Philadelphia University

Lecture Notes for 650364

Probability & Random Variables

Chapter 1:

Lecture 2: Probability Definition, Joint, Marginal and Conditional Probability

Department of Communication & Electronics Engineering

Instructor

Dr. Qadri Hamarsheh

Email: qhamarsheh@philadelphia.edu.jo
Website: http://www.philadelphia.edu.jo/academics/qhamarsheh

Probability

1) Introduction

2) Set Definitions

3) Set Operations

4) Probability Introduced Through Sets and Relative Frequency

5) Joint and Conditional Probability

6) Total Probability and Bayes' Theorem

7) Independent Events

8) Combined Experiments

9) Bernoulli Trials

4)Probability Introduced through Sets and Relative Frequency

- ✓ Most **three approaches** used for the definition and discussion of probability are:
  - **Classical Probability** based on sample space (mathematical approach based on probability theory, **set theory** and using the axiomatic definition of probability.
  - **Relative frequency or empirical probability** which is based more on engineering or scientific **observations.**
  - **Subjective probability** made by a person's knowledge of the situation.
- ✓ **Probability Experiments**:
  - **Chance processes**, such as flipping a coin, rolling a die, or drawing a card at random from a well-shuffled deck are called probability experiments. A probability experiment is a chance process that leads to **well defined outcomes** or results.
- ✓ **An outcome** of a probability experiment is the result of a single trial of a probability experiment.

- Each outcome of a probability experiment occurs at **random**. This means you cannot predict with certainty which outcome will occur when the experiment is conducted.
- Each outcome of the experiment is **equally likely** unless otherwise stated. That means that each outcome has the same probability of occurring.
- **Sample Points**: These outcomes are called sample points.

✓ The set of all possible outcomes (realizations) of a probability experiment is called a **sample space** and denoted by **S**.

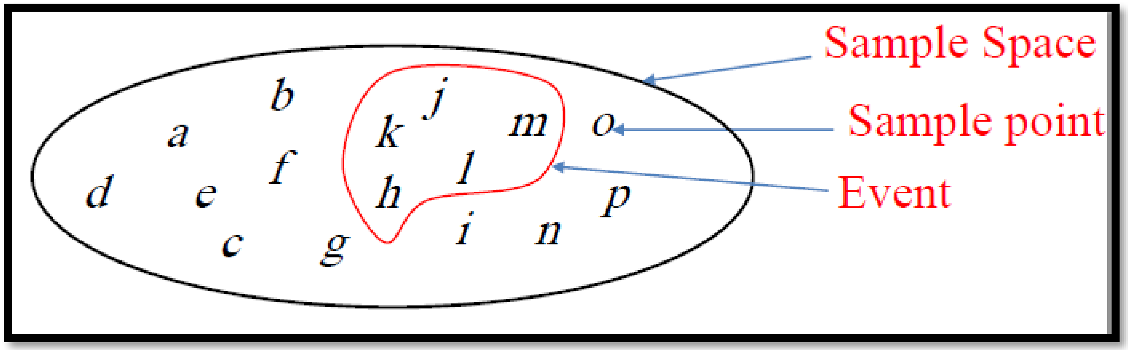
- Sample spaces for various probability experiments

Experiment	Sample Space
Toss one coin	H, T*
Roll a die	1, 2, 3, 4, 5, 6
Toss two coins	HH, HT, TH, TT

*H = heads; T = tails.

✓ An **event** then usually consists of **one** or **more outcomes** of the sample space. Certain **subsets** of **S** are referred to as events.

- An event with one outcome is called a **simple event**.
- When an event consists of two or more outcomes, it is called a **compound event**.
- **Mutually exclusive**: If two events have no common outcomes they are called **mutually exclusive**.
- Events A and B are mutually exclusive, if only one of them can occur, that is,



✓ **Example: Dice**

- Sample space $S = \{1;\ 2; 3; 4; 5\ ; 6\}$
- Events are the subsets of **S**, e.g.,
  - **A = "outcome is even"** = $\{2;\ 4;\ 6\}$.
  - **B = "outcome is > 4"** = $\{5;\ 6\}$.

