

### TUTORIAL 2

#### Learning outcomes:

- 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.

**Probability** is the mathematical likelihood of an event occurring, a value expressed on a scale from 0 (impossible) to 1 (certainty), or as a percentage from 0% to 100%.

**The sample space**,  $S$ , is the set of all possible outcomes of an experiment.

An **event** is a collection of basic outcomes from the sample space  $S$  of an experiment, which is a subset of  $S$ .

While determining the probabilities, there are three approaches:

- 1) *Classical probability* – used to calculate the probabilities of equally likely outcomes (e.g., getting an even number from rolling a die).
- 2) *Empirical probability* (relative frequency) – based on dividing the frequency of successes to total number of trials (e.g., computing the probability of an accident using the accident data for the last 5 years).
- 3) *Subjective probability* – based on own judgment of the individual.

Some notations in probability:

Symbol	Short description	Probabilistic meaning	Example
$\cup$	or	union	$P(A \cup B)$
$\cap$	and	intersection	$P(A \cap B)$
$^c$	not	complement of	$P(A^c)$
$ $	given	conditional on	$P(A   B)$

### 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

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

*Example.*

In a class of 25 students, 20 have taken Finance, 16 have taken Economics, 14 have taken both.

What is the probability that a student chosen at random has taken at least one of these subjects?

*Solution.* Let E – Economics and F – Finance, then

$$P(E \cup F) = \frac{16}{25} + \frac{20}{25} - \frac{14}{25} = \frac{22}{25}$$

### Multiplicative Rule of Probability

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

*Example.*

You bought a M&M's chocolate bag. Suppose there are 10 red, 5 blue, 5 brown and 15 green chocolate chips in the bag. You randomly took out 2 chips out of the bag without a replacement. What is the probability that you pick red first, and blue?

*Solution.*

$$P(\text{red} \cap \text{blue}) = P(\text{red}) * P(\text{blue} | \text{red}) = \frac{10}{35} * \frac{5}{34} = 0.042$$

**Mutually exclusive events** (disjoint events) those types that cannot happen or exist at the same time.

**Collectively exhaustive events** are a set of events, or categories, when taken together, covers all possible outcomes.

**Complementary events** are two outcomes of an experiment that are mutually exclusive and collectively exhaustive.

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  $P(B|A) = P(B)$ .

## Conditional probability

Let A, B be events. The conditional probability of event A, given that event B has occurred, is denoted by  $P(A | B)$  and is found by:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad \text{provided that } P(B) \neq 0$$

Similarly,

$$P(B | A) = \frac{P(A \cap B)}{P(A)} \quad \text{provided that } P(A) \neq 0$$

**The Law of Total Probability** calculates the overall probability of an event by summing the probabilities of that event occurring under each condition of a set of mutually exclusive and exhaustive events, which together form a complete partition of the sample space.

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

here, B and  $B^C$  are mutually exclusive, and collectively exhaustive.

*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 and 0.01 for low-risk clients calculate the overall probability that any randomly selected borrower will default.

*Solution.* Let's denote:

$$H - \text{high-risk} \Rightarrow P(H) = 0.20$$

$$H^C - \text{low-risk} \Rightarrow P(H^C) = 1 - 0.20 = 0.80$$

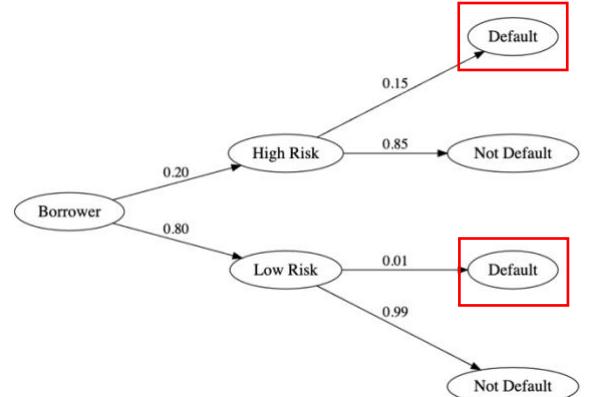
$$D - \text{Default}, \quad P(D | H) = 0.15 \text{ and } P(D | H^C) = 0.01$$

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

## Bayes' theorem (or Bayes' rule)

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:

