WESTMINSTER
International University in Tashkent

WEEK 2
PROBABILITY TOPICS

Lecturer: Olmos Isakov

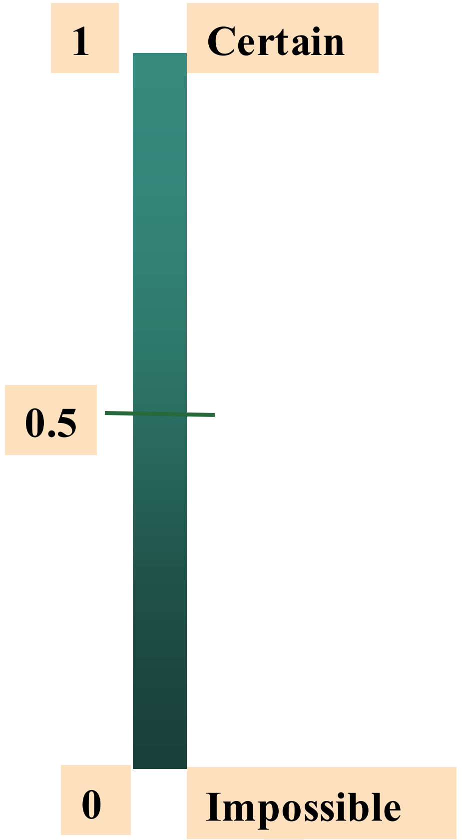
Office hours: Tuesdays, 09:00 – 11:00 (ATB 216)

AGENDA

- To understand basic probability concepts and approaches.
- To understand some rules of probability.
- To understand conditional probability.
- To learn Bayes’ theorem.
- To understand probability with permutations and combinations.

WHAT IS PROBABILITY?

Probability – the numerical measure of the likelihood that an event will occur.



The probability of any event must be between 0 and 1, inclusively

$$\boxed{0 \leq P(A) \leq 1 \quad \text{For any event } A}$$

Certain Event – an event that is sure to occur (probability = 1)

Impossible Event – an event that has no chance of occurring (probability = 0)

The sum of the probabilities of all mutually exclusive and collectively exhaustive events is 1.

$$P(A) + P(B) + P(C) = 1$$

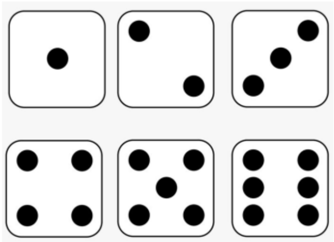
If A, B, and C are mutually exclusive and collectively exhaustive

APPROACHES TO PROBABILITY

WIUT

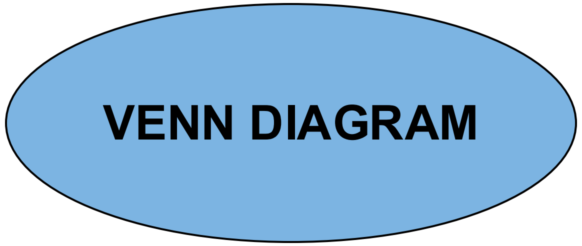
CLASSICAL	EMPIRICAL	SUBJECTIVE
It is based on the assumption that the outcomes of an experiment are equally likely	It is based on the number of times an event occurs as a proportion of a known number of trials	The probability of an event happening that is assigned by an individual based on whatever information is available
The probability of rolling a 4 with a 6-sided die	The probability of loan default if 650 out of 10000 loans defaulted last 5 years.	The probability of Uzbek national football team advancing to knockout stage in FIFA World Cup 2026.

EVENTS & SAMPLE SPACE



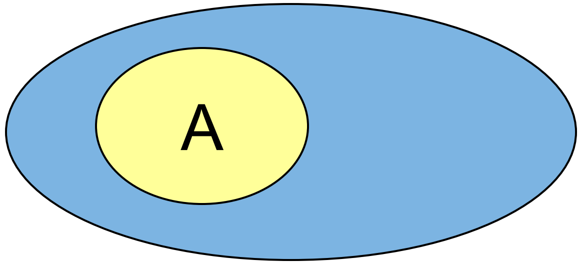
Consider an experiment

Sample space S:



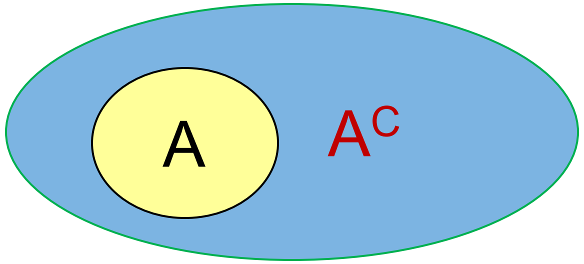
Example:
S={1, 2, 3, 4, 5, 6} rolling a die