✓ **Example: Rolling two dice**

- Sample space is

$$S = \{(1;\ 1);\ (1;\ 2);\ (1;\ 3);\ (1;\ 4);\ (1;\ 5);\ (1;\ 6);$$
$$(2;\ 1);\ (2;\ 2);\ (2;\ 3);\ (2;\ 4);\ (2;\ 5);\ (2;\ 6);$$
$$(3;\ 1);\ (3;\ 2);\ (3;\ 3);\ (3;\ 4);\ (3;\ 5);\ (3;\ 6);$$
$$(4;\ 1);\ (4;\ 2);\ (4;\ 3);\ (4;\ 4);\ (4;\ 5);\ (4;\ 6);$$
$$(5;\ 1);\ (5;\ 2);\ (5;\ 3);\ (5;\ 4);\ (5;\ 5);\ (5;\ 6);$$
$$(6;\ 1);\ (6;\ 2);\ (6;\ 3);\ (6;\ 4);\ (6;\ 5);\ (6;\ 6)\}:$$

- Events are all subsets of **S**, e.g.,
  - **A = "outcomes are the same":**

$$= \{(1;\ 1);\ (2;\ 2);\ (3;\ 3);\ (4;\ 4);\ (5;\ 5);\ (6;\ 6)\}.$$

  - **B = "outcome of the 1. roll is 1":**

$$= \{(1;\ 1);\ (1;\ 2);\ (1;\ 3);\ (1;\ 4);\ (1;\ 5);\ (1;\ 6)\}.$$

✓ **Discrete Sample Space:**

- **Example**: Experiment of rolling a single die and observing the number that shows up.

There are six elements $S = \{1, 2, 3, 4, 5, 6\} \leftarrow$ Discrete Sample Space.

✓ **Continuous Sample Space:**

- **Example**: Experiment of obtaining a number by spinning the pointer on a wheel of chance numbered from 0 to 12.

$S = \{0 \le s \le 12\} \leftarrow$ Continuous Sample Space

✓ **Probability Definition and Axioms:**

- Let S be the sample space of a random Experiment. Suppose that to each event A of S, a number denoted by $P(A)$ is associated with A. If P satisfies the following axioms, then it is called a probability and the number $P(A)$ is said to be the probability of A.

Axiom 1: $P(A) \ge 0$

Axiom 2: $P(S) = 1$

Axiom 3: If $\{A_1, A_2, A_3, \ldots\}$ is a sequence of mutually exclusive events, then

$$P\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} P(A_n)$$

Note that:

Equally likely: $P(A) = P(B)$

The probability of the empty set ϕ is 0. That is, $P(\phi) = 0$.

- **Sample spaces** are used in **classical probability** to determine the **numerical probability** that an event will occur. The formula for determining the probability of an event **E** is

$$\boxed{P(E) = \frac{\text{number of outcomes contained in the event } E}{\text{total number of outcomes in the sample space}}}$$

✓ **Mathematical Model of Experiments:**

- A real experiment is defined mathematically by three things:
  - **Assignment of the sample space S**
  - **Definition of events of interest**
  - **Making probability assignments to the events such that the axioms are satisfied.**

✓ **Example**: A die is tossed; find the probability of each event:

a) Getting a two
b) Getting an even number
c) Getting a number less than 5

- **Solution**: The sample space is **1, 2, 3, 4, 5, 6**, so there are six outcomes in the sample space.

a) P(2) = 1/6.
b) P(even number)= 3/6=1/2;

c) P(number less than 5) = 4/6= 2/3;

- ✓ **Example:** The experiment of tossing a coin once, the sample space is $S = \{H, T\}$. Let events $A = \{H\}$ and $B = \{T\}$ then:

  $P(A) = P(B) = 1/2 = 0.5$ and $P(S) = 1$

- ✓ **Example:** The experiment of rolling a single die, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Let events $A = \{1, 3, 5\}$ and $B = \{2, 4\}$ then:

$P(A) = 3/6 = 0.5$, $P(B) = 2/6$, $P(AUB) = P(A) + P(B) = 5/6$

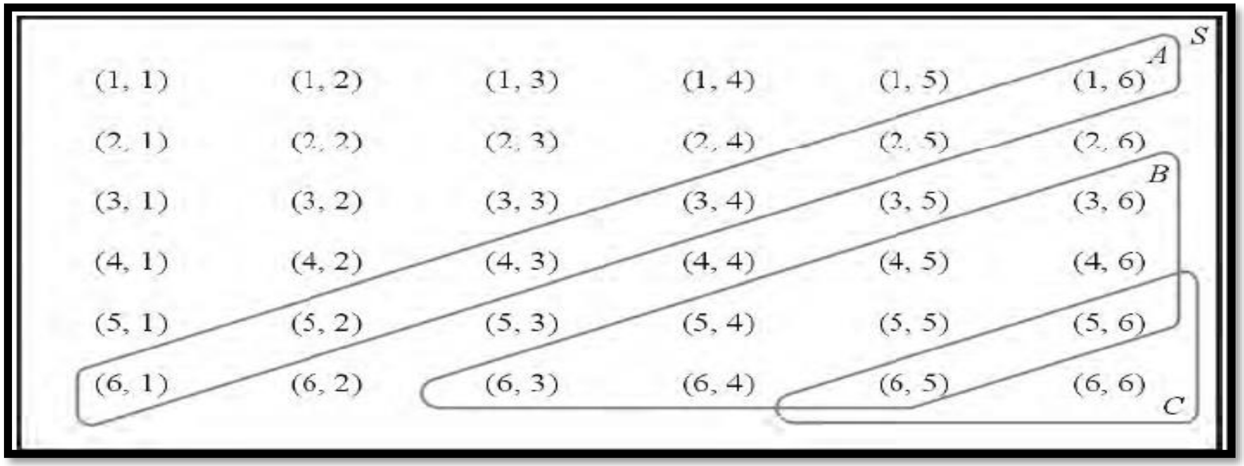
- ✓ **Example:** An experiment consists of observing the sum of the numbers showing up when two dice are thrown. If three events define by:

  $A = \{sum = 7\}$,

  $B = \{8 < sum \leq 11\}$,

  and $C = \{10 < sum\}$, then

  $P(A) = 6/36$, $P(B) = 9/36$ *and* $P(C) = 3/36$



✓ The **Relative frequency** probability P of event A is defined as

$$\boxed{P(A) = \lim_{n \to \infty} \frac{n_A}{n}}$$

- The ratio $\frac{n_A}{n}$ is the **relative frequency** (or average number of success) for the event A.

✓ **Example:** If a fair coin is flipped n times, the side that shows up will be "**heads**" about n_H times, then

$$\boxed{P(H) = \lim_{n \to \infty} \frac{n_H}{n} = \frac{1}{2}}$$

✓ **Example**:

a) In a box there are **80** resistors each having the same size and shape. From the 80 resistors **18** are **10Ω**, **12** are **22Ω**, **33** are **27Ω**, and **17** are **47Ω**. If the experiment is to randomly draw out one resistor from the box with each one being "**equally likely**" to be drawn, then

P(draw 10Ω) = 18/80 ,

P(draw 22Ω) = 12/80

P(draw 27Ω) = 33/80 ,

P(draw 47Ω) = 17/80

b) Next suppose a **22Ω** resistor is drawn from the box and not replaced, a second resistor is then drawn from the box. What are the probabilities of drawing a resistor of any one of the four values?

P(draw 10Ω / 22Ω) = 18/79 ,

P(draw 22Ω / 22Ω) = 11/79

P(draw 27Ω / 22Ω) = 33/79 ,

P(draw 47Ω / 22Ω) = 17/79

- ✓ Probabilities can be computed for situations that do not use sample spaces. In such cases, **frequency distributions** are used and the probability is called empirical probability.
- ✓ Empirical probability is sometimes called relative frequency probability

$$P(E) = \ frequency\ of\ E/sum\ of\ the\ frequencies$$

- ✓ **Example**: Suppose a class of students consists of 4 freshmen, 8 sophomores, 6 juniors, and 7 seniors. The information can be summarized in a frequency distribution as follows:

Rank	Frequency
Freshmen	4
Sophomores	8
Juniors	6
Seniors	7
TOTAL	25

Using the frequency distribution shown previously, find the probability of selecting a junior student at random.