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

*Example.*

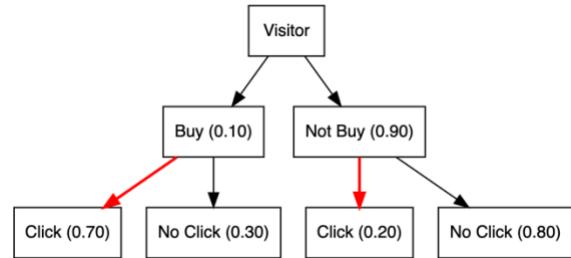
A company wants to know how likely a customer is to buy a product after clicking on an ad. According to historical data, 10% of all website visitors buy the product. But 70% of the buyers have clicked the ad, and this percentage is only 20% for non-buyers.

Find the probability of buying the product after clicking on an ad.

*Solution.*  $P(\text{Click}) = 0.70 * 0.10 + 0.20 * 0.90 = 0.25$

$$P(\text{Buy} | \text{Click}) = \frac{P(\text{Buy}) * P(\text{Click} | \text{Buy})}{P(\text{Click})} = \frac{0.10 * 0.70}{0.25} = 0.28$$

If a customer clicks an ad, the chance they will buy increases from 10% (prior) to 28% (posterior).



## Permutations and combinations

Permutations involve arrangements where the order of items matters, while combinations are selections where the order is irrelevant.

	With replacement	Without replacement
<b>Combinations (unordered)</b>	$n+r-1 C_r = \frac{(n+r-1)!}{r!(n-1)!}$	$n C_r = \frac{n!}{r!(n-r)!}$
<b>Permutations (ordered)</b>	$n^r$	$n P_r = \frac{n!}{(n-r)!}$

Here,  $n$  represents the total number of distinct objects available in a set, and  $r$  represents the number of these objects that are selected.

For example, if I choose two letters out of three (A, B, C), and do not care about the order, then the following *combinations* can be chosen:

- without replacement: AB, AC, BC.  $Formula: {}_3C_2 = \frac{3!}{2!(3-2)!} = 3$
- with replacement: AB, AC, BC, AA, BB, CC.  $Formula: {}_{3+2-1}C_2 = {}_4C_2 = \frac{4!}{2!(4-2)!} = 6$

If the order is important, then the following **permutations** can be chosen:

- without replacement: AB, AC, BC, BA, CA, CB.  $Formula: {}_3P_2 = \frac{3!}{(3-2)!} = 6$
- with replacement: AB, AC, BC, AA, BB, CC, BA, CA, CB.  $Formula: n^r = 3^2 = 9$

*Example.*

A company has 10 employees: 3 managers (M1, M2, M3) and 7 staff (S1–S7). A 4-person task force is formed at random. What is the probability the task force contains at least 2 managers?

*Solution.*

We count teams with exactly 2 managers and exactly 3 managers (4 managers impossible).

- Exactly 2 managers (2 from managers, 2 from staff):  ${}_3C_2 * {}_7C_2 = 3 * \frac{6*7}{2} = 63$  different ways.
- Exactly 3 managers (3 from managers, 1 from staff):  ${}_3C_3 * {}_7C_1 = 7$  different ways.

There are total of  ${}_{10}C_4 = \frac{10!}{4!*6!} = \frac{7*8*9*10}{1*2*3*4} = 210$  ways to select 4 people out of 10.

Then, total favorable outcomes =  $63+7 = 70$  and total outcomes = 210.

$$P(\text{at least 2 managers}) = \frac{70}{210} \approx 0.33$$

## TASKS

1. For the following exercises, decide whether the statement makes sense (or is clearly true) or does not make sense (or is clearly false).
  - a. When randomly selecting a day of the week, it is certain that you will select a day containing the letter y, so  $P(y) = 1$ .
  - b. If there is a 0.9 probability that it will rain sometime today, then there is a probability of 0.1 that it will not rain sometime today.
  - c. An insurance company states that the probability that a particular car will be involved in a car crash this year is 0.6 and the probability that the car will not be involved in a car crash this year is 0.3.
  - d. Jack estimates that the subjective probability of his being struck by lightning sometime next year is  $1/2$ .
  - e. The probability of rolling 7 on a die is  $\frac{1}{6}$ .
  - f. Let A be the event that you throw a six, and B be the event that you throw a three from rolling a die. Then the events A and B are independent.
  - g. If events A and B are independent, then  $P(A \cap B) = P(A) + P(B)$
  - h. When rolling a die, the probability of getting a number no higher than 6 is 1.
  - i. If two events, A and B, are independent, then the conditional probability of A given B, denoted  $P(A|B)$ , is always equal to 1.
  - j. The number of ways to select 3 objects from a set of 5 objects without a replacement is smaller when using combinations than when using permutations.
2. Consider the following Venn diagram, where

