Probability and Statstics

Dr. Syed Faisal Bukhari

Associate Professor

Department of Data Science

Faculty of Computing and Information Technology

University of the Punjab

Textbooks

- **Probability & Statistics for Engineers & Scientists**, Ninth Edition, Ronald E. Walpole, Raymond H. Myer
- **Elementary Statistics: Picturing the World,** 6th Edition, Ron Larson and Betsy Farber
- **Elementary Statistics**, 13th Edition, Mario F. Triola

Reference books

- **Probability Demystified**, Allan G. Bluman
- **Schaum's Outline of Probability and Statistics**
- **MATLAB Primer**, Seventh Edition
- **MATLAB Demystified** by McMahon, David

References

Readings for these lecture notes:

- Probability & Statistics for Engineers & Scientists, Ninth edition, Ronald E. Walpole, Raymond H. Myer
- Probability Demystified, Allan G. Bluman
- Elementary Statistics, 10^{th} Edition, Mario F. Triola

These notes contain material from the above three books.

Target Journals

Some of the journals that are relevant to health care and the medical field, based on computer science.

1. **Medical Decision Making**, JCR Impact Factor (2017-18) = 2.793
2. **Health Informatics Journal**, JCR Impact Factor (2017-18) = 2.297
3. **Informatics for Health and Social Care**, JCR Impact Factor (2017-18) = 1.137
4. **Health Care Analysis**, , JCR Impact Factor (2017-18) = 1.043
5. **International Journal of Health Care Quality Assurance**, JCR Impact Factor (2017-18) = 1.218

Target Journals

Some of the journals that are relevant to education, based on computer science.

1. **Computers & Education**, JCR Impact Factor (2017-18) = 5.627
2. **Computer Applications in Engineering Education**, JCR Impact Factor (2017-18) = **1.435**
3. **Journal of Computing in Higher Education**, JCR Impact Factor (2017-18) = **1.870**
4. **Acm Transactions on Computing Education**, , JCR Impact Factor (2017-18) = **1.356**
5. **Assessment & Evaluation In Higher Education**, JCR Impact Factor (2017-18) = **2.473**
6. **Educational Assessment Evaluation and Accountability**, JCR Impact Factor (2017-18) = **1.772**
7. **Computer Applications in Engineering Education = Impact Factor: 1.435**

Subjective Probability

A third type of probability is called **subjective probability**. Subjective probability is based upon an **educated guess**, **estimate**, **opinion**, or **inexact information**.

Sample Spaces

There are two **specific devices** that will be used to find sample spaces for probability experiments. They are **tree diagrams** and **tables**.

A **tree diagram** consists of **branches** corresponding to the outcomes of two or more probability experiments that are done in sequence.

Sample Spaces

- In order to construct a tree diagram, use branches corresponding to the outcomes of the **first experiment**. These branches will emanate from a single point.
- Then from each branch of the first experiment draw branches that represent the outcomes of the **second experiment**.
- You can continue the process for further experiments of the sequence if necessary.

Tree Diagram [1]

Example: A coin is tossed and a die is rolled. Draw a tree diagram and find the sample space.

Solution:

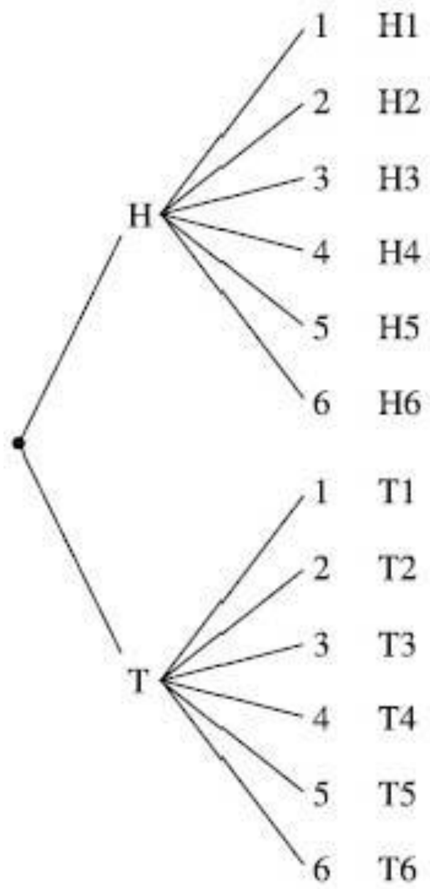
Since there are two outcomes (heads and tails for the coin), draw two branches from a single point and label one **H** for head and the other one **T** for tail.

From each one of these outcomes, draw and label six branches representing the outcomes **1, 2, 3, 4, 5,** and **6** for the die.

Trace through each branch to find the outcomes of the experiment.

Tree Diagram [2]

Example: A coin is tossed and a die is rolled. Draw a tree diagram and find the sample space.