Event A:



Example:
A={1,6} when rolling a die

Complementary event A^C (or A'):
$P(A) + P(A^C) = 1$



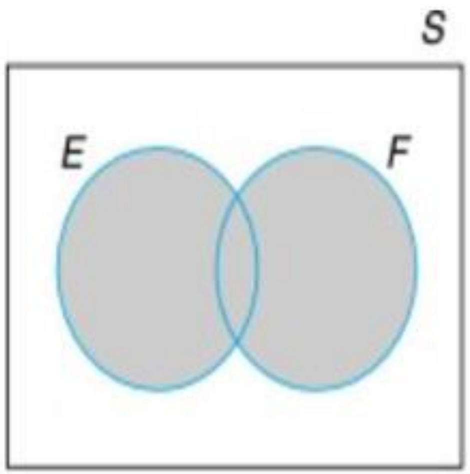
Example:
A^C={2,3,4,5} rolling a die

UNION AND INTERSECTION

The Union of Events

denoted by **E ∪ F (E OR F)**

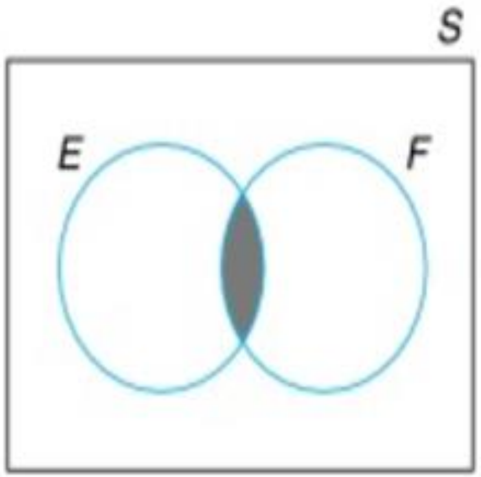
consists of all outcomes that are either in E or in F or in both E and F.



The Intersection of Events

denoted by **E ∩ F (E AND F),**

consists of all outcomes that are in both E and F



Example

You are walking at Shahrisabz street in Tashkent and randomly meet one person on the street.

Which of the following is the most likely event about this person?

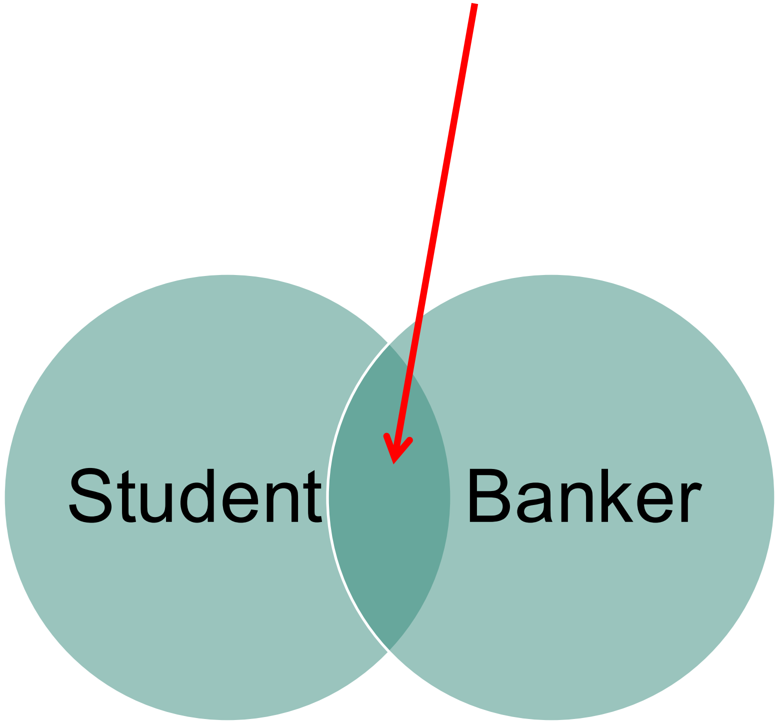
a. He is a WIUT student and a banker.

b. He is a banker.

