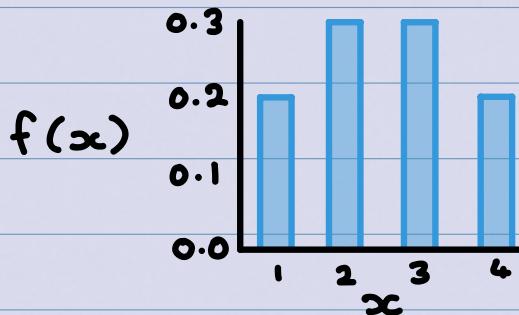
a) Find  $f(3)$ .

$$\hookrightarrow f_x(3) = F_x(3) - F_x(2) = 0.8 - 0.5 = 0.3$$

b) Calculate  $P(1 < x \leq 3)$ .

$$\begin{aligned} \hookrightarrow P(1 < x \leq 3) &= P(x \leq 3) - P(x \leq 1) \\ &= F(3) - F(1) = 0.8 - 0.2 = 0.6. \end{aligned}$$

c) Plot  $f(x)$ .



paka, "mean of  $x$ " or "first moment of  $x$ "

**Expected Value:** suppose  $X$  is a discrete random variable with probability function  $f_x(x)$ . Then,  $E(X)$  is called the expected value of  $X$  and is defined by :

$$E(X) = \sum_{x \in A} x \cdot f_x(x)$$

**Example:**  $X$  denotes the outcome of one fair six-sided die roll. Compute  $E(X)$ .

paka, the odds of  $x=1, \dots, 6 = 1/6$ .

we know  $f(1) = f(2) = \dots = f(6) = 1/6$

$$\therefore E(X) = \sum_{x \in \{1, 2, \dots, 6\}} x \cdot f(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5.$$

## Note(s):

- 1) The expected value of  $X$  may be a value that  $X$  can never actually take.
- 2) both  $\mu$  and  $E(x)$  refer to the same quantity which is the expected value of random variable  $X$ .

Example: a random variable  $X$  only takes two values, 0 or 1. If  $P(X=0) = 0.4$ , find  $E(x)$ .

$$P(X=1) = \overline{P(X=0)} = 1 - P(X=0) = 1 - 0.4 = 0.6$$

$$E(x) = \sum_{x \in \{0,1\}} x \cdot f(x) = 0 \times 0.4 + 1 \times 0.6 = 0.6.$$

Example: consider rolling 2 fair dice. Let  $x$  be the sum of the rolls, and let  $Y = X \pmod{4}$ . Find  $Y$ 's pmf.

→ we know  $S = \{(1,1), (1,2), \dots, (2,1), (2,2), \dots, (6,1), (6,2), \dots, (6,6)\}$   
 $\therefore A = \{2, 3, 4, \dots, 12\}$

$x$	$f_x(x)$	$y = x \pmod{4}$
2	$1/36$	2
3	$2/36$	3
4	$3/36$	0
5	$4/36$	1
6	$5/36$	2
7	$6/36$	3
8	$5/36$	0
9	$4/36$	1
10	$3/36$	2
11	$2/36$	3
12	$1/36$	0

$$f_y(0) = P(Y=0) = P(X=4) + P(X=8) + P(X=12) \\ = \frac{3}{36} + \frac{5}{36} + \frac{1}{36} = \frac{9}{36} = \frac{1}{4}.$$

$$f_y(1) = P(Y=1) = P(X=5) + P(X=9) \\ = \frac{4}{36} + \frac{4}{36} = \frac{8}{36} = \frac{2}{9}$$

$$f_y(2) = P(Y=2) = P(X=2) + P(X=6) + P(X=10) \\ = \frac{1}{36} + \frac{5}{36} + \frac{3}{36} = \frac{9}{36} = \frac{1}{4}$$

$$f_y(3) = P(Y=3) = P(X=3) + P(X=7) + P(X=11) \\ = \frac{2}{36} + \frac{6}{36} + \frac{2}{36} = \frac{10}{36} = \frac{5}{18}$$