Tree Diagram [3]

Example: A coin is tossed and a die is rolled. Find the probability of getting

a. A head on the coin and a 3 on the die.

b. A head on the coin.

c. A 4 on the die.

Solution:

a. $P(H3) = {}^{1}/_{12}$ = **0.0833 (or 8.33%)**

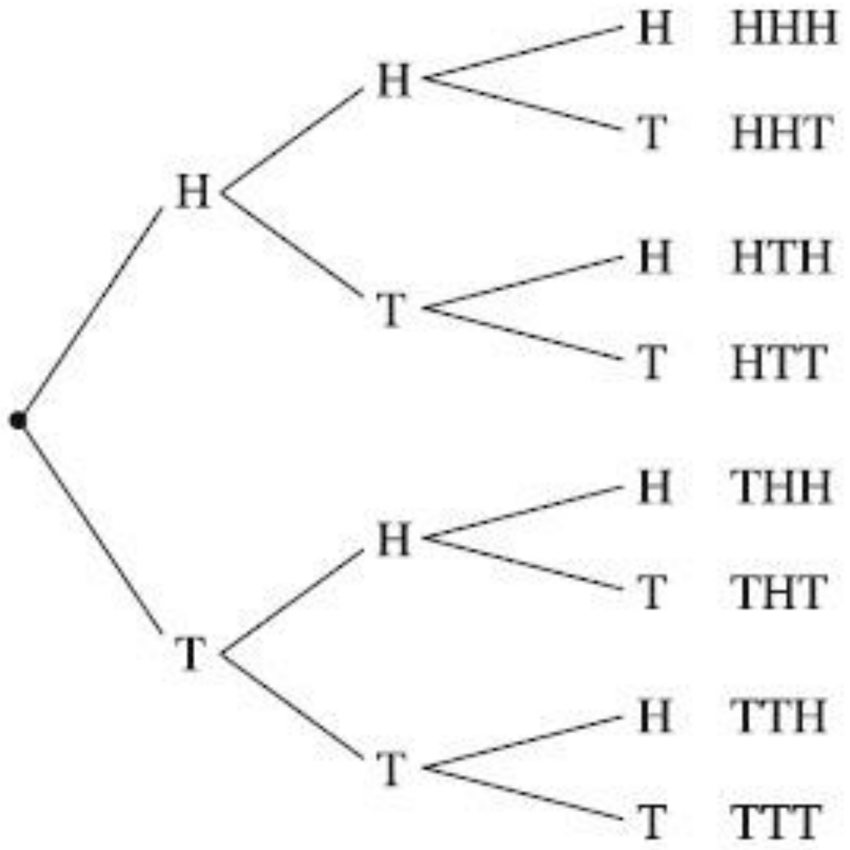
b. P(head on the coin) = ${}^{6}/_{12} = {}^{1}/_{2}$ = **0.5 (or 50%)**

c. P(4 on the die) = ${}^{2}/_{12} = {}^{1}/_{6}$ = **0.1667 (16.67%)**

Tree Diagram [4]

Example: Three coins are tossed. Draw a tree diagram and find the sample space.

Solution



Tree Diagram [5]

Example: Three coins are tossed. Find the probability of getting

a. Two heads and a tail in any order.

b. Three heads.

c. No heads.

d. **At least** two tails.

e. **At most** two tails.

Solution:

a. P(2 heads and a tail) = 3/8 = **0.375 (or 37.5 %)**

b. P(HHH) = $^{1}/_{8}$ = **0.125 (or 12.5 %)**

c. P(TTT) = $^{1}/_{8}$ = **0.125 (or 12.5 %)**

d. P(at least two tails) = $^{4}/_{8}$ = $^{1}/_{2}$ = **0.5 (or 50 %)**

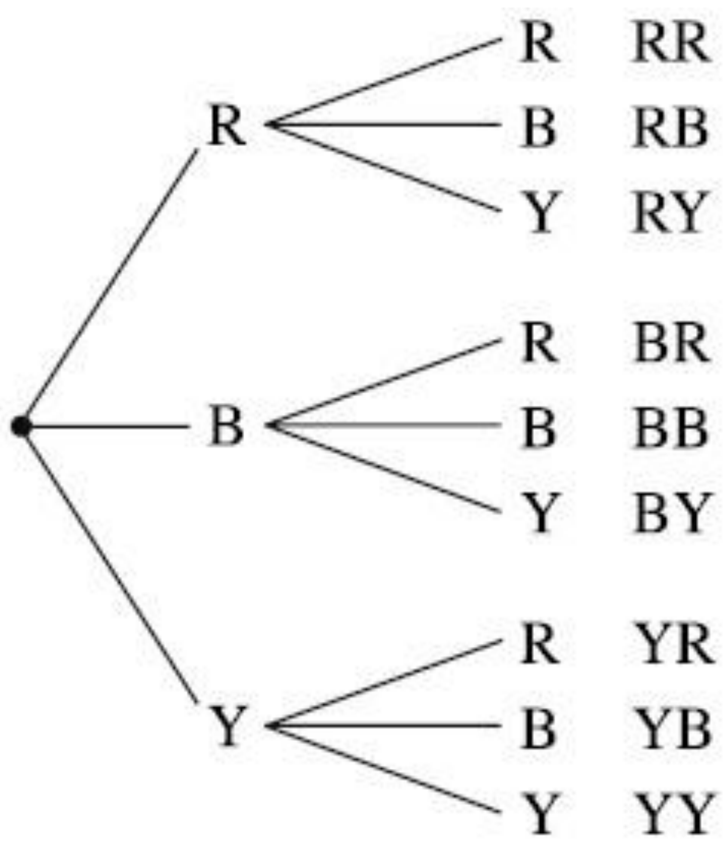
e. P(at most two tails) = $^{7}/_{8}$ = **0.875 (or 87.5 %)**

$\Rightarrow \boldsymbol{At\ most\ two\ tails\ mean\ no\ three\ tails}$

Tree Diagram [6]

Example: A box contains a **red ball (R),** a **blue ball (B),** and a **yellow ball (Y**). **Two balls** are selected at random in succession. Draw a **tree diagram** and find the sample space if the first ball is **replaced** before the second ball is selected.

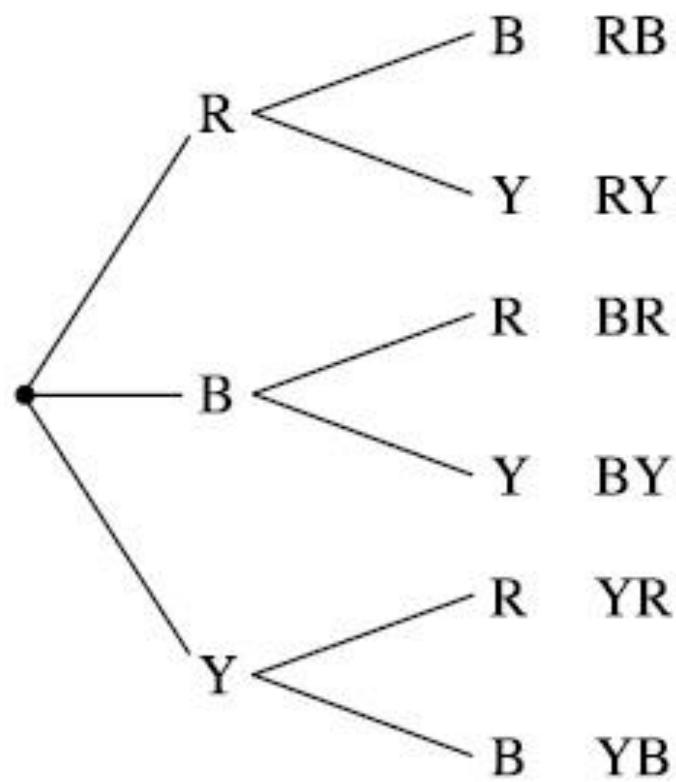
Solution



Tree Diagram [7]

Example: A box contains a **red ball (R),** a **blue ball (B)**, and a **yellow ball (Y)**. Two balls are selected at random in succession. Draw a **tree diagram** and find the sample space if the first ball is **not replaced** before the second ball is selected.