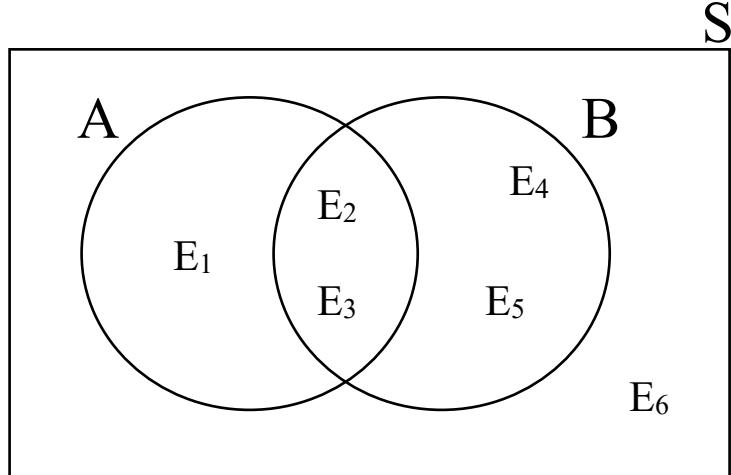
$$P(E_2) = P(E_3) = 1/5, \quad P(E_4) = P(E_5) = 1/10, \quad P(E_6) = 1/20$$

here,  $E_1, E_2, \dots, E_6$  are mutually exclusive

and collectively exhaustive events.

Find each of the following probabilities:

- |                  |                    |
|------------------|--------------------|
| a. $P(A)$        | e. $P(A^c)$        |
| b. $P(B)$        | f. $P(B^c)$        |
| c. $P(A \cup B)$ | g. $P(A \cup A^c)$ |
| d. $P(A \cap B)$ | h. $P(A^c \cap B)$ |



**3.** Suppose that 80% of the employees of a company received cash or company stock as a bonus at the end of the year. If 120 employees (60% of the total employees) received a cash bonus and 30% received stock, then how many employees received neither cash nor stock as a bonus?

**4.** You work in a small company with 10 employees, 3 of whom are close friends of yours. If the directors of this company are chosen at random, what is the probability that you are named director and one of your friends is named deputy director?

**5.** A personnel officer has 8 candidates to fill 4 similar positions. 5 candidates are men, and 3 are women. If, in fact, every combination of candidates is equally likely to be chosen, what is the probability that no women will be hired?

**6.** Suppose that there are 7 students in your classroom who were born in the month of October. Compute the probability that at least two of them share the same birthday.

**7.** In an audit Ali analyses 70% of the audit items and Ahmed analyses the remaining items. Ali's error rate is 5.5% and Ahmed's error rate is 3.4%. Suppose an item is sampled at random.

