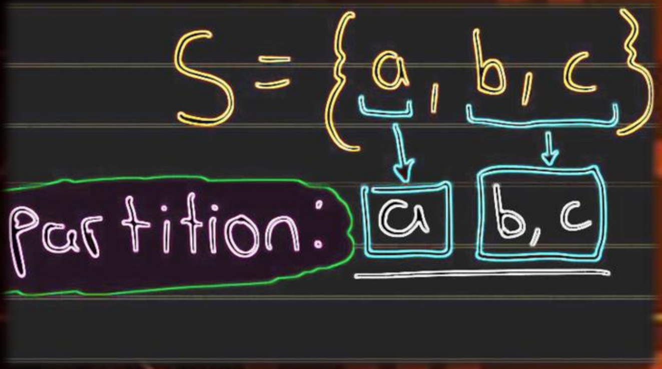
Probability and Statistics

ENG-PST



KARABÜK ÜNİVERSİTESİ 2007

Lecture 6

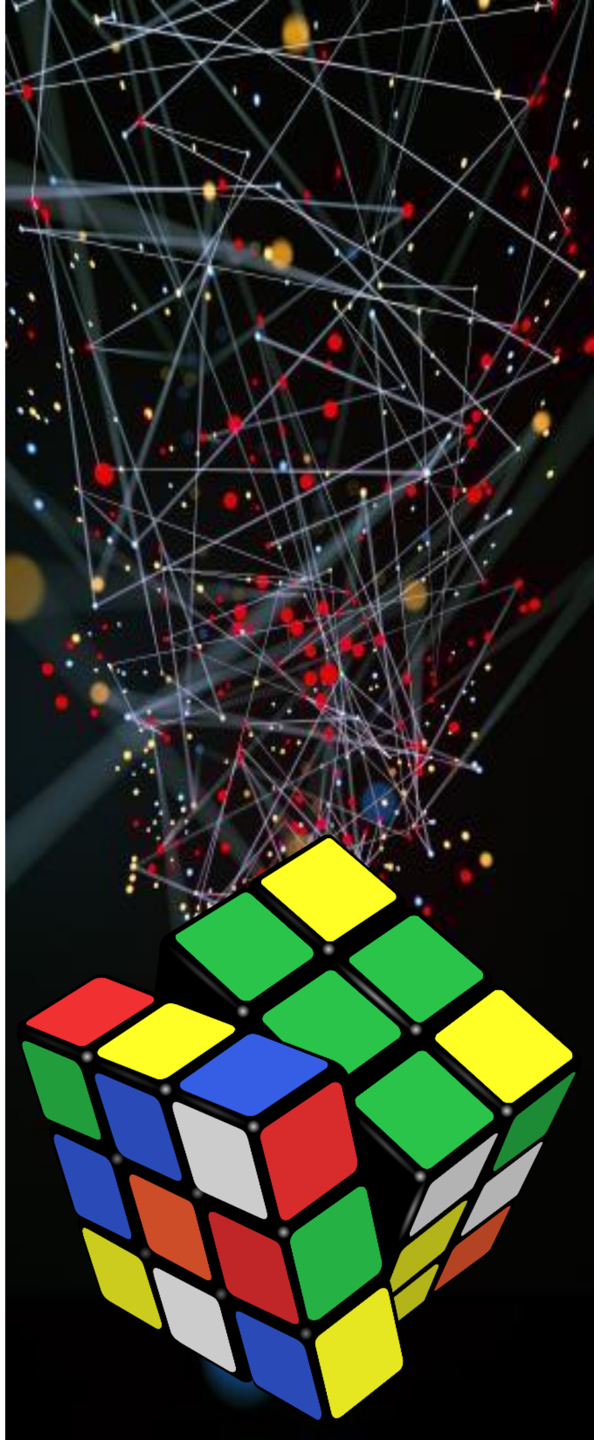
Dr. Tarik Adnan Almohamad
Email: tarikalmohamad@karabuk.edu.tr
Office: Kat.1, # 164

Probability of an Event

- The likelihood of the occurrence of an event resulting from such a statistical experiment is evaluated by means of a set of real numbers, called weights or probabilities, ranging from **0** to **1**.

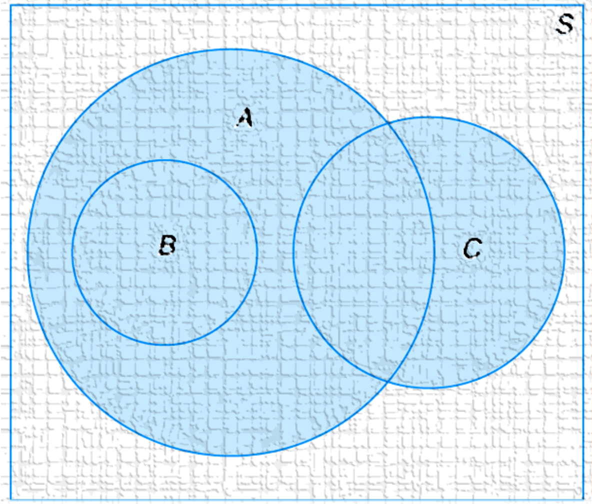
- To every point (or element) in the sample space, we assign a probability such that the sum of all probabilities is **1**.

- If we have reason to believe that a certain sample point is quite likely to occur when the experiment is conducted, the probability assigned should be close to 1. On the other hand, a probability closer to 0 is assigned to a sample point that is not likely to occur.



Cont.

- To find the probability of an event A, we sum all the probabilities assigned to the sample points in A. This sum is called the probability of A and is denoted by $P(A)$.



Note that **B** & **C** events are not related to each other. Hence, they are mutually exclusive

The **probability** of an event A is the sum of the weights of all sample points in A. Therefore,

$$0 \leq P(A) \leq 1, \quad P(\phi) = 0, \quad \text{and} \quad P(S) = 1.$$

Furthermore, if $A_1, A_2, A_3, \ldots$ is a sequence of mutually exclusive events, then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots.$$

Example

- A coin is tossed twice. What is the probability that at least one head occurs?