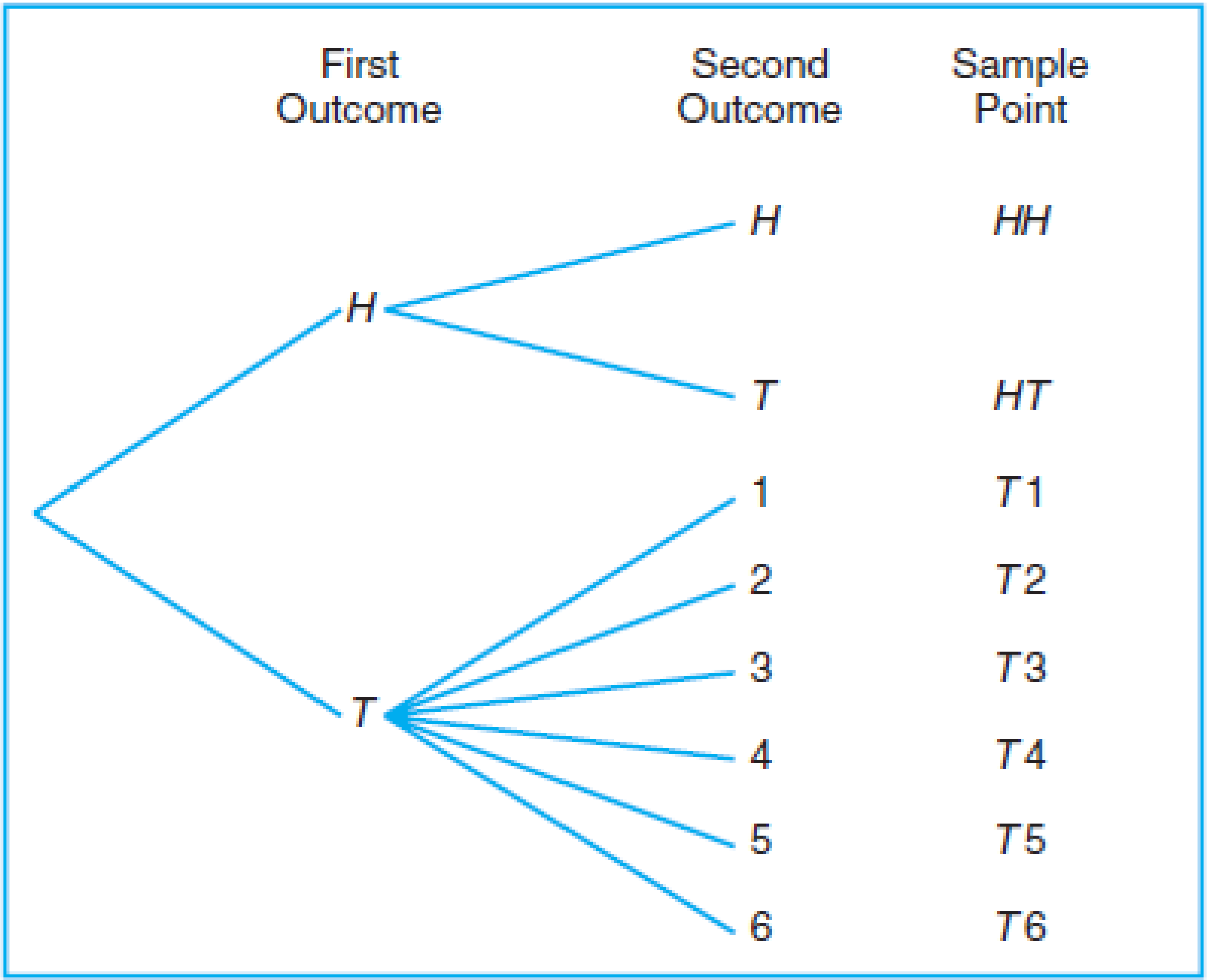
Solutions



Tree Diagram [9]

Example An experiment consists of **flipping a coin** and then flipping it a **second time** if a **head occurs**. If a **tail** occurs on the **first flip**, then a **die is tossed** once. To list the elements of the sample space providing the most information, construct the **tree diagram**