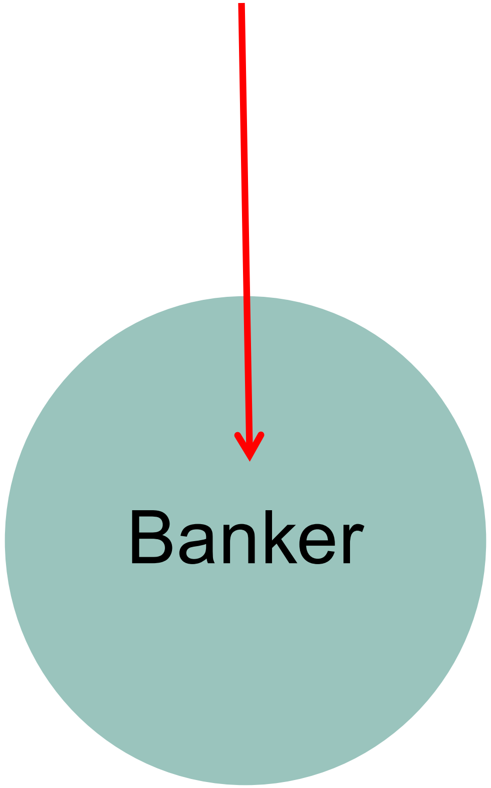
c. He is a WIUT student or a banker.

Example

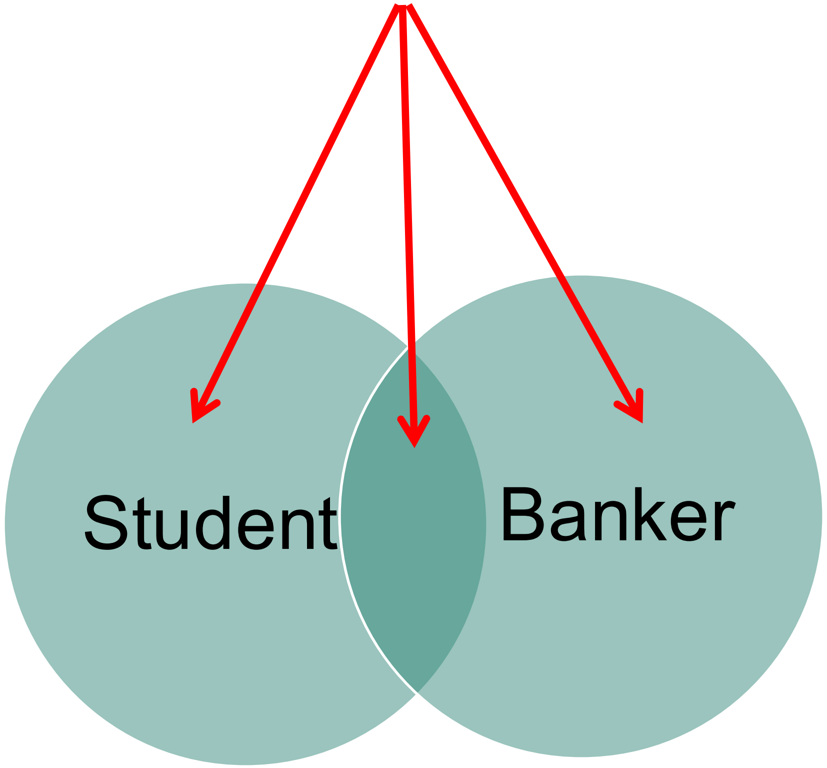
A ∩ B
WIUT Student **AND** Banker



Banker



A U B
WIUT Student **OR** Banker



ADDITIVE RULE

Additive Rule of Probability

The probability of the union of events A and B is the sum of the probability of event A and the probability of event B, minus the probability of the intersection of events A and B; that is

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

Ex: If **Event A**: Being a Wednesday in 2024 and **Event B**: Being a day in January 2024,

Find the probability of a day being either a Wednesday or in January 2024 **P(A ∪ B)**

	January	Not January	Total
Wednesday	$\frac{5}{366}$	$\frac{48}{366}$	$\frac{53}{366}$
Not Wednesday	$\frac{26}{366}$	$\frac{287}{366}$	$\frac{313}{366}$
Total	$\frac{31}{366}$	$\frac{335}{366}$	$\frac{366}{366}$

$$P(A \cup B = P(A) + P(B) - P(A \cap B) = \frac{31}{366} + \frac{53}{366} - \frac{5}{366} = \frac{79}{366}$$

MULTIPLICATIVE RULE

Multiplicative Rule of Probability

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

Ex: **If you randomly choose two days in 2024 without replacement,**

a. What is the probability that both days are Wednesday?
b. What is the probability that the first chosen day is Wednesday and second chosen day is not Wednesday?

Note:
There are 366 days in 2024
There are 53 Wednesdays in 2024
The rest 313 are not Wednesdays

MULTIPLICATIVE RULE

Ex: **If you randomly choose two days in 2024 without replacement,**

a. What is the probability that both days are Wednesday?
b. What is the probability that the first chosen day is Wednesday and second chosen day is not Wednesday?