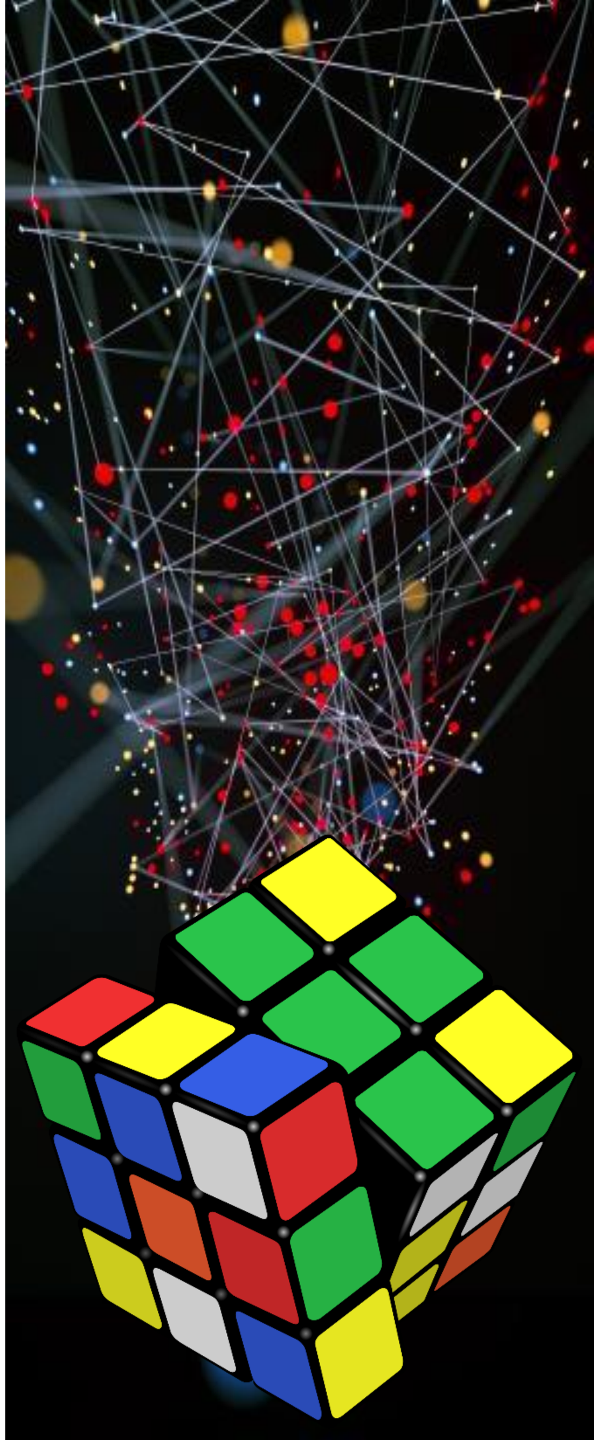
Solution:

- The sample space for this experiment is $S = \{HH, HT, TH, TT\}$

- If the coin is balanced, each of these outcomes is equally likely to occur. Therefore, we assign a probability of w to each sample point. Then $4w = 1$, or $w = 1/4$.

- If A represents the event of at least 1 head occurring, then

$$A = \{HH, HT, TH\} \text{ and } P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$$



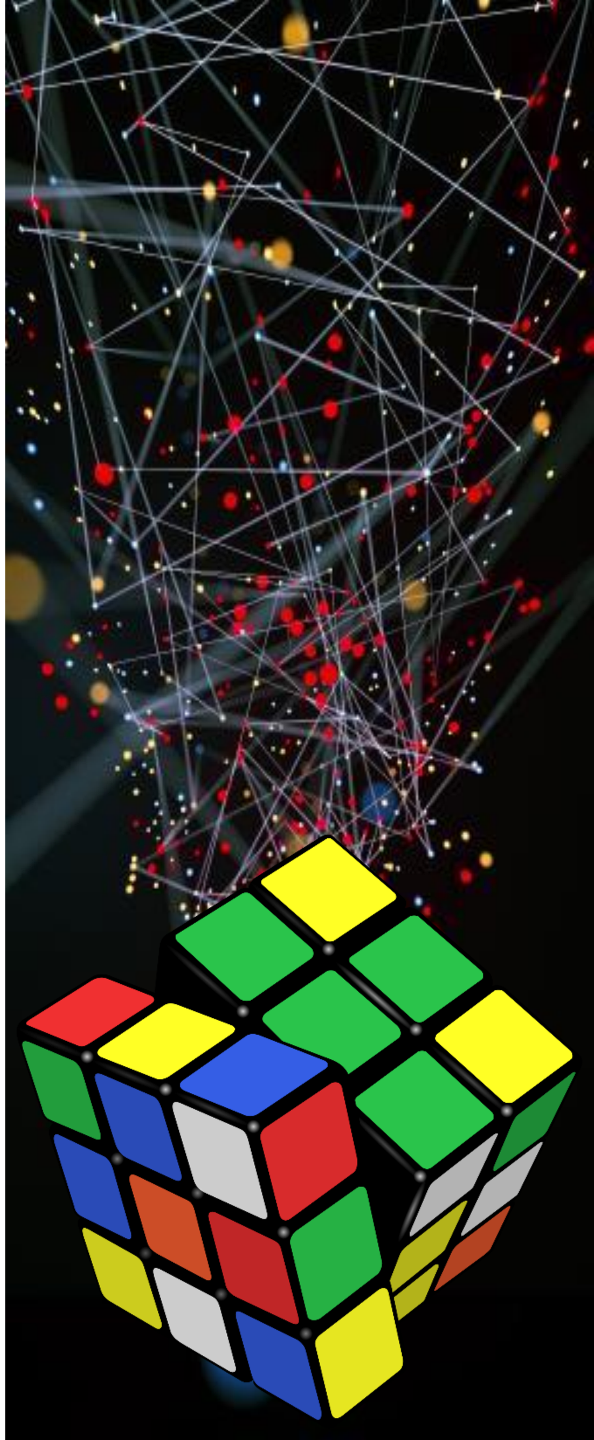
Example

- A die (*zar*) is loaded in such a way that an even number is twice as likely to occur as an odd number. If E is the event that a number <u>less</u> than 4 occurs on a single toss of the die, find $P(E)$.

Solution:

- The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. We assign a probability of w to each odd number and a probability of $2w$ to each even number.
- Since the sum of the probabilities must be 1, we have $9w = 1$ or $w = 1/9$. Hence, probabilities of 1/9 and 2/9 are assigned to each odd and even number, respectively. Therefore,

$$E = \{1, 2, 3\} \text{ and } P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}.$$



Example

- In a dice game, let A be the event that an even number turns up and let B be the event that a number divisible by 3 occurs. Assume that the even number is twice as likely to occur as an odd number. Find $P(A \cup B)$ and $P(A \cap B)$.