Solution: Since there are **6 juniors** and a total of **25 students**,

P(junior)=6/25

✓ **Combining events | Summary**

Interpretation	Set theory expression
Certain event	S
Impossible event	$\emptyset$
A occurs	A
A and B occur	$A \cap B$
A or B occur	$A \cup B$
A does not occur	A^c
B occurs but A does not	$B \setminus A$
A and B are mutually exclusive	$A \cap B = \emptyset$

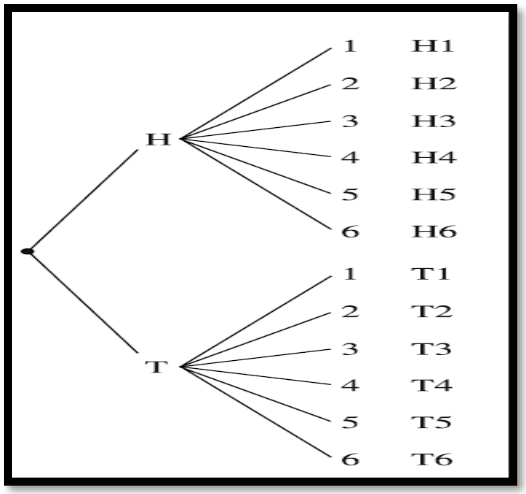
✓ **Finding Sample spaces | Summary**

✓ **Two specific devices** 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.

- In order to **construct a tree diagram**, use branches corresponding to the outcomes of the first experiment. 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.

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



Hence there are twelve outcomes. They are **H1, H2, H3, H4, H5, H6, T1,T2, T3, T4, T5, and T6**.

Once the sample space has been found, probabilities for events can be computed.

- **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) Since there are 12 outcomes in the sample space and only one way to get a head on the coin and a three on the die, **P(H3)=1/12**

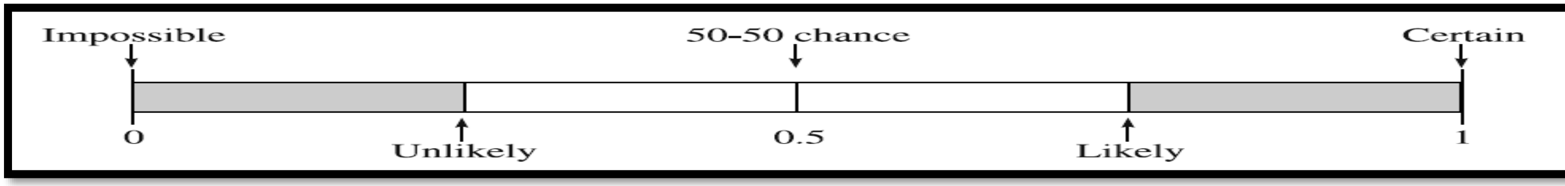
b) Since there are six ways to get a head on the coin, namely H1, H2, H3,H4, H5, and H6, **P(head on the coin) =6/12=1/2**

c) Since there are two ways to get a 4 on the die, namely H4 and T4, **P(4 on the die) = 2/12= 1/6**

- ✓ A **table** can be used for finding the sample space
- ✓ **Example**: when two dice are rolled. Since the first die can land in 6 ways and the second die can land in 6 ways, there are 6×6 or 36 outcomes in the sample space

	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)

- ✓ **Probability Rules** | **Summary**
- ✓ **Rule 1**: The probability of any event will always be a number from **zero to one**. This can be denoted mathematically as $\mathbf{0 \leq P(E) \leq 1}$. Probabilities **cannot be negative** nor can they be **greater than one**



- ✓ **Rule 2:** When an **event cannot occur**, the probability will be **zero**.
  - **Example**: A die is rolled; find the probability of getting a 7.

    Solution: Since the sample space is **1, 2, 3, 4, 5**, and **6**, and there is no way to get a 7, **P(7)=0**.
- ✓ **Rule 3:** When an event is certain to occur, the probability is 1.
  - **Example**: A die is rolled; find the probability of getting a number less than 7.

    Solution: Since all outcomes in the sample space are less than 7, the probability is **6/6=1.**
- ✓ **Rule 4:** The sum of the probabilities of all of the outcomes in the sample space is 1.