Solution

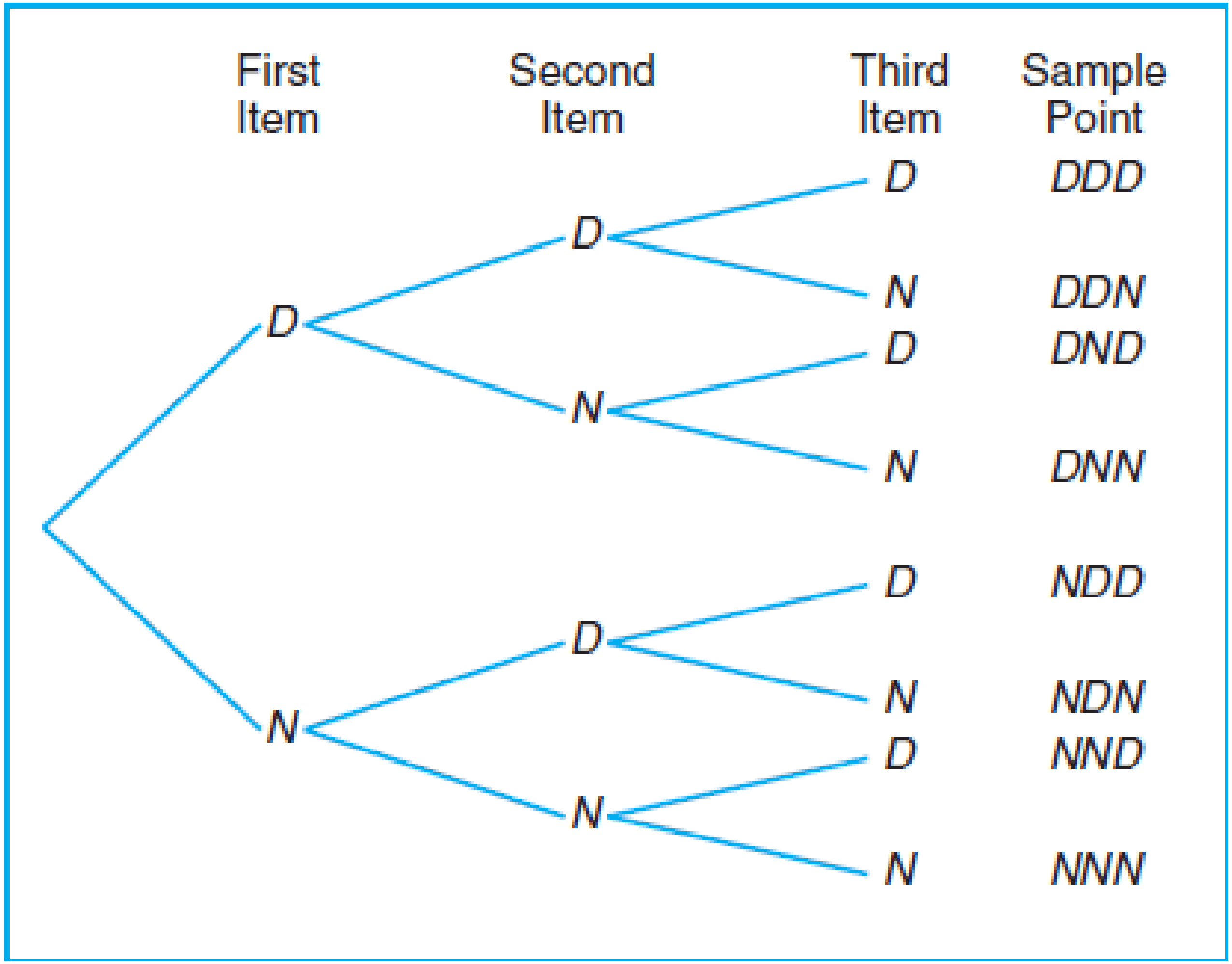


First Outcome
Second Outcome
Sample Point
H
H
HH
T
HT
T
1
T1
2
T2
3
T3
4
T4
5
T5
6
T6

Tree Diagram [10]

Example Suppose that **three items** are selected at random from a manufacturing process. Each item is inspected and classified **defective**, **D**, or **nondefective**, **N**. To list the elements of the sample space providing the most information, construct the **tree diagram.**

Solution



First Item
Second Item
Third Item
Sample Point
D
D
D
DDD
N
DDN
N
D
DND
N
DNN
N
D
D
NDD
N
NDN
N
D
NND
N
NNN

Tables [1]

Another way to find a sample space is to use a **table.**

Example: Find the sample space for selecting a card from a standard deck of 52 cards.

- There are four suits—**hearts** and **diamonds**, which are **red**, and **spades** and **clubs**, which are **black**.

- Each suit consists of 13 cards—ace through king. **Face cards** are kings, queens, and jacks.

A standard deck of 52 cards

Heart	A ♥	2 ♥	3 ♥	4 ♥	5 ♥	6 ♥	7 ♥	8 ♥	9 ♥	10 ♥	J ♥	Q ♥	K ♥
Diamond	A ♦	2 ♦	3 ♦	4 ♦	5 ♦	6 ♦	7 ♦	8 ♦	9 ♦	10 ♦	J ♦	Q ♦	K ♦
Spade	A ♠	2 ♠	3 ♠	4 ♠	5 ♠	6 ♠	7 ♠	8 ♠	9 ♠	10 ♠	J ♠	Q ♠	K ♠
Club	A ♣	2 ♣	3 ♣	4 ♣	5 ♣	6 ♣	7 ♣	8 ♣	9 ♣	10 ♣	J ♣	Q ♣	K ♣

Tables [2]

Example: A single card is drawn at random from a standard deck of cards. Find the probability that it is

a. The 4 of diamonds.

b. A queen.

Solution:

a. P(The 4 of diamonds) = ${}^{1}/_{52}$ = **0.0192 (or 1.9231%)**

b. P(A queen) = ${}^{4}/_{52}$ = ${}^{1}/_{13}$ = **0.0769 (or 7.6923 %)**

Tables [3]

A table can be used for the sample space when two dice are rolled.

	Die 2					
Die 1	1	2	3	4	5	6
1	(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)
2	(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
4	(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
5	(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
6	(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

Tables [4]

Example: When two dice are rolled, find the probability of getting a sum of nine.

Solution:

Let A be the event of getting a "sum of 9"

P(A) = $^{4}/_{36}$ = $^{1}/_{9}$ = **0.1111 (or 11.11 %)**

Intersection [1]

Intersection: The **intersection** of two events A and B, denoted by the symbol **$A \cap B$**, is the event containing all elements that are **common to A and B**.