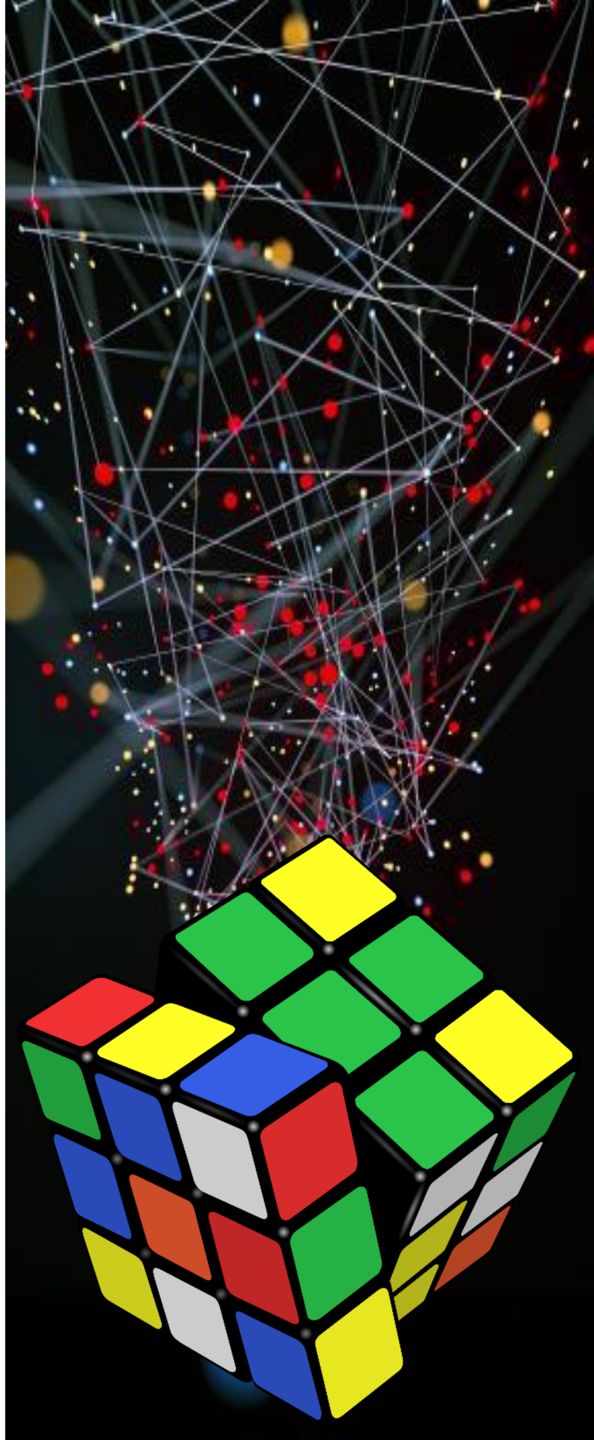
Solution:

- For the events $A = \{2, 4, 6\}$ and $B = \{3, 6\}$, we have

$$A \cup B = \{2, 3, 4, 6\} \; and \; A \cap B = \{6\}.$$

- By assigning a probability of 1/9 to each odd number and 2/9 to each even number, we have

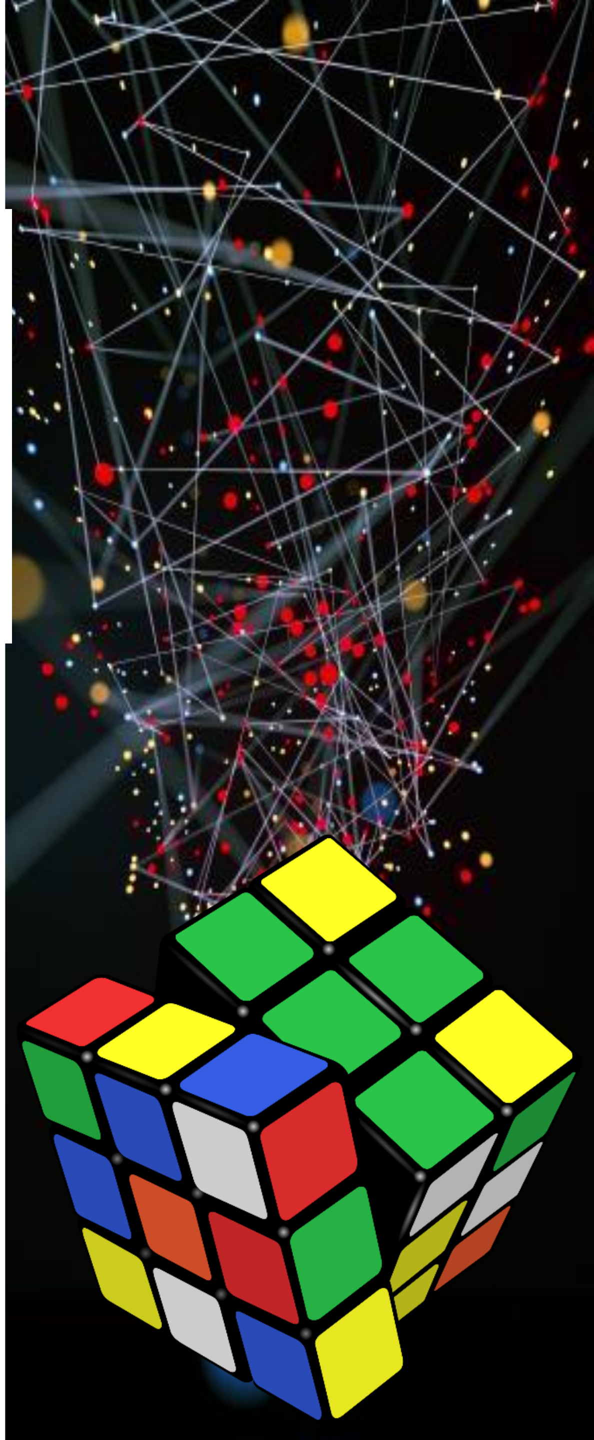
$$P(A \cup B) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9} \quad \text{and} \quad P(A \cap B) = \frac{2}{9}.$$



Cont.

> If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then the probability of event A is
>
> $$P(A) = \frac{n}{N}.$$

- If the sample space for an experiment contains N elements, all of which are equally likely to occur, we assign a probability equal to $1/N$ to **each** of the N points.
- The probability of any event A containing n of these N sample points is then the ratio of the number of elements in A to the number of elements in S.