Solution: There are 366 days and 53 Wednesdays in 2024

$P(1^{st}\text{: Wednesday}) = \frac{53}{366}$ ➔ After the 1st pick, there are 52 Wednesdays available to pick from 365 days

$P(2^{nd}\text{: Wednesday}) = \frac{52}{365}$

a. $P(\text{Wednesday} \cap \text{Wednesday}) = \frac{53}{366} * \frac{52}{365} = \frac{2\,756}{133\,590} = 0.0206 = 2.06\%$

b. $P(\text{Wednesday} \cap \text{Not Wednesday}) = \frac{53}{366} * \frac{313}{365} = \frac{16\,589}{133\,590} = 0.1242 = 12.42\%$

INDEPENDENT EVENTS

Events A and B are **independent events** if the occurrence of B does not alter the probability that A has occurred, that is **$P(A|B) = P(A)$** or **$P(B|A) = P(B)$.**

$P(A \cap B) = P(A)*P(B)$ ➔ Independent event

Whatever happens in one event has absolutely nothing to do with what will happen next, because:

- **Two events are unrelated**

  Ex: The weather in New York & your willingness to study FoS

- **You repeat an event with an item whose numbers will not change**

  Ex: Rolling the die or tossing the coin

- **You repeat the same activity but you replace the item that was previously chosen**

  Ex: If I am allowed to choose the same Wednesday that I chose in my first pick

Conditional Probability

A, B – events.

The **conditional probability** of event A, given that event B has occurred, is denoted by P(A | B) and is found by:

$$\boxed{\mathbf{P(A \mid B)} = \frac{\boldsymbol{P(A \cap B)}}{\boldsymbol{P(B)}}} \quad \textit{provided that } P(B) \neq 0$$

Similarly,

$$\boxed{\mathbf{P(B \mid A)} = \frac{\boldsymbol{P(A \cap B)}}{\boldsymbol{P(A)}}} \quad \textit{provided that } P(A) \neq 0$$

Example

You invite a group of friends to a traditional Uzbek meal, where you serve **Uzbek pilaf (plov)** and **fresh salad**. Based on previous experiences, you know that:

- **80%** of your friends like **Uzbek pilaf**.
- **60%** of your friends like **fresh salad**.
- **50%** of your friends like **both** Uzbek pilaf and fresh salad.
- **90%** of your friends like **either** Uzbek pilaf **or** fresh salad.

Calculate the **conditional probability** that a friend likes fresh salad **given that they like Uzbek pilaf**. Do you think whether it is higher than 60% or lower or the same?

Solution

Let's denote, P – pilaf, S – salad, then

$P(P) = 0.80$, $P(S) = 0.60$, $P(P \cap S) = 0.50$, $P(P \cup S) = 0.90$, $P(S \mid P) = ?$

$$P(S \mid P) = \frac{P(P \cap S)}{P(P)} = \frac{0.50}{0.80} = 0.625 \text{ or } 62.5\%$$

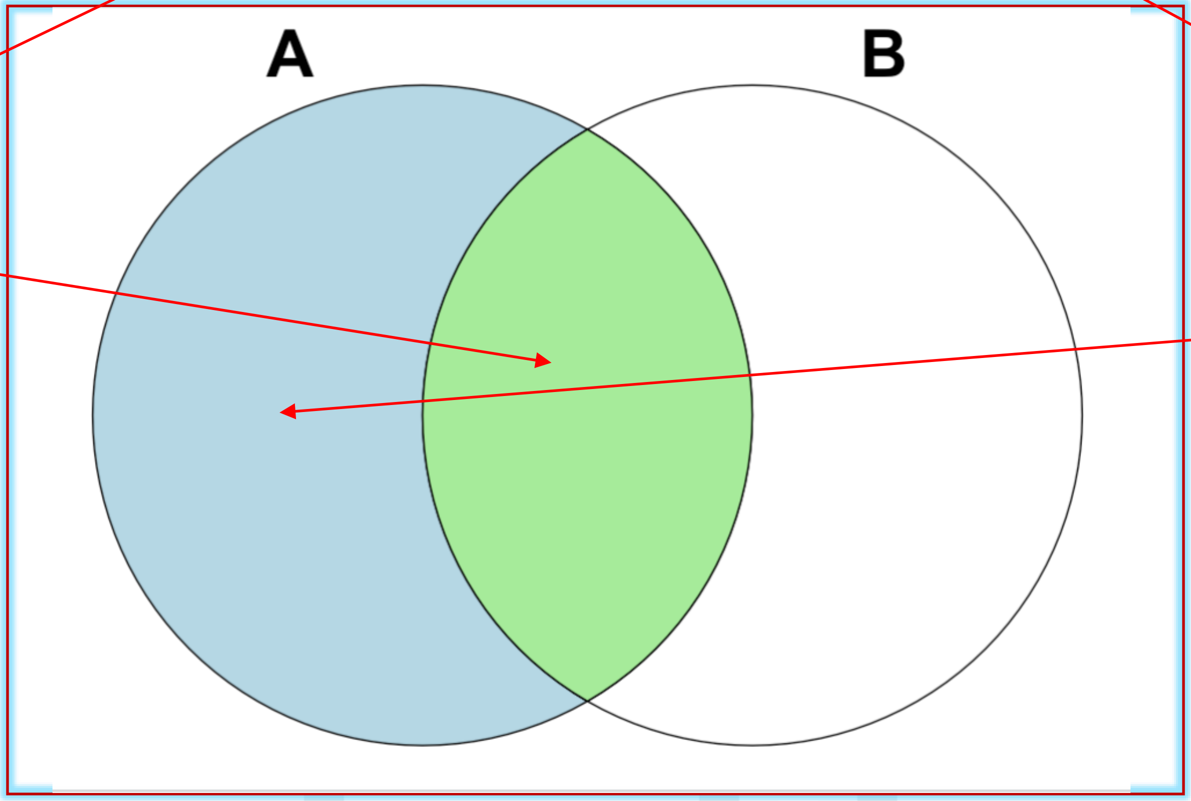
This means that if someone already likes Uzbek pilaf, there's a 62.5% chance they will also enjoy the fresh salad!