Example: Let **E** be the event that a person selected at random in a classroom is majoring in **engineering**, and let **F** be the event that the person is **female**. Then **$E \cap F$** is the event of all female engineering students in the classroom.

.

Intersection [2]

Example: Let **$V = \{a, e, i, o, u\}$** and **$C = \{l, r, s, t\}$;** then it follows that **$V \cap C = \emptyset$**. That is, V and C have no elements in common and, therefore, **cannot both simultaneously occur**.

Union

Union: The **union** of the two events A and B, denoted by the symbol **AUB**, is the event containing all the elements that belong to A or B or both.

Notation for Addition Rule

P(A or B) = **P**(in a single trial, event *A* occurs or event *B* occurs or they both occur)

Mutually Exclusive or Disjoint

Two events A and B are **mutually exclusive, or disjoint,**

if **$A \cap B$ = { }** or **Ø**

OR

Two events A and B are **mutually exclusive**, **or disjoint,**

if **$A \cap B$ = Ø**, that is, if A and B have no elements in common.

OR

Events ***A*** and ***B*** are **disjoint** (or **mutually exclusive**) if they **cannot occur at the same time**. (That is, disjoint events do not overlap.)

Addition Rule I

Addition Rule I: When two events are **mutually exclusive** or **disjoint events**

P(A or B) = P(A) + P(B)

OR

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

Example: When a die is rolled, find the probability of getting a **2** or a **3**.

Solution:

S = { 1, 2, 3, 4, 5, 6}; n(S) = 6

Let **A** be the event of getting a **"2"**

A = {2}; n(A) = 1

P(A) = n(A)/n(S) = $\frac{1}{6}$ = **0.1667 (or 16.67%)**

Let **B** be the event of getting a **"3"**

B = {3}; n(B) = 1

P(B) = n(B)/n(S) = $\frac{1}{6}$ = **0.1667 (or 16.67%)**

Since events **A** and **B** are **mutually exclusive**, so

P(A ∪ B) = $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$ = 0.3333 **(or 33.33%)**

Addition Rule I [2]

A cable television company offers programs on **eight** different channels, **three** of which are affiliated with **ABC**, **two** with NBC, and **one** with **CBS**. The other **two** are an **educational channel** and the **ESPN sports** channel. Suppose that a person subscribing to this service turns on a television set without first selecting the channel. Let ***A*** be the event that the program belongs to the **NBC network** and ***B*** the event that it belongs to the **CBS network**. Since a television program cannot belong to more than one network, the events *A* and *B* have no programs in common.

Therefore, the intersection $\mathbf{A \cap B}$ contains no programs, and consequently the events *A* and *B* are **mutually exclusive**.

Addition Rule I [3]

Example: In a committee meeting, there were **5** freshmen, **6** sophomores, **3** juniors, and **2** seniors. If a student is selected at random to be the chairperson, find the probability that the chairperson is a **sophomore** or a **junior**.

Addition Rule I [4]

Solution:

Let A be the event of selecting a chairperson as a **"sophomore"**

$P(A) = \frac{6}{16} = \frac{3}{8} =$ **0.3750**

Let B be the event of selecting a chairperson as a **"junior"**

$P(B) = \frac{3}{16} =$ **0.1875**

Since **A** and **B** are **mutually exclusive or disjoint events**

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

$= \frac{6}{16} + \frac{3}{16} = \frac{9}{16} =$ **0.5625 (or 56.25%)**

Addition Rule I [5]

Example: A card is selected at random from a deck. Find the probability that the card is an **ace or** a **king**.

Heart	A ♥	2 ♥	3 ♥	4 ♥	5 ♥	6 ♥	7 ♥	8 ♥	9 ♥	10 ♥	J ♥	Q ♥	K ♥
Diamond	A ♦	2 ♦	3 ♦	4 ♦	5 ♦	6 ♦	7 ♦	8 ♦	9 ♦	10 ♦	J ♦	Q ♦	K ♦
Spade	A ♠	2 ♠	3 ♠	4 ♠	5 ♠	6 ♠	7 ♠	8 ♠	9 ♠	10 ♠	J ♠	Q ♠	K ♠
Club	A ♣	2 ♣	3 ♣	4 ♣	5 ♣	6 ♣	7 ♣	8 ♣	9 ♣	10 ♣	J ♣	Q ♣	K ♣

Solution:

Let **A** be the event of selecting an **ace**

P(A) = $\frac{4}{52} = \frac{1}{13}$ = **0.0769 (or 7.69%)**

Let **B** be the even of selecting a **king**

P(B) = $\frac{4}{52} = \frac{1}{13}$ = **0.0769 (or 7.69%)**

Since **A** and **B** are **mutually exclusive or disjoint events**

P(AUB) = P(A) + P(B)