$\forall y \notin \{0, 1, 2, 3\}, f_y(y) = 0.$

• General Formula when  $Y = aX + b$

↳ in the case of a linear function of the form

$Y = g(x) = ax + b$ , where  $a \neq 0$ :

$$f_y(y) = f_x\left(\frac{y-b}{a}\right)$$

• General Formula when  $Y = g(x)$

↳ Let  $Y = g(x)$  where  $g(\cdot)$  is an injective function.  
Since  $g(\cdot)$  is injective, its inverse  $g^{-1}(\cdot)$  exists.

aka one-to-one

Then, we see that:

$$\begin{aligned}
 f_Y(y) &= P(Y = y) \\
 &= P(g(X) = y) \\
 &= P(X = g^{-1}(y)) \\
 &= f_X(g^{-1}(y))
 \end{aligned}$$

$$\therefore f_Y(y) = f_X(g^{-1}(y))$$

Let  $A$  be a discrete random variable with the range  $A$  and pmf  $f(x)$ . The expected value of some function  $g(\cdot)$  of  $X$  is given by:

$$E[g(x)] = \sum_{x \in A} g(x) \cdot f(x)$$

Example: a TV station sells 15, 30, and 60 second advertising spots.  $X$  is the length of a random commercial and suppose that its probability distribution is given by:

$x$	15	30	60
$f(x)$	0.1	0.3	0.6

a) Find  $E(x)$ .

$$\hookrightarrow E(x) = \sum_{x \in \{15, 30, 60\}} x \cdot f(x) = 15 \times 0.1 + 30 \times 0.3 + 60 \times 0.6 = 46.5 \text{ seconds.}$$

b) if a 15s spot is \$500, a 30s spot is \$800, and a 60s spot is \$1000, find the average amount paid for a commercial on this station.

$$\hookrightarrow E(\text{amount paid}) = 500 \times 0.1 + 800 \times 0.3 + 1000 \times 0.6 = \$890.$$

Example: Consider the discrete random variable  $X$  with pmf of the form  $f_X(x) = \frac{1}{3}$ , for  $x = -1, 0, 1$ . If we define the random variable  $Y = X^2$ , calculate  $E(Y)$ .

$$E(Y) = \sum_{x \in A} g(x) f(x) = \frac{1}{3}(-1)^2 + \frac{1}{3}(0)^2 + \frac{1}{3}(1)^2 = \frac{2}{3} !$$

### Linearity Property of Expectation

The expectation  $E(\cdot)$  is a linear operator, satisfying the properties:

1) For constants  $a$  and  $b$ ,

$$E[a g(x) + b] = a E[g(x)] + b$$

2) For constants  $a$  and  $b$  and for functions  $g_1$  and  $g_2$ ,

$$E[a g_1(x) + b g_2(x)] = a E[g_1(x)] + b E[g_2(x)]$$

• If  $g(x)$  is a linear function, then  $E[g(x)] = g[E(x)]$ .

This is not necessarily true if  $g(x)$  is non-linear!

Example: let random variable  $X$  have the following probability function:

$x$	1	2	3	4
$f(x)$	0.4	0.3	0.1	0.2

a) Find  $E(x)$ .

$$\hookrightarrow \sum_{x \in \{1, 2, 3, 4\}} x \cdot f(x) = 1 \times 0.4 + 2 \times 0.3 + 3 \times 0.1 + 4 \times 0.2 = 2.1$$

b) Find  $E(\frac{1}{x})$ .

$$\hookrightarrow \sum_{x \in \{1, 2, 3, 4\}} \frac{1}{x} \cdot f(x) = \frac{1}{1} \times 0.4 + \frac{1}{2} \times 0.3 + \frac{1}{3} \times 0.1 + \frac{1}{4} \times 0.2 = 0.633.$$