Referring to the sample space for tossing two coins (HH, HT, TH, TT), each outcome has a probability of 1/4 and the sum of the probabilities of all of the outcomes is

$$\mathbf{1/4 + 1/4 + 1/4 + 1/4 = 1}$$

- ✓ **Rule 5:** The probability that an event will not occur is equal to 1 minus the probability that the event will occur.

5)Joint and Conditional Probability

- ✓ In many engineering applications we often perform an experiment that consists of many experiments, examples are:
  - Observation the input and output digits of a binary communication system.
  - Observation of the trajectories of several objects in space.
- ✓ Suppose: E is the random experiment contains two subexperiments E_1 and E_2.
  - If
    - S_1 sample space for E_1 with $\{a_1, a_2, \ldots, a_{n1}\}$ outcomes.
    - S_2 sample space for E_2 with $\{b_1, b_2, \ldots, b_{n2}\}$ outcomes.
  - Then

- the sample space of the combined experiment is the Cartesian product:

$$S_1 * S_2 = \{(a_i, b_j): i = 1, 2, \dots, n_1; j = 1, 2, \dots, n_2\}$$

- If events $A_1, A_2, \dots, A_n$ are defined for S_1
- If events $B_1, B_2, \dots, B_m$ are defined for S_2
- Then events $A_i B_j$ are the events of the total experiment (see the table below).
- We can define probability measures on S_1, S_2 and S as the following:

Assume that

N : Number of outcomes for S;

N_A : Number of outcomes for A_i

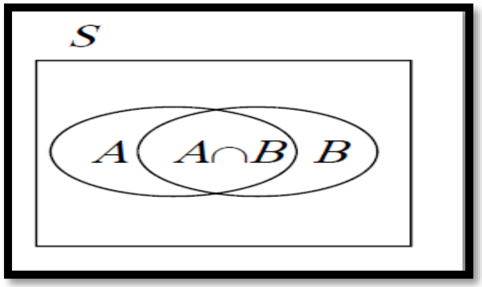
N_B : Number of outcomes for B_j

N_{AB} : Number of outcomes for $A_i \cap B_j$

- **Joint Probability**: the probability

$$P(A_i \cap B_j) = \frac{N_{AB}}{N}$$

  - The probability $P(A \cap B)$ is called the **joint probability** for two events A and B.



- It can be shown that:

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

- Equivalently:

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

- For **mutually exclusive** events:

$$P(A \cup B) = P(A) + P(B) \text{ where } P(A \cap B) = 0$$

- **Marginal Probability :** if events $A_1, A_2, \ldots, A_n$ are mutually exclusive and exhaustive, then

$$B_j = \sum_{i=1}^{n} P(A_i \cap B_j) \text{ Marginal Probability for } B_j$$

And

$$A_i = \sum_{j=1}^{m} P(A_i \cap B_j) \text{ Marginal Probability for } A_i$$

Experiment E		Experiment E_1				
		A_1	A_2	...	A_n	Marginal Prob. For B_j
Experiment E_2	B_1	$A_1 \cap B_1$	$A_2 \cap B_1$	...	$A_n \cap B_1$	$P(B_1)$
	B_2	$A_1 \cap B_2$	$A_2 \cap B_2$	...	$A_n \cap B_2$	$P(B_2)$
	...	...	...	...	...	...
	B_m	$A_1 \cap B_m$	$A_2 \cap B_m$	...	$A_n \cap B_m$	$P(B_m)$
	Marginal Prob. For A_i	$P(A_1)$	$P(A_1)$	...	$P(A_n)$	1

$$P(B_1) = P(A_1 \cap B_1) + P(A_2 \cap B_1) + \cdots + P(A_n \cap B_1)$$

$$P(B_2) = P(A_1 \cap B_2) + P(A_2 \cap B_2) + \cdots + P(A_n \cap B_2)$$

...