**a.** What is the probability that it is in error (i.e. audited incorrectly)?

**b.** If the chosen item is incorrect what is the probability that Ali is to blame?

**c.** If the chosen item is incorrect what is the probability that Ahmed is to blame?

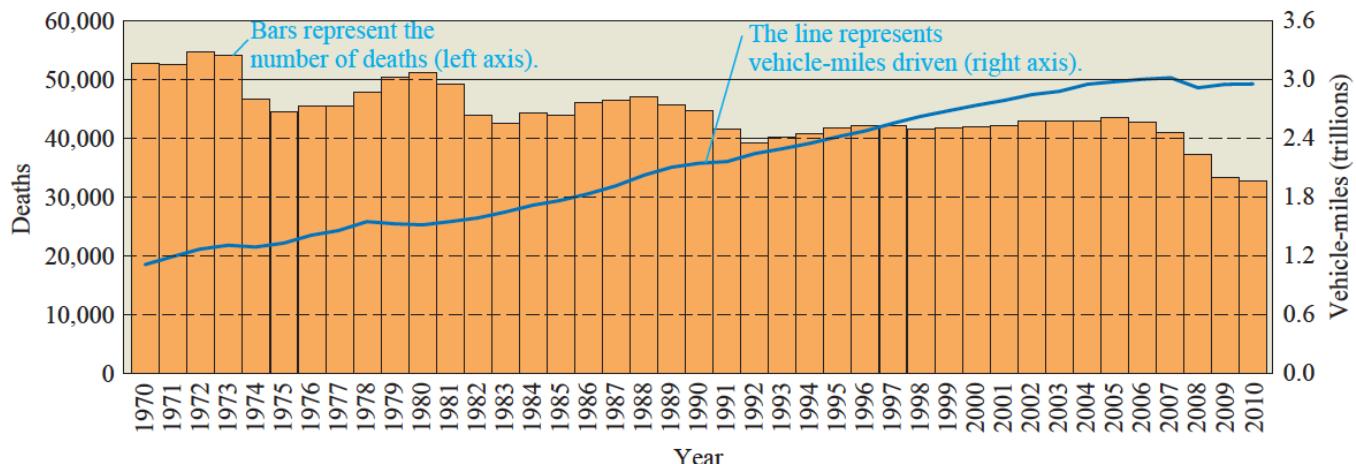
**8.** The successful operation of three separate switches is needed to control a machine. If the probability of failure of each switch is 0.1 and the failure of any switch is independent of any other switch, what is the probability that the machine will break down?



**9. Which Is Safer: Flying or Driving? (work in groups of 3-4 students to discuss this task)**

For the period 1990 through 2010, the average (mean) number of deaths in commercial airplane accidents in the United States was roughly 60 per year. (The actual number varies significantly from year to year.) As of 2010, airplane passengers in the United States travel a total of about 8 billion miles per year. Use these numbers to calculate the death rate per mile of air travel.

The figure below shows the number of automobile fatalities and the total number of miles driven (among all Americans) for each year over a period of more than four decades.



**Figure 9** This graph shows the annual number of deaths (bars) and the annual number of vehicle-miles driven (line) in the United States from 1970 to 2010. Source: National Transportation Safety Board.

Compare the risk of flying to the risk of driving (you can compare year 2010 for driving and flying since this was the safest driving year according to the figure).

### HOMEWORK

10. In a study of checkout scanning systems, samples of purchases were used to compare the scanned prices to the posted prices. Table 1 summarizes results for a sample of 819 items.

**TABLE 1** Scanner Accuracy

	Regular-priced items	Advertised-special items
Undercharge	20	7
Overcharge	15	29
Correct price	384	364

Based on these data,

- a. what is the probability that a regular-priced item has a scanning error?
  - b. what is the probability that a randomly selected item is either advertised-special or overcharged item?
11. Let event A = learning Spanish. Let event B = learning German. Then  $A \cap B$  = learning Spanish and German. Suppose  $P(A) = 0.4$  and  $P(B) = 0.2$ .  $P(A \cap B) = 0.08$ . Are events A and B independent?

**12.** Suppose you randomly select a family with three children. Assume that births of boys and girls are equally likely. What is the probability that the family has each of the following?

- a. Three girls
- c. A girl, a boy, and a boy, in that order
- b. Two boys and a girl
- d. At least one girl

**13.** In the past, Energy construction's main competitor, Skyline Company, has submitted bids 75% of the time. If Skyline does not bid on a job, then there is 0.70 probability that Energy will get the job. If Skyline submits a bid, then the probability that Energy gets the job drops to 0.40.

- a. What is the probability that Energy gets the job for the next bid?
- b. If Energy gets the job, what is the probability that Skyline made a bid?
- c. If Energy did not get the job, what is the probability that Skyline did not make a bid?

**14.** At a local high school, 80% of the students took IELTS test, and 15% of the students took both IELTS and SAT. Based on the information provided, which of the following calculations are not possible, and why? Here, IELTS = A, SAT = B.

- a.  $P(B | A)$
- b.  $P(A | B)$
- c.  $P(A \cup B)$

**15.** Ahmed is in a dark room selecting socks from his drawer. He has only six socks in his drawer, a mixture of black and white. If he chooses two socks, the chances that he picks a white pair is  $\frac{2}{3}$ . Then what are the chances that he selects a black pair?