• Variance: the variance of a random variable  $X$  is:

$$\sigma^2 = \text{Var}(X) = E[(X - \mu)^2]$$

aka  $\mu!$

↳ can also be calculated as:  $\sigma^2 = \text{Var}(X) = E(X^2) - E(X)^2$

Standard Deviation: the standard deviation of a random variable  $X$  is:

$$\sigma = \text{sd}(X) = \sqrt{\text{Var}(X)} = \sqrt{E[(X - \mu)^2]}$$

• Important Properties:

- for real constants  $a$  and  $b$ ,  $\text{Var}(aX + b) = a^2 \text{Var}(X)$
- if  $Y = aX + b$ , then
  - 1)  $E(Y) = aE(X) + b$
  - 2)  $\text{Var}(Y) = a^2 \text{Var}(X)$
  - 3)  $\text{sd}(Y) = |a| \text{sd}(X)$

Example:  $X$  has  $\text{Var}(X) = 2$ . Compute the variance of random variable  $Y$ , where  $Y = -2X + 3$ .

↳ we know that  $\text{Var}(Y) = a^2 \text{Var}(X) = (-2)^2 (2) = 8$ .

• Properties of Variance:

- 1) If  $X$  is a r.v.,  $\text{Var}(X) \geq 0$ .
- 2) big variances indicate that the probability distribution of  $X$  is more "spread out" around the mean.
- 3)  $\text{Var}(X) = 0$  if and only if  $P(X = E(X)) = 1$ .  
↳ no variance if the expected value is guaranteed!

Example:  $X$  is a random variable with probability function given by:

$x$	1	2	3	4	5
$f(x)$	0.1	0.1	0.3	0.35	0.15

If  $Y = 4 - 2X$ , find  $E(Y)$ ,  $\text{Var}(Y)$ , and  $\text{sd}(Y)$ :

$$\cdot E(X) = \sum_{x \in \{1, \dots, 5\}} x \cdot f_x(x) = 1 \cdot 0.1 + 2 \cdot 0.1 + 3 \cdot 0.3 + 4 \cdot 0.35 + 5 \cdot 0.15 = 3.35$$

$$E(Y) = 4 - 2E(X) = 4 - 2(3.35) = -2.7$$

$$\begin{aligned} E(X^2) &= \sum_{x \in \{1, \dots, 5\}} x^2 \cdot f_x(x) = 1^2 \cdot 0.1 + 2^2 \cdot 0.1 + 3^2 \cdot 0.3 + 4^2 \cdot 0.35 + 5^2 \cdot 0.15 \\ &= 12.55 \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - \mu^2 = E(X^2) - E(X)^2 = 12.55 - 3.35^2 \\ &= 1.3275 \end{aligned}$$

$$\text{Var}(Y) = (-2)^2 \text{Var}(X) = 4(1.3275) = 5.31$$

$$\text{sd}(Y) = \sqrt{\text{Var}(Y)} = \sqrt{5.31} = 2.3043.$$

$$\therefore E(Y) = -2.7, \text{Var}(Y) = 5.31, \text{and } \text{sd}(Y) = 2.3043!$$

**Discrete Uniform Distribution:** suppose the range of random variable  $X$  is  $\{a, a+1, \dots, b\}$ , where  $a$  and  $b$  are integers and all values are equally likely. Then,  $X$  is said to be a discrete uniform distribution on the set  $\{a, a+1, \dots, b\}$ . We may write  $X \sim U[a, b]$ , where  $a$  and  $b$  are the parameters of the distribution.

Example: if  $X$  is the number obtained from rolling a fair die, then  $X$  has a discrete uniform distribution over the set  $\{1, 2, \dots, 6\}$ . Here,  $a=1$ ,  $b=6$ , and  $X \sim U[1, 6]$ .

Suppose  $X \sim U[a, b]$ . This means that we have  $b-a+1$  possible values for  $X$ , which are all equally likely. Therefore, the probability function of  $X$  is:

$$f(x) = P(X=x) = \frac{1}{b-a+1} \text{ for } x=a, a+1, \dots, b !$$

otherwise,  $f(x)=0$