

# 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  $i = 1, 2, 3, 4, 5, 6$ .

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

↳ 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  $\boxed{P(A \cup B)} = P(A) + P(B)$

"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.

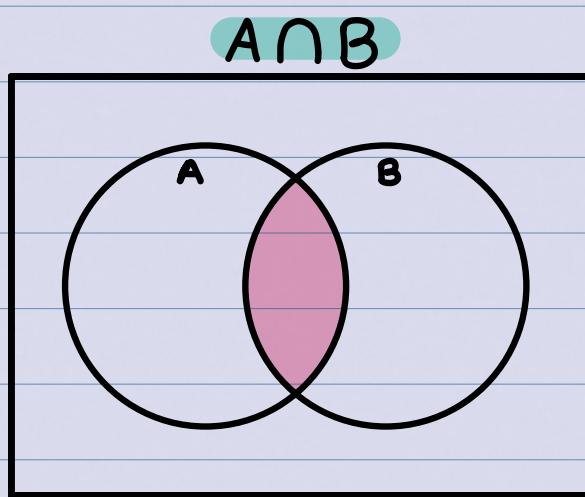
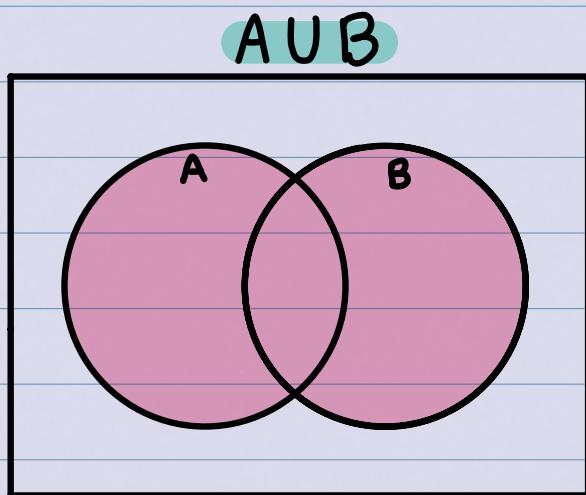
$$\hookrightarrow \therefore C = \{(1,5), (3,3), (5,1)\}.$$

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

↳ made up of:

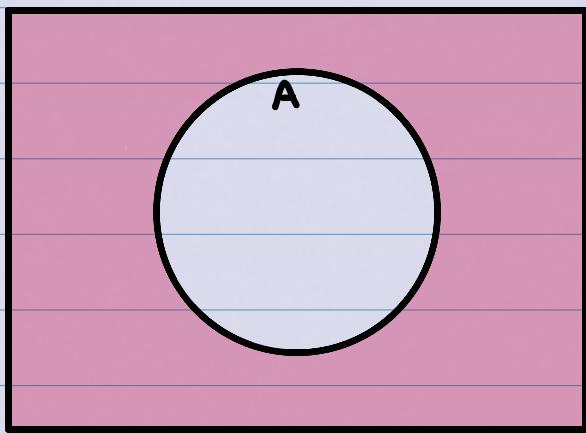
- a rectangle to represent the sample space S.
- Circles within the rectangle to illustrate the events.

### 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,

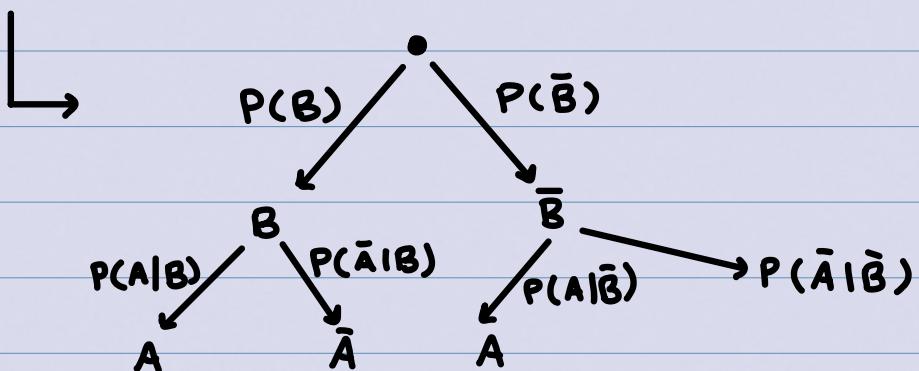
$$P(A) = \sum_{i=1}^k P(B_i) \cdot P(A|B_i)$$

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

only for  
 $k=2$  !!