Total probability formula

$$P(A) = \boxed{P(A|B) * P(B)} + \boxed{P(A \mid B^C) * P(B^C)}$$

- B and B^c are mutually exclusive, and collectively exhaustive.

$P(A \cap B)$



$P(A \cap B^C)$

Example.

Suppose, 20% of bank clients are high-risk, and the remaining are low-risk. If the probability of default among high-risk clients is 0.15 (15% of high-risk borrowers default) and 0.01 for low-risk clients (1% of low-risk borrowers default), calculate the **overall probability** that any randomly selected borrower will default.

Total probability formula

$$P(A) = P(A|B) * P(B) + P(A \mid B^C) * P(B^C)$$

Solution.

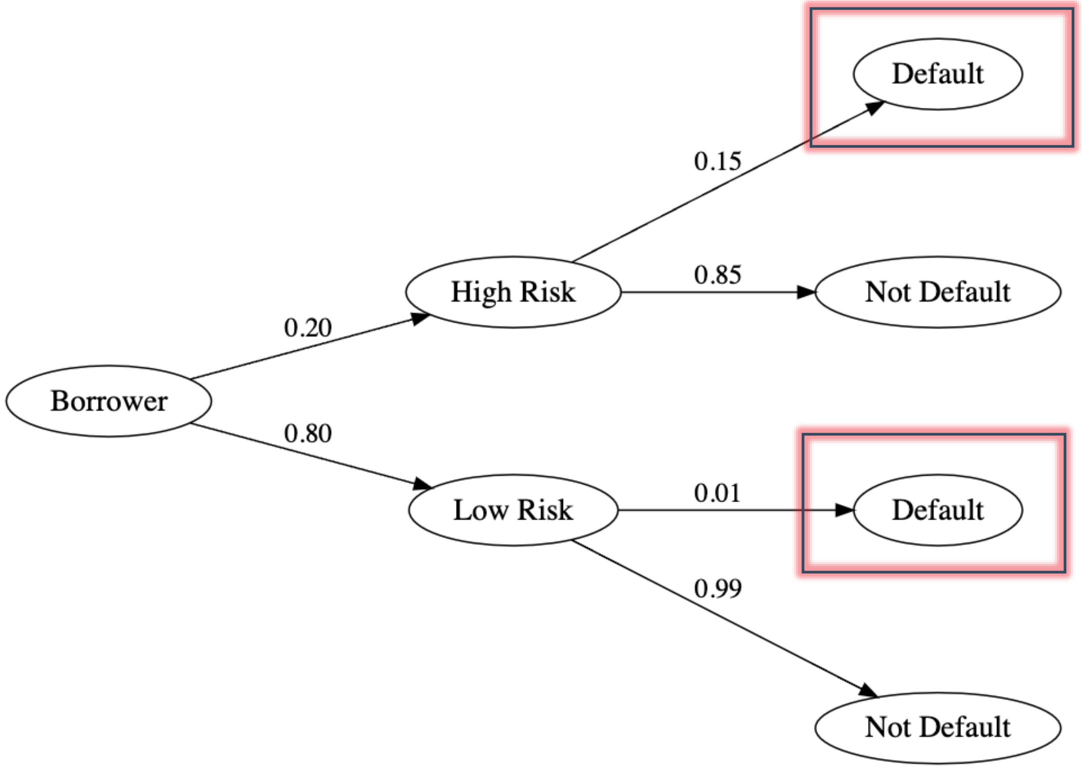
H – high-risk => $P(H) = 0.20$

H^C – low-risk => $P(H^C) = 1 - 0.20 = 0.80$

D – Default, $P(D \mid H) = 0.15$ and $P(D \mid H^C) = 0.01$

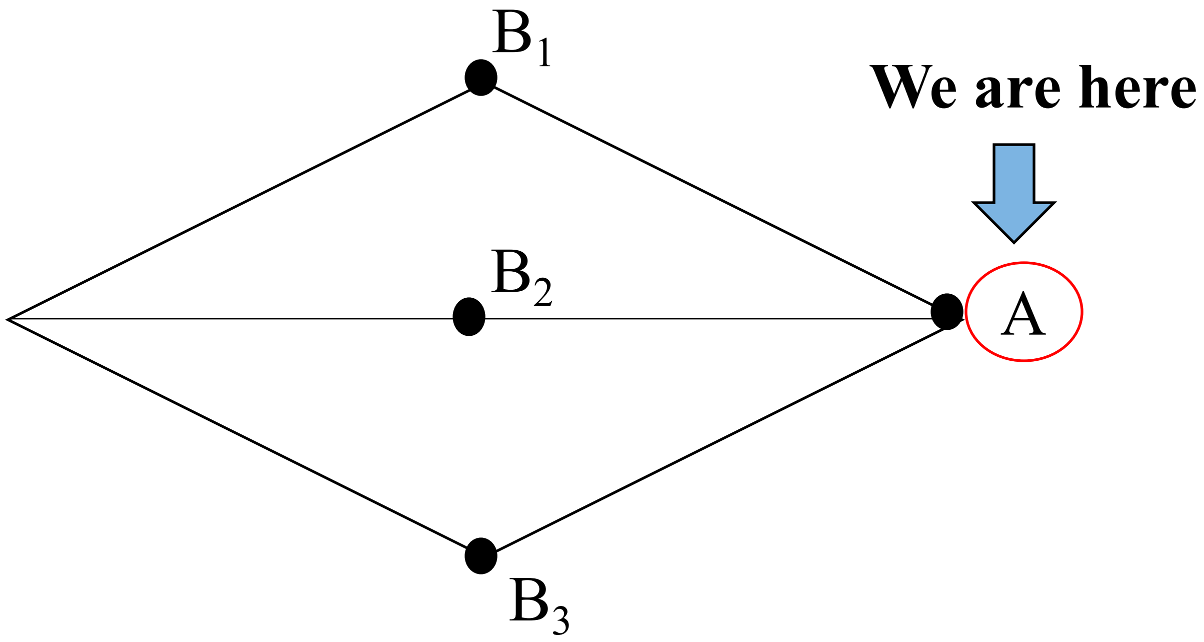
$P(D) = P(D \mid H) * P(H) + P(D \mid H^C) * P(H^C) =$
$0.15*0.20 + 0.01*0.80 = 0.038$

So, the overall probability that any borrower will default is **0.038 (or 3.8%).**