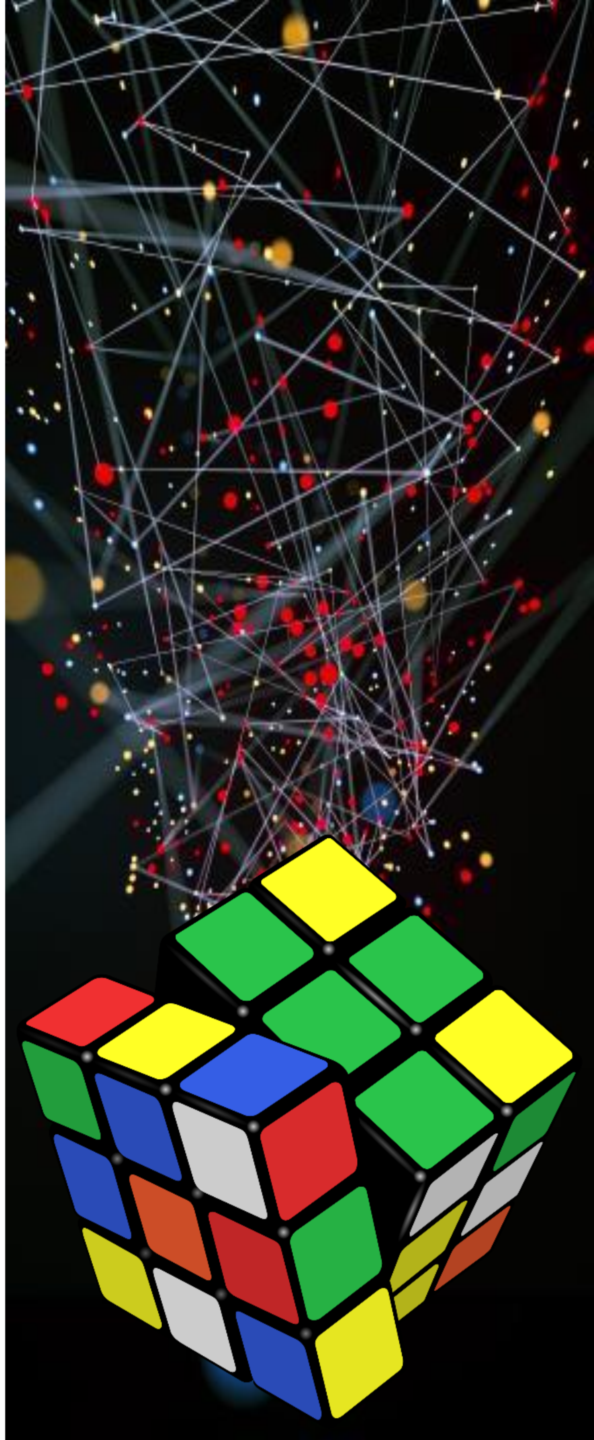


Example

- A statistics class for engineers consists of 25 industrial, 10 mechanical, 10 electrical, and 8 civil engineering students. If a person is randomly selected by the lecturer to answer a question, find the probability that the student chosen is:

- (a) an <u>industrial</u> engineering major,
- (b) a <u>civil</u> engineering or an <u>electrical</u> engineering major.

Solution?



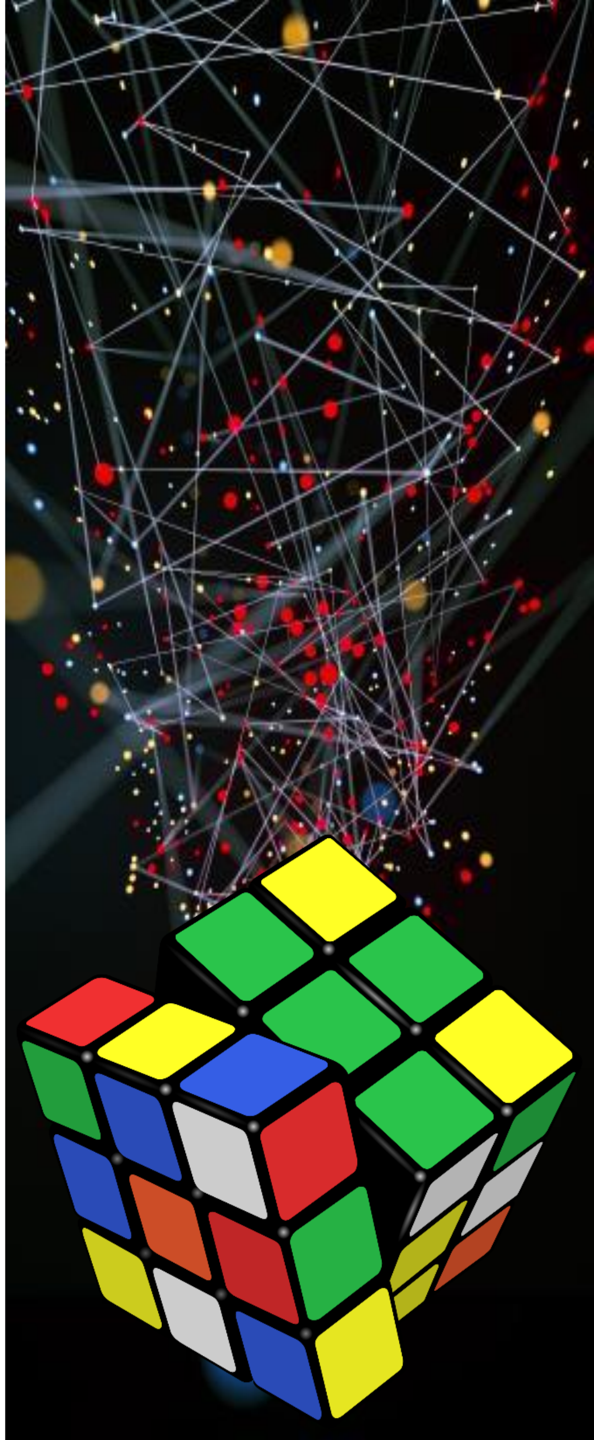
Solution:

- Let's denote by I: industrial, M, E, and C the students majoring, mechanical, electrical, and civil engineering, respectively. The total number of students in the class is 53, all of whom are equally likely to be selected.

(a) Since 25 of the 53 students are majoring in industrial engineering, the probability of event I, selecting an industrial engineering major at random, is $P(I) = \frac{25}{53}$

(b) Since 18 of the 53 students are civil or electrical engineering majors, it follows that

$$P(C \cup E) = \frac{18}{53}$$



Additive Rules

- Several important laws that frequently simplify the computation of probabilities follow. The first, called the additive rule, applies to unions of events.

If A and B are two events, then

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

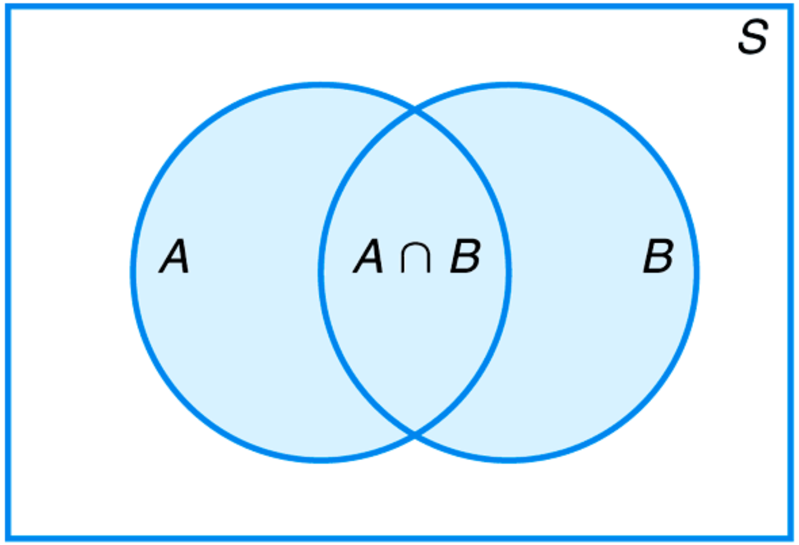


Figure 2.7: Additive rule of probability

Cont.

