

# **Probability and Statistics**

**Dr. Faisal Bukhari**

**Associate Professor**

**Department of Data Science**

**Faculty of Computing and Information Technology**

**University of the Punjab**

**Three things cannot be long hidden: the sun, the moon, and the truth**

**Buddha**

# Textbooks

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

❑ **Elementary Statistics: Picturing the World**, 6<sup>th</sup> Edition, Ron Larson and Betsy Farber

❑ **Elementary Statistics**, 13<sup>th</sup> 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 McMahan, 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
- ❑ An Introduction with Statistical Applications, Second Edition, John J. Kinney

These notes contain material from the above three books.

# Conditional Probability

A **conditional probability** is the probability of an event occurring, given that another event has already occurred. The conditional probability of event  $B$  occurring, given that event  $A$  has occurred, is denoted by  **$P(B|A)$**  and is read as “probability of  $B$ , given  $A$ .”

**Example:** Two cards are selected in sequence from a standard deck of 52 playing cards. Find the probability that the second card is a queen, given that the first card is a king. (Assume that the king is not replaced.)

## Solution

Because the first card is a king and is not replaced, the remaining deck has 51 cards, 4 of which are queens. So,

$$P(B|A) = \frac{4}{51} = 0.078.$$



# Independent and Dependent Events

## [1]

Two events A and B are **independent** if and only if  **$P(B|A) = P(B)$  or  $P(A|B) = P(A)$** , assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**.

OR

Two events, A and B, are said to be **independent** if the fact that **event A** occurs does not affect the probability that **event B** occurs.

# Independent and Dependent Events

## [2]

Two events are **independent** when the occurrence of one of the events does not affect the probability of the occurrence of the other event. Two events  $A$  and  $B$  are independent when

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

Events that are **not independent** are **dependent**.

# Independent and Dependent Events [3]

To determine whether  $A$  and  $B$  are independent, first calculate  $P(B)$ , the probability of **event  $B$** . Then calculate  $P(B|A)$ , the probability of  **$B$ , given  $A$** . If the values **are equal**, then the events are **independent**. If  $P(B) \neq P(B|A)$ , then  $A$  and  $B$  are dependent events.

# Independent and Dependent Events

## [4]

**Example 1:** If a coin is tossed and then a die is rolled, the **outcome of the coin in no way affects** or changes the probability of the outcome of the die.

**Example 2:** Selecting a card from a deck, replacing it, and then selecting a second card from a deck. The outcome of the first card, as long as it is **replaced**, has no effect on the probability of the outcome of the second card.

# Classifying Events as Independent or Dependent

**Example** Determine whether the events are independent or dependent.

1. Selecting a king ( $A$ ) from a standard deck of 52 playing cards, not replacing it, and then selecting a queen ( $B$ ) from the deck
2. Tossing a coin and getting a head ( $A$ ), and then rolling a six-sided die and obtaining a 6 ( $B$ )
3. Driving over 85 miles per hour ( $A$ ), and then getting in