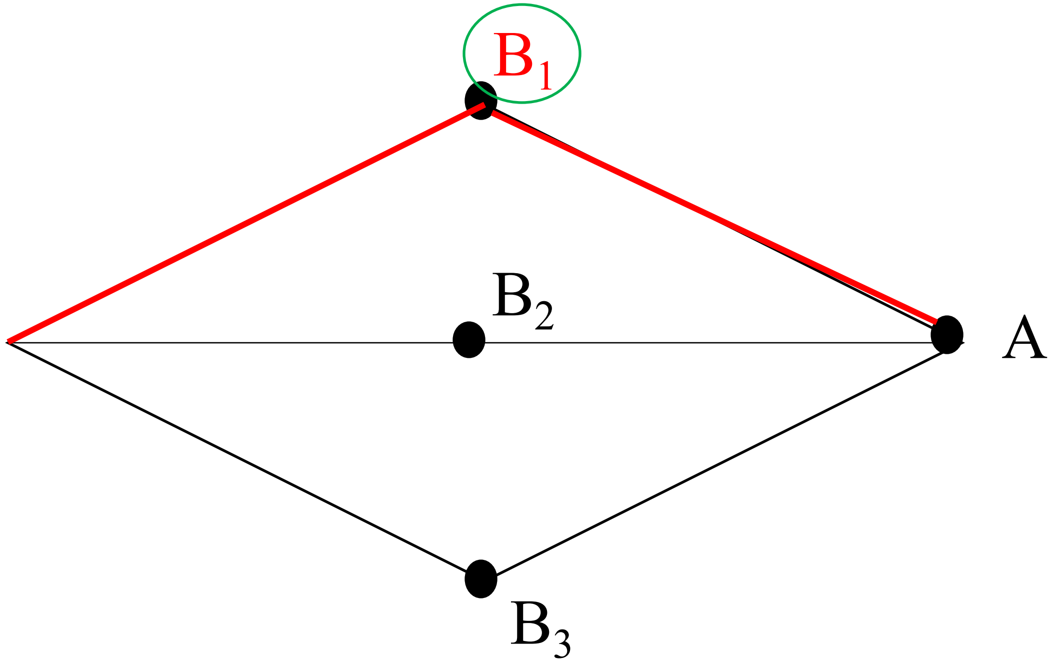


Bayes' theorem (or Bayes' rule)

Paths to A indicating that A has occurred.



What is the probability that we got there via B_1?



$$P(B_1|A) = \frac{P(A \cap B_1)}{P(A)}$$

Bayes' theorem (or Bayes' rule)

Definitions

A **prior probability** is an initial probability value originally obtained before any additional information is obtained.

A **posterior probability** is a probability value that has been revised by using additional information that is later obtained.

Bayes' Theorem

The probability of event A, given that event B has subsequently occurred, is

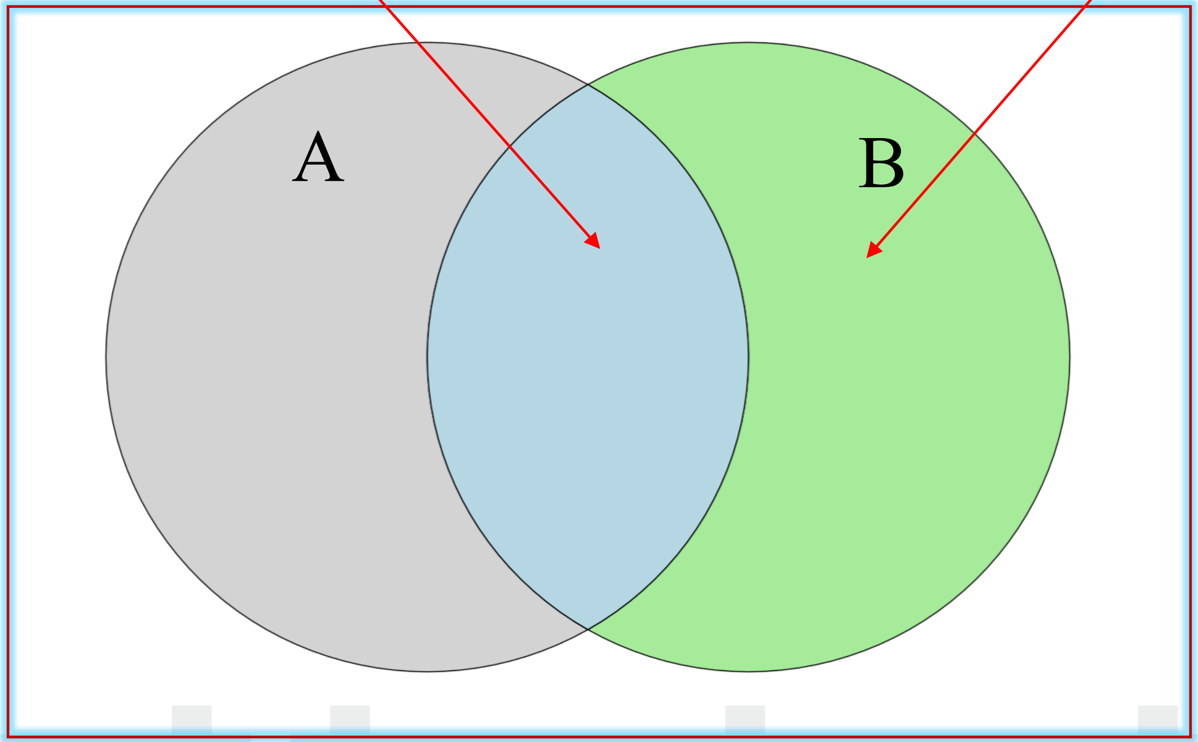
$$\mathbf{P(A|B)} = \frac{P\,(A\ \cap B)}{P(B)} = \frac{P(A) * P(B|A)}{P(B)} = \frac{\boldsymbol{P(A) * P(B|A)}}{\boldsymbol{P(A) * P(B|A) + P\left(A^C\right) * P(B|A^C)}}$$

Proof

$$P(A|B) = \frac{P(A)*P(B|A)}{P(B)} = \frac{P(A)*P(B|A)}{\boxed{P(A)*P(B|A)}+\boxed{P(A^C)*P(B|A^C)}}$$

$P(A \cap B)$

$P(A^C \cap B)$



Example

Imagine you are developing a spam filter for your company's email system, and you want to calculate the probability that a received email is actually spam, given that the email contains certain suspicious words like "win" or "free". You know that roughly 15% of all emails are spams. Approximately 80% of spam emails contain those suspicious words whereas this percentage is only 10% for legitimate emails.

a. Find the probability that the randomly selected email contains a suspicious word.
b. You later read that email and found some suspicious words. Use this additional information to find the probability that the selected email is a spam.