$$P(B_m) = P(A_1 \cap B_m) + P(A_2 \cap B_m) + \cdots + P(A_n \cap B_m)$$

$$P(B_j) = \sum_{i=1}^{n} P(A_i \cap B_j)$$

$$P(A_i) = \sum_{j=1}^{m} P(A_i \cap B_j)$$

$$i = 1, 2, \dots, n \text{ and } j = 1, 2, .., m$$

- **Conditional Probability**: The probability $P(A|B)$ is called the **conditional probability** for two events A and B.
  - Given some event **B** with nonzero probability; $P(B) \neq 0$. The conditional probability of A given B is:

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

  - An expression for the **Conditional Probability** in terms of joint and marginal probabilities

    $$P(B|A) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{N_{AB}}{N}}{\frac{N_A}{N}} = \frac{N_{AB}}{N_A}, P(A) \neq 0$$

    $$\text{Conditional probability} = \frac{\textit{Joint} \text{ probability}}{\text{Marginal probability}}$$

  - Since **B** is known to have occurred, it becomes the **new sample space** replacing the original.

- For **mutually exclusive** events A and B:

$$A \cap B = \emptyset, and\ P(A|B) = 0$$

✓ **Example:** The HBO cable network took a survey of **500** subscribers to determine people's favorite show, the data are shown in the following table: (for details click here to see the YouTube video)

	Male	Female
Game of Thrones	80	120
West World	100	25
Others	50	125

- Compute the total sum for each row and column, we obtain the following table:

	Male	Female	Total
Game of Thrones	80	120	200
West World	100	25	125
Others	50	125	175
Total	230	270	500

- Each square can be called **joint event**, because the data written in the square depended on two different variables.
- Obtain the **probability distribution table** by dividing each value by the total number of subscribers "observations".

	Male (M)	Female (F)	Total
Game of Thrones (GoT)	0.16	0.24	0.4
West World (WW)	0.2	0.05	0.25
Others (Ot)	0.1	0.25	0.35
Total	0.46	0.54	1

- 0.24 is the **joint probability** of the intersection of **Female** event (**F**) and the **Game of Thrones** (**GoT**) event and can be written as $P(\text{F AND GoT}) = P(\text{F} \cap \text{GoT}) = 0.24$.
- **Joint Probability Distribution** is the collectively of the six probabilities **0.16, 0.2, 0.1, 0.24, 0.05** and **0.25**

$$P(M \cap GoT) + P(M \cap WW) + P(M \cap Ot) + P(F \cap GoT) + P(F \cap WW) + P(F \cap Ot) = 1$$

- Computation of the **Marginal Probabilities**

	Male (M)	Female (F)	Total
Game of Thrones (GoT)	0.16	0.24	0.4
West World (WW)	0.2	0.05	0.25
Others (Ot)	0.1	0.25	0.35
Total	0.46	0.54	1

- 0.4 is the **marginal probability** (called like that because it is in the margin of the table

$$P(GoT) = 0.4$$

- The **marginal probability** is the **sum of the joint Probabilities** of the given variable.

$$P(GoT) = P(M \cap GoT) + P(F \cap GoT) = 0.16 + 0.24 = 0.4$$

- **Marginal Probability Distribution** for Game of Thrones, West World and Others

$$P(GoT) + P(WW) + P(Ot) = 1$$

	Male (M)	Female (F)	Total
Game of Thrones (GoT)	0.16	0.24	0.4
West World (WW)	0.2	0.05	0.25
Others (Ot)	0.1	0.25	0.35
Total	0.46	0.54	1

- **Marginal Probability Distribution** for Male, and Female

$$P(M) + P(F) = 1$$

	Male (M)	Female (F)	Total
Game of Thrones (GoT)	0.16	0.24	0.4
West World (WW)	0.2	0.05	0.25
Others (Ot)	0.1	0.25	0.35
Total	0.46	0.54	1

- **Questions:**
  - **Q1**: What is the probability of an HBO subscriber being **male**?

$\boldsymbol{P(Male) = 0.46}$ Marginal Probability of Male Column.

$$\boldsymbol{or}$$

$$\boldsymbol{P(Male) = P(M \cap GoT) + P(M \cap WW) + P(M \cap Ot)}$$
$$\boldsymbol{= 0.16 + 0.2 + 0.1 = 0.46}$$

  - **Q2**: What is the probability of an HBO subscriber **preferring West World**?

$$\boldsymbol{P(WW) = P(M \cap WW) + P(F \cap WW) = 0.2 + 0.05 = 0.25}$$

  - **Q3**: What is the probability of an HBO subscriber being **male AND preferring West World**?