If A and B are mutually exclusive, then

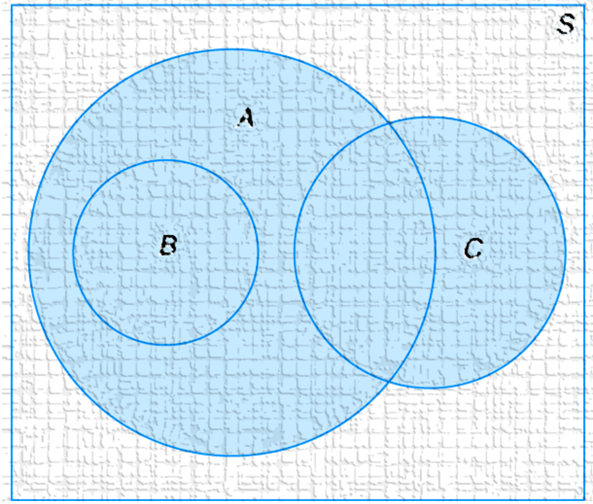
$$P(A \cup B) = P(A) + P(B).$$

If $A_1,\ A_2, \ldots, A_n$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n).$$

If $A_1,\ A_2, \ldots, A_n$ is a partition of sample space S, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n) = P(S) = 1.$$



Note that **B** & **C** events are not related to each other. Hence, they are mutually exclusive

Cont.

For three events A, B, and C,

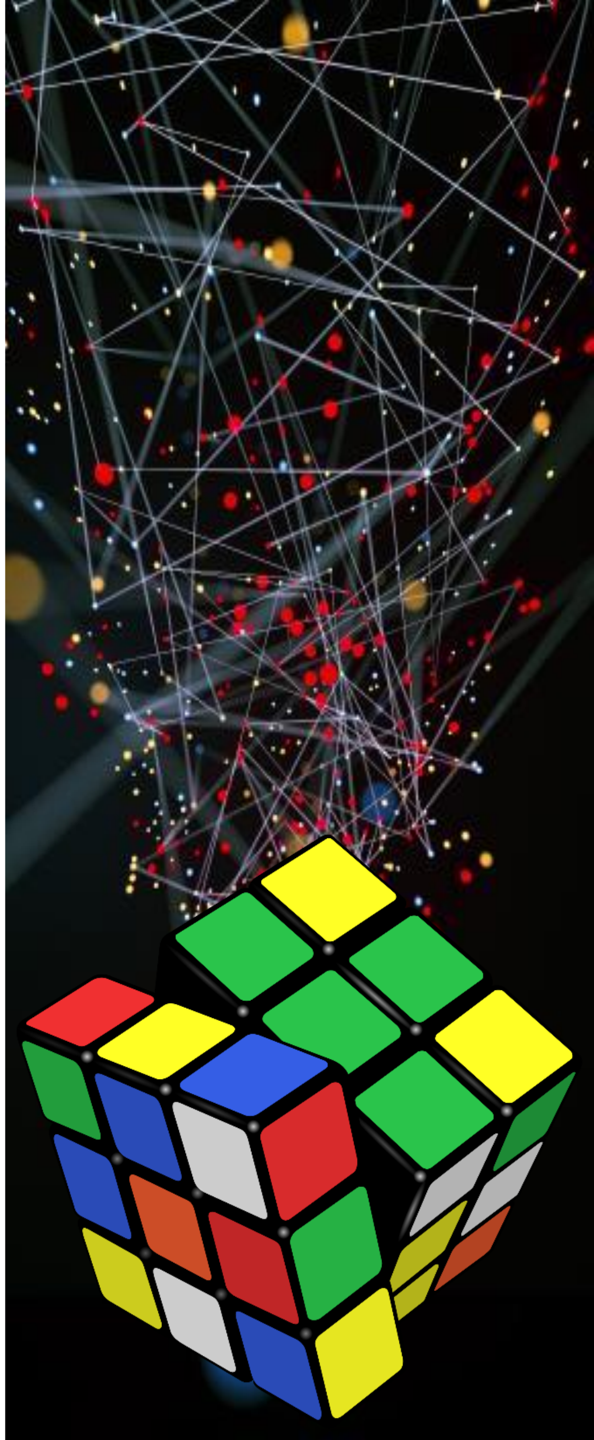
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) \\ - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

If A and A' are complementary events, then

$$P(A) + P(A') = 1.$$

The conditional probability of B, given A, denoted by $P(B|A)$, is defined by

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



Cont.

Two events A and B are **independent** if and only if

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

assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**.

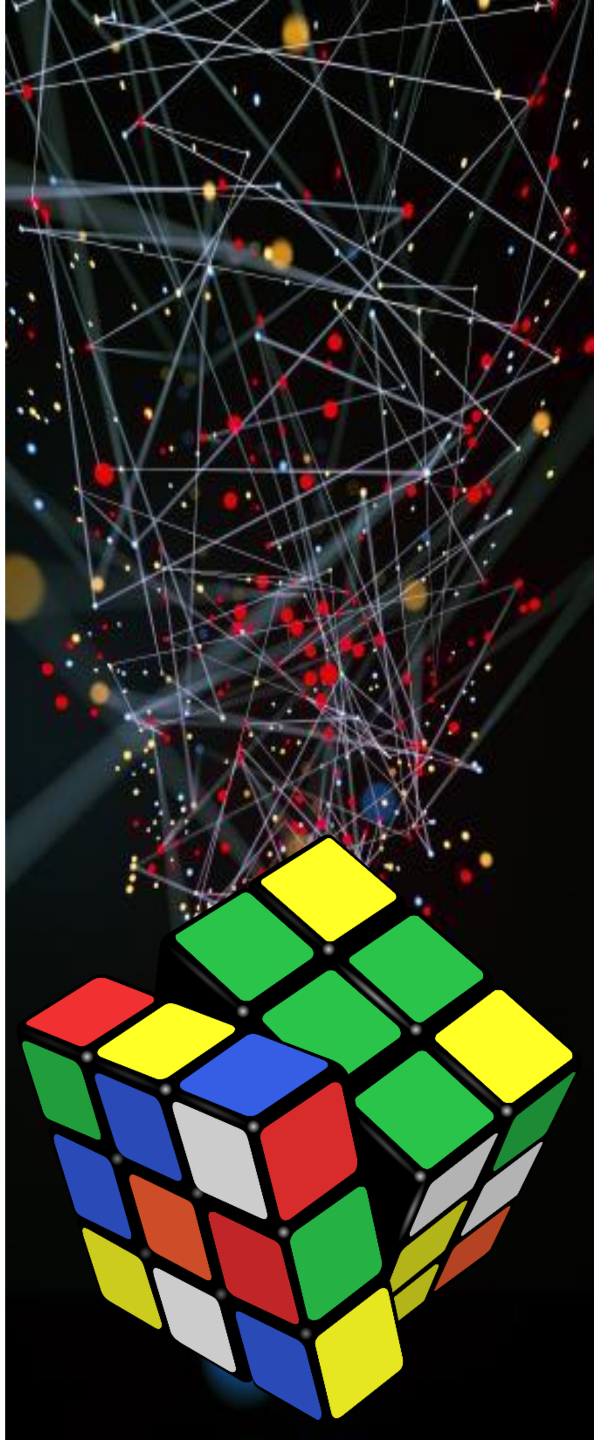
If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0.$$

Two events A and B are independent if and only if

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

Therefore, to obtain the probability that two independent events will both occur, we simply find the product of their individual probabilities.



Cont.

If, in an experiment, the events $A_1, A_2, \ldots, A_k$ can occur, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k)$$
$$= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}).$$

If the events $A_1, A_2, \ldots, A_k$ are independent, then

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \cdots P(A_k).$$

The probability of the event A

$$P(A) = P[(E \cap A) \cup (E' \cap A)] = P(E \cap A) + P(E' \cap A)$$
$$= P(E)P(A|E) + P(E')P(A|E').$$

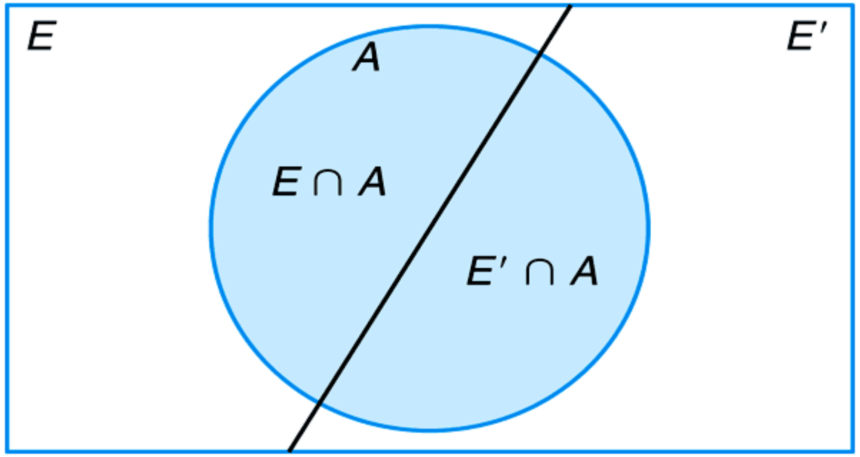
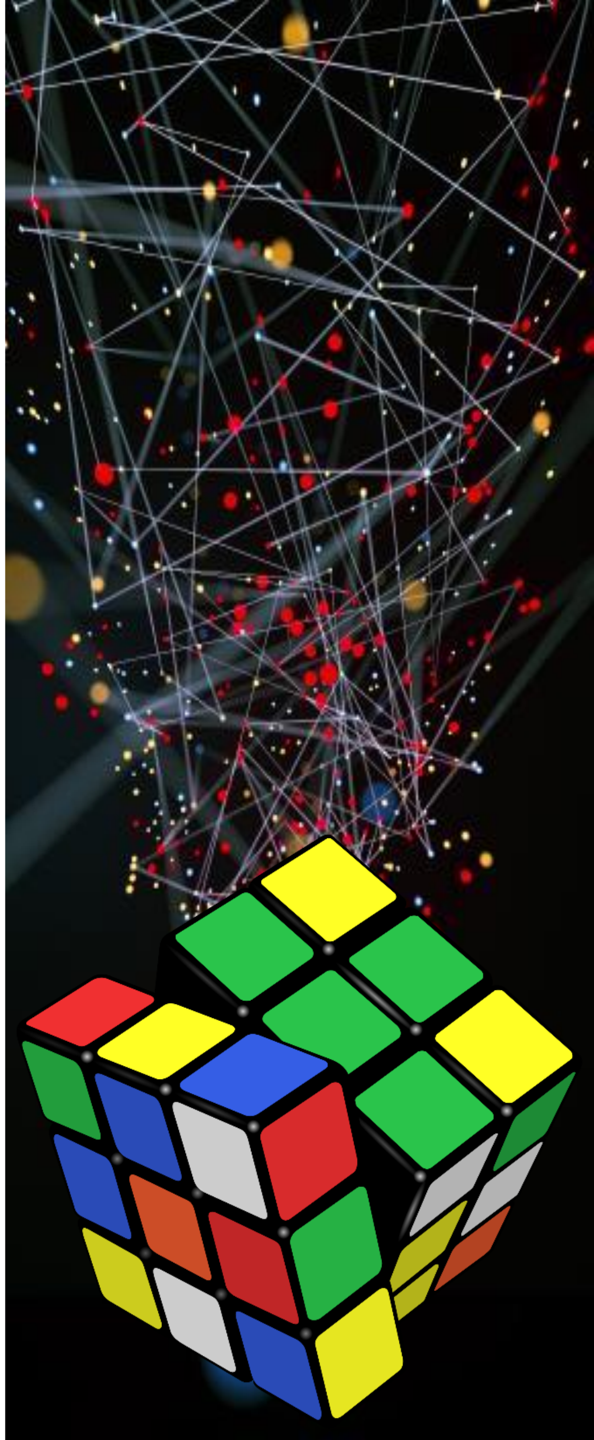


Figure 2.12 Venn diagram for the events *A*, *E* and *E′*



Total probability

- If the sample space is partitioned into k subsets is covered by the following theorem, sometimes called the theorem of **total probability** or the rule of elimination

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A of S,

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

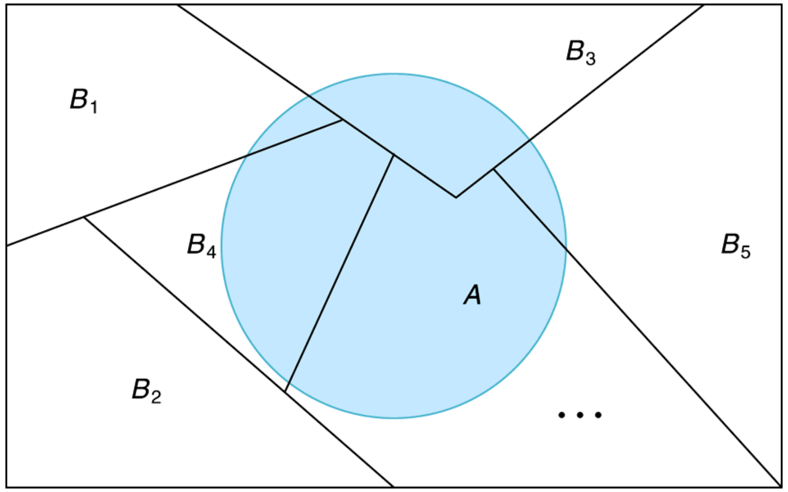
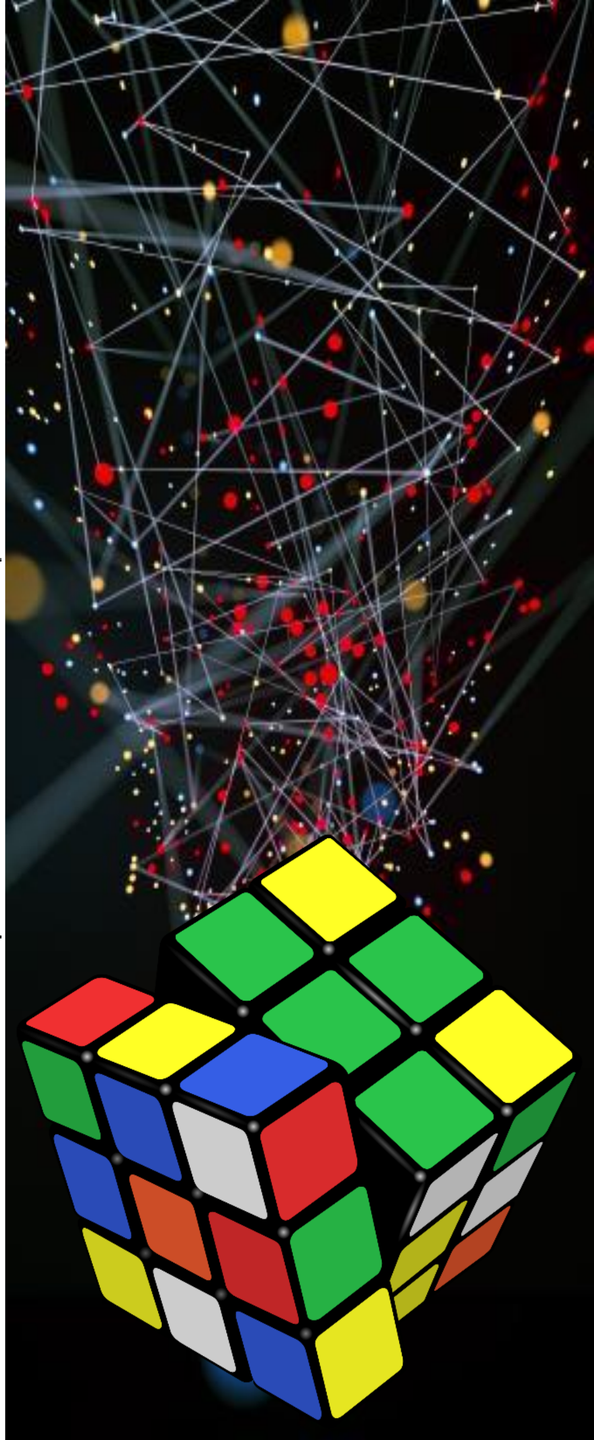


Figure 2.14: Partitioning the sample space S.

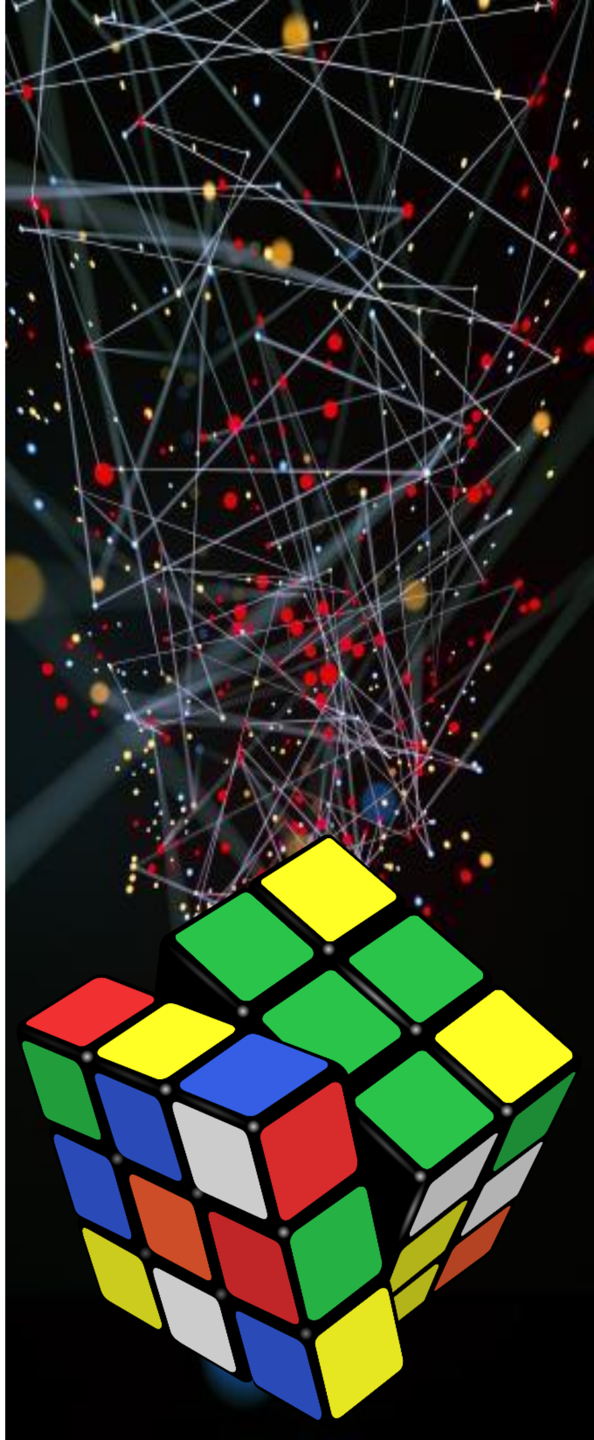


Bayes' Rule

- Bayesian statistics is a collection of tools that is used in a special form of statistical inference which applies in the analysis of experimental data in many practical situations in science and engineering. Bayes' rule is one of the most important rules in probability theory.

(Bayes' Rule) If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ for $i = 1, 2, \ldots, k$, then for any event A in S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)} \quad \text{for } r = 1, 2, \ldots, k.$$



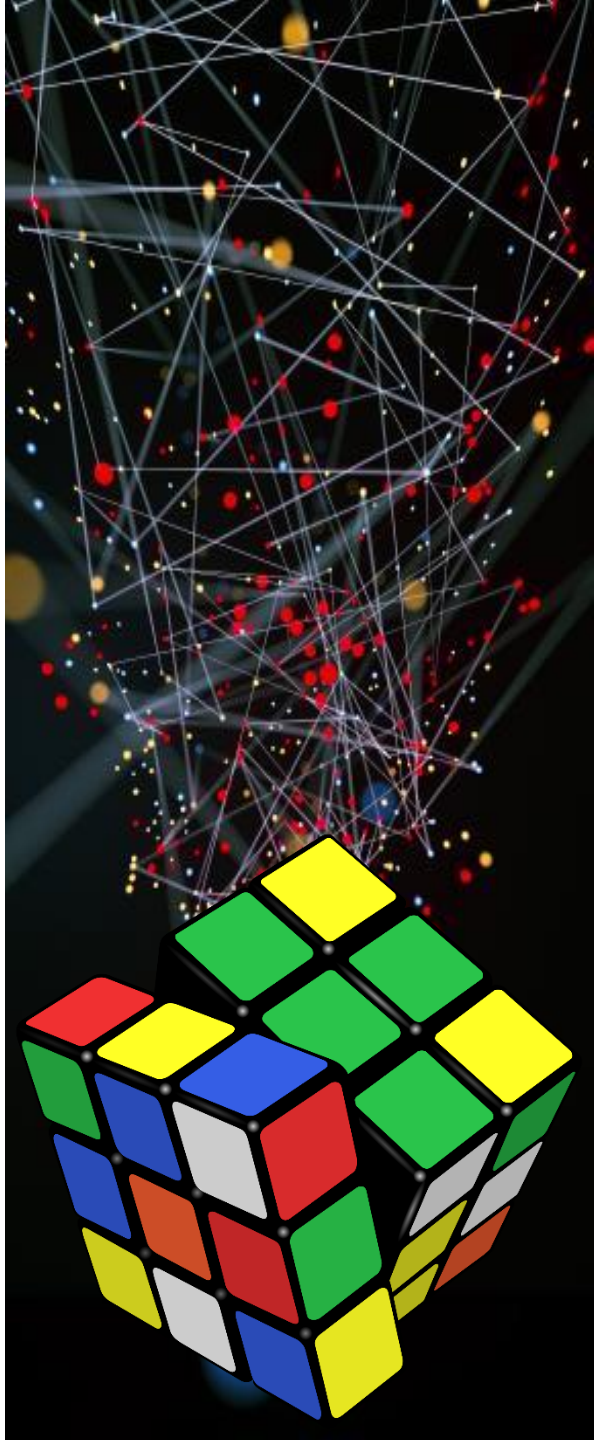
Example

- A manufacturing firm employs three analytical plans for the design and development of a particular product. For cost reasons, all three are used at varying times. In fact, plans 1, 2, and 3 are used for 30%, 20%, and 50% of the products, respectively. The defect rate is different for the three procedures as follows:

$$P(D|P_1) = 0.01, \qquad P(D|P_2) = 0.03, \qquad P(D|P_3) = 0.02$$

- where $P(D|P_j)$ is the probability of a defective product, given plan j. If a random product was observed and found to be defective, which plan was most likely used and thus responsible?

Solution?



Solution?

- From the statement of the problem

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad \text{and} \quad P(P_3) = 0.50,$$

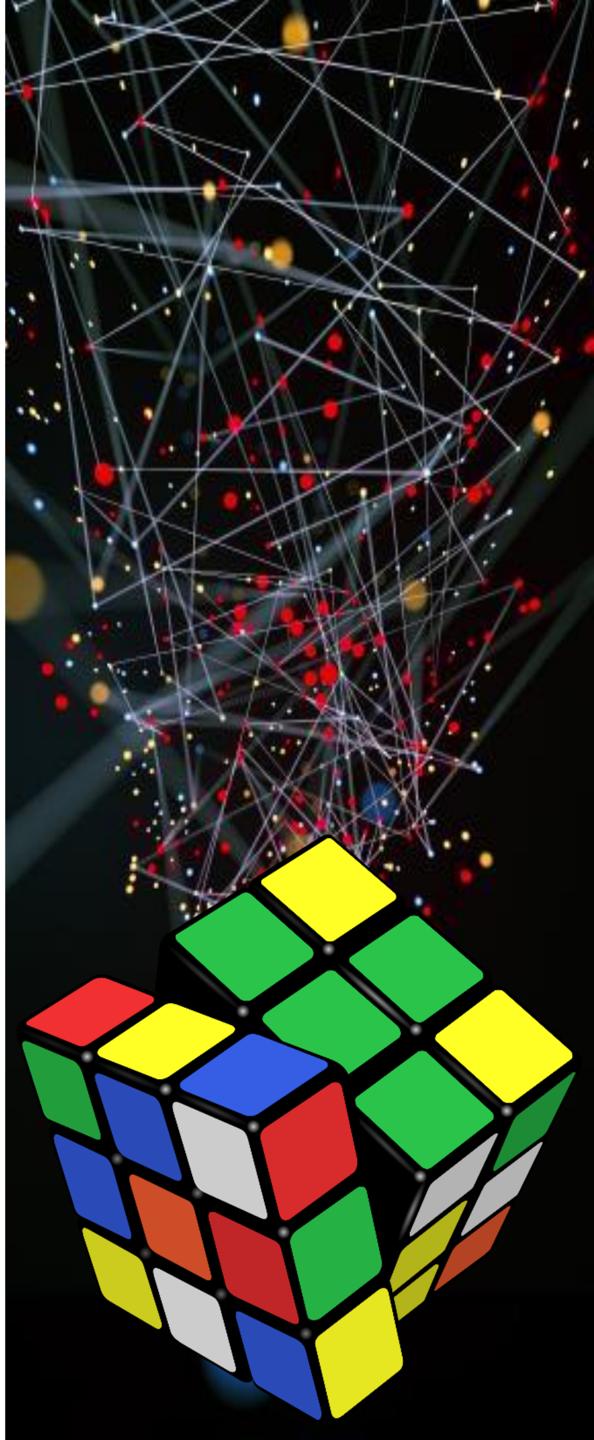
- we must find $P(P_j\,|D)$ for $j = 1, 2, 3$. Bayes' rule shows

$$P(P_1|D) = \frac{P(P_1)P(D|P_1)}{P(P_1)P(D|P_1) + P(P_2)P(D|P_2) + P(P_3)P(D|P_3)}$$
$$= \frac{(0.30)(0.01)}{(0.3)(0.01) + (0.20)(0.03) + (0.50)(0.02)} = \frac{0.003}{0.019} = 0.158.$$

Similarly,

$$P(P_2|D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \text{ and } P(P_3|D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

The conditional probability of a defect given plan 3 is the largest of the three; thus a defective for a random product is most likely the result of the use of plan 3.



Thank you

Feel free to ask questions in the class or through Microsoft Team.