Solution

Let's use the following notation:

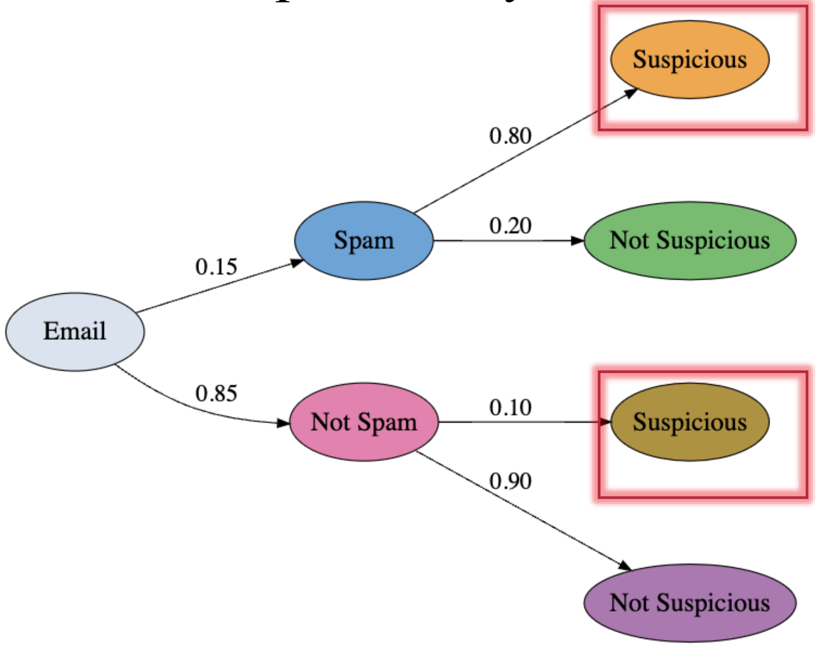
W – suspicious word

S – spam => $P(S) = 0.15$

S^C – not spam => $P(S^C) = 0.85$

prior probabilities

a. Find the probability that the randomly selected email contains a suspicious word.



= 0.15*0.80 = 0.12

$$P(W) = 0.15*0.80 + 0.85*0.10 = 0.205$$

= 0.85*0.10 = 0.085

Solution

b. We have the following:

$P(W) = 0.205$, $P(W|S^C) = 0.10$, $P(W|S) = 0.80$

Knowing that the email contains suspicious word, the probability of being a spam *(posterior probability):*

$$P(S|W) = \frac{P(S)*P(W|S)}{P(W)} = \frac{0.15*0.80}{0.205} = 0.585$$

Meaning: If there are suspicious words ”win” or “free” in the email, then there is a 58.5% chance that the selected email is spam.

Permutations & Combinations

	With replacement	Without replacement
Combinations (unordered)	${}_{n+r-1}C_r = \frac{(n+r-1)!}{r!*(n-1)!}$	${}_nC_r = \frac{n!}{r!*(n-r)!}$
Permutations (ordered)	n^r	${}_nP_r = \frac{n!}{(n-r)!}$

Example

A 4-digit PIN number is selected for the customer's debit card. What is the probability that there are no repeated digits?

Solution

$$\boxed{10} \times \boxed{10} \times \boxed{10} \times \boxed{10} = 10^4$$

There are 10 possible values for each digit of the PIN (namely: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9), so there are $10 \cdot 10 \cdot 10 \cdot 10 = 10^4 = 10000$ total possible PIN numbers.

To have no repeated digits, all four digits would have to be different, which is selecting without replacement. We could either compute $10 \cdot 9 \cdot 8 \cdot 7$, or notice that this is the same as the permutation ${}_{10}P_4 = 5040$.

The probability of no repeated digits is the number of 4 digit PIN numbers with no repeated digits divided by the total number of 4 digit PIN numbers. This probability is $\frac{{}_{10}P_4}{10^4} = \frac{5040}{10000} = 0.504$

References

1. Bennett et. al. Statistical Reasoning for Everyday Life. Chapter 6.
2. Lind et. al. Basic Statistics for Business and Economics. Chapter 5.
3. Newbold et al. Statistics for Business and Economics. Chapter 3.

Thank You!