$\boldsymbol{P(M \cap WW) = 0.2}$ (Joint Prob.)

  - **Q4**: What is the probability of an HBO subscriber being **male OR preferring West World**?

    It is the union of probabilities Male and West World

$$\boldsymbol{P(M \cup WW) = P(M) + P(WW) - P(M \cap WW)}$$
$$\boldsymbol{= 0.46 + 0.25 - 0.2 = 0.51}$$

Remark: **0.2** is subtracted because this value is added twice in the calculation process.

OR

$$P(M \cup WW) = P(M \cap GoT) + P(M \cap WW) + P(M \cap Ot) + P(F \cap WW) = 0.16 + 0.2 + 0.1 + 0.05 = 0.51$$

	Male (M)	Female (F)	Total
Game of Thrones (GoT)	0.16	0.24	0.4
West World (WW)	0.2	0.05	0.25
Others (Ot)	0.1	0.25	0.35
Total	0.46	0.54	1

- **Q5**: Fatimah just got an HBO subscription. What is the chance that her favorite show will be Game Of Thrones?

Because Fatimah is Female, we take into consideration only the **Female** Column

	Female (F)
Game of Thrones (GoT)	0.24
West World (WW)	0.05
Others (Ot)	0.25
Total	0.54

$$P(GoT|F) = \frac{P(GoT \cap F)}{P(F)} = \frac{0.24}{0.54} = 0.4444$$

○ **Conditional Probability Distribution** for Male, and Female

	Male (M)	Female (F)	Female (F) Conditional PD	Total
Game of Thrones (GoT)	0.16	0.24	0.4444	0.4
West World (WW)	0.2	0.05	0.0925	0.25
Others (Ot)	0.1	0.25	0.4629	0.35
Total	0.46	0.54	1	1

- We can analyze and compare the **conditional and marginal probabilities distributions**
- Q5: Given that a subscriber's favorite show is **West World**. What is the probability that they are **male**.

$$P(M|WW) = \frac{P(M \cap WW)}{P(WW)} = \frac{0.2}{0.25} = 0.8$$

✓ **Example:** In a box there are 100 resistors having resistance and tolerance as shown in Table below. Let a resistor be selected from the box and assume each resistor has the same **likelihood** of being chosen. Define three events:

- **A as draw a 47Ω resistor**
- **B as draw a resistor with 5% tolerance,**
- **C as draw a 100Ω resistor.**
  - From the table, the **applicable probabilities**:

TABLE 1.4-1
Numbers of resistors in a box having given resistance and tolerance

	Tolerance		
Resistance (Ω)	5%	10%	Total
22	10	14	24
47	28	16	44
100	24	8	32
Total	62	38	100

$$P(A) = P(47\Omega) = 44/100$$
$$P(B) = P(5\%) = 62/100$$
$$P(C) = P(100\Omega) = 32/100$$

- The **joint probabilities** are:

$$P(A\ \cap B) = P(47\Omega\ \cap\ 5\%) = 28/100$$
$$P(A\ \cap C) = P(47\Omega\ \cap\ 100\Omega) = 0$$
$$P(B\ \cap C) = P(5\% \cap\ 100\ \Omega) = 24/100$$

- The **conditional probabilities** are:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{28}{62}$$
$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} = 0$$
$$P(B \mid C) = \frac{P(B \cap C)}{P(C)} = \frac{24}{32}$$

✓ **Example: Dice:** $S = \{1,2,3,4,5,6\}$, $A = \{1,2,6\}$, $B = \{2,4,6\}$

What is the probability $(A|B)$?

Simple way: $P(A|B) = \dfrac{2\ (number\ of\ outcomes\ in\ A\ that\ is\ in\ B)}{3\ (number\ of\ outcomes\ in\ B\ as\ new\ sample\ space)}$

Standard way : $(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{\frac{2}{6}}{\frac{3}{6}} = \dfrac{2}{3}$

Conditional probability ---easy way :

$$P(A|B) = \frac{how\ much\ outcomes\ of\ A\ in\ B}{how\ much\ B}$$