$= \frac{4}{52} + \frac{4}{52}$ = 0.1538 **(or 15.38%)**

Addition Rule II [1]

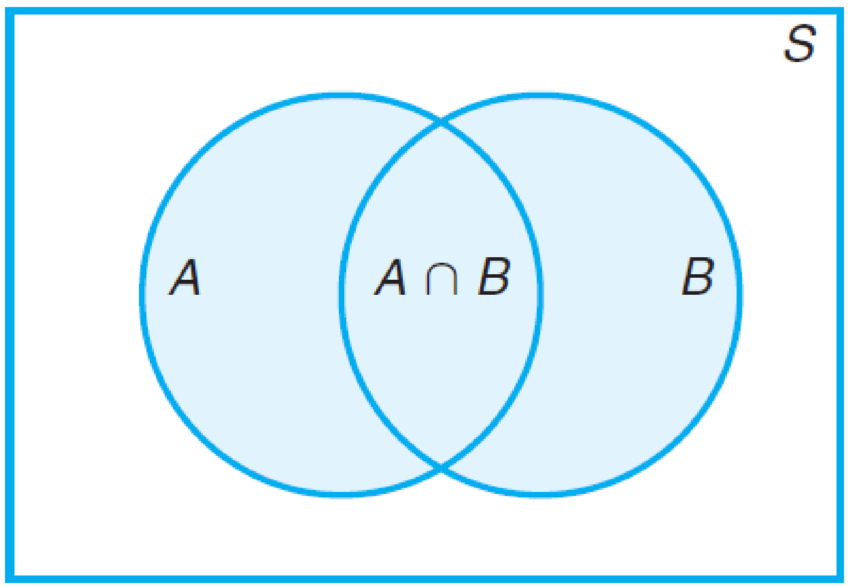
- When two events are **not mutually exclusive**, you need to add the probabilities of each of the two events and **subtract the probability of the outcomes that are common** to both events. In this case, addition **rule II** can be used.

- **Addition Rule II:** If A and B are two events that are not mutually exclusive, then

$$P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$$

When A and B are two events that are not mutually exclusive

If A and B are two events, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.



Additive rule of probability

Addition Rule II [2]

Example: A card is selected at random from a deck of 52 cards. Find the probability that it is a **6** or a **diamond**.

Addition Rule II [3]

Solution:

Let A be the event of getting a **"6".**

P(A) = $\frac{4}{52} = \frac{1}{13}$ = **0.0769 (or 7.69%)**

Let B be the event of getting a **"diamond".**

P(B) = $\frac{13}{52} = \frac{1}{4}$ = **0.2500 (or 25%)**

Addition Rule II [4]

Let **A ∩ B** be the event of getting a **"6"** and a **"diamond"**

P(A∩B) = $\frac{1}{52}$ = **0.0192 (or 1.9231%)**

Since A and B are **not mutually exclusive**, so

P(AUB) = P(A) + P(B) – P(A ∩ B)

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

$$= \mathbf{\frac{4}{13}} = \mathbf{0.3077\ (or\ 30.77\%)}$$

Addition Rule II [5]

Example: A die is rolled. Find the probability of getting an even number or a number less than 4.

Addition Rule II [6]

Solution:

S = {1, 2, 3, 4, 5, 6}

n(S) = 6

Let A be the event of getting an even number

A = { 2, 4, 6}, n(A) = 3

P(A) = $\frac{3}{6} = \frac{1}{2}$ **= 0.50 (or 50%)**

Let B be the event of getting a number less than 4

B = {1, 2, 3}, n(B) = 3

P(B) = $\frac{3}{6} = \frac{1}{2}$ **= 0.50 (or 50%)**

Addition Rule II [7]

Let $A \cap B$ be the event of getting an **"even number"** and a **"number less than 4"**

$A \cap B = \{2\}$

P(A∩B) $= \frac{1}{6} = 0.1667$ or (16.67%)

Since A and B are **not mutually exclusive**, so

P(A∪B) = P(A) + P(B) – P(A ∩ B)

$$= \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \mathbf{\frac{5}{6} = 0.8333\ (or\ 83.3333\ \%)}$$

Table

A table can be used for the sample space when two dice are rolled.

	Die 2					
Die 1	1	2	3	4	5	6
1	(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)
2	(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
4	(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
5	(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
6	(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

Addition Rule II [8]

Example: Two dice are rolled; find the probability of getting **doubles** or a **sum of 8**.