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

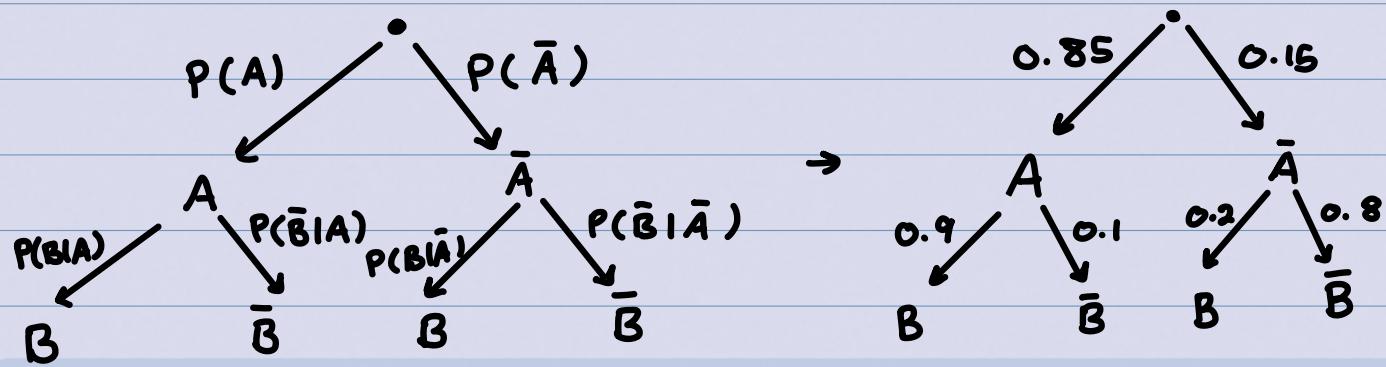
for event A to occur, it must happen with either B or  $\bar{B}$ .  $\therefore$  only  $A \cap B$  or  $A \cap \bar{B}$  can occur.



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.

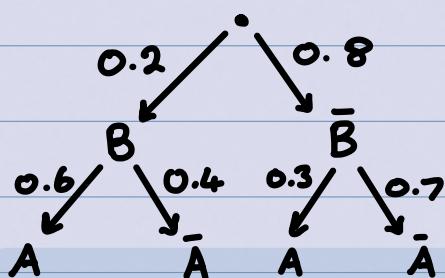
Let A be the event that I remember to water the plant, and B be the event that the plant survives.



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.

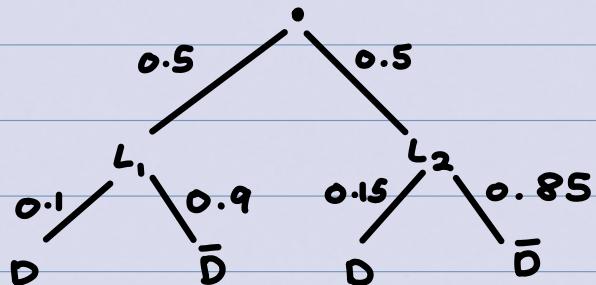
$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})} = \frac{(0.9)(0.85)}{(0.9)(0.85) + (0.2)(0.15)} \\ &= 0.9623 = 96.23\% ! \end{aligned}$$

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.



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

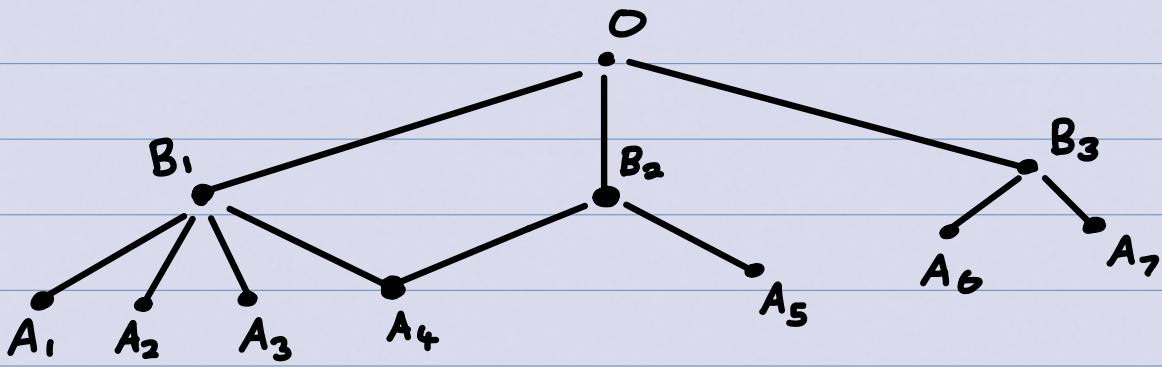
↳ 12.5% chance.

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  $B_1$ ?

$$P(B_1 | A_4) = \frac{P(A_4 | B_1) \cdot P(B_1)}{P(A_4)} = \frac{\frac{1}{4} \cdot \frac{1}{3} \cdot \frac{1}{3}}{\frac{1}{4}} = \frac{1/36}{1/4}$$

$$= \frac{1}{36} \times 4 = \frac{1}{9} = 0.11$$

$\therefore$ , there was an 11.11% chance he came from  $B_1$ .