Addition Rule II [9]

Solution:

Let A be the event of getting doubles

A = {(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)}, n(A) = 6

P(A) = $\frac{6}{36} = \frac{1}{6}$ = **0.1667 (or 16.67%)**

Let B be the event of getting a sum of 8

B = {(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)}, n(A) = 5

P(B) = $\frac{5}{36}$ = **0.1389 (or 13.89%)**

Addition Rule II [10]

Let **A∩B** be the event of getting a **'doubles'** and a **'sum of 8'**

A∩B = {(4, 4)}

P(A∩B) = $\frac{1}{36}$ = 0.0277 (or 2.7777 %)

Since A and B are **not mutually exclusive**, so

P(AUB) = P(A) + P(B) – P(A ∩ B)

$$= \frac{6}{36} + \frac{5}{36} - \frac{1}{36} = \frac{10}{36}$$

$= \frac{5}{18} =$ **0.2777or (27.7777%)**

Addition Rule II [11]

Let **P** be the event that an employee selected at random from an oil drilling company **smokes cigarettes**.

Let **Q** be the event that the employee selected drinks **alcoholic beverages**.

Then the event **P U Q** is the set of all employees who either **drink** or **smoke** or do **both**.

Example : A coin is tossed twice. What is the probability that at **least 1 head** occurs?

Solution : The sample space for this experiment is

S = {HH, HT, TH, TT}

n(S) = 4

Let **A** be the event of getting at **least 1 head**

A = {HH, HT, TH}

n(A) = 3

$$\therefore P(A) = \frac{n(A)}{n(s)}$$

$$= \frac{3}{4} = \mathbf{(0.75 \text{ or } 75\%)}$$

Example : A die 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 less than 4** occurs on a single toss of the die, find ***P*(*E*).**

Solution

P(Even number) = 2p

P(Odd number) = p

S = { 1, 2, 3, 4, 5 ,6}

$\because$ Sum of probability $= 1$

$\therefore$ **p + 2p + p + 2p + p + 2p** = 1

$\Rightarrow 9\mathrm{p} = 1$

$\Rightarrow \mathbf{p = \frac{1}{9}}$

$\Rightarrow$ P(Even number) = $\mathbf{\frac{2}{9}}$

$\Rightarrow$ P(Odd number) = $\mathbf{\frac{1}{9}}$

E = {1, 2, 3}

P(E) = P(1) + P(2) + P(3)

$= \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \mathbf{\frac{4}{9}}$ (or 0.4444 or 44%)

Example A die is loaded in such a way that **an even number** is **twice** as likely to occur as an odd number.

Let ***A*** be the event that an **even number** turns up and let ***B*** be the event that a **number divisible** by **3** occurs. Find $\mathbf{P(A \cup B)}$ and $P(A \cap B)$.

Solution

S = {1, 2, 3, 4, 5, 6}

P(Even number) = 2p

P(Odd number) = p

S = { 1, 2, 3, 4, 5 ,6}

$\because$ Sum of probability $= 1$

$\therefore$ **p + 2p + p + 2p + p +2p** = 1

$\Rightarrow 9p = 1$

$\Rightarrow \mathbf{p = \frac{1}{9}}$

$\Rightarrow$ P(Even number) = $\frac{2}{9}$

$\Rightarrow$ P(Odd number) = $\frac{1}{9}$

A = { 2, 4, 6}

P(A) = $\frac{2}{9} + \frac{2}{9} + \frac{2}{9} = \frac{6}{9} = \frac{2}{3}$

B = {3, 6}

P(B) = $\frac{1}{9} + \frac{2}{9} = \frac{3}{9} = \frac{1}{3}$

$A \cap B = \{6\}$

$P(A \cap B) = \frac{2}{9}$ **$\because P(\text{Even number}) = \frac{2}{9}$**

Since A and B are **not mutually exclusive**, so

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

$= \frac{6}{9} + \frac{3}{9} - \frac{2}{9} = \frac{7}{9}$ (or 0.7778 or 77.7778%)

- To find ***P*(*A* or *B*),** begin by associating use of the word **"or"** with addition.

- Consider whether **events *A*** and ***B*** are **disjoint**; that is, can they happen at the same time?

- If they are **not disjoint** (that is, they can happen at the same time), be sure to avoid (or at least compensate for) **double-counting** when adding the relevant probabilities.

- If you understand the importance of not double counting when you find *P*(*A* or *B*), you don't necessarily have to calculate the value of **$P(A) + P(B) - P(A \cap B)$**

Errors made when applying the addition rule

- Errors made when applying the addition rule often involve **double-counting**; that is, events that **are not disjoint** are treated as if they were. One indication of such an error is a total probability that **exceeds 1.**

- However, errors involving the addition rule do not **always cause** the total probability to **exceed 1**.

Suggested Readings

Probability & Statistics for Engineers & Scientists, Ninth Edition, Ronald E. Walpole, Raymond H. Myer

2.1 Sample space

2.2 Events

2.3 Counting Sample Points

2.4 Probability of an Event

2.5 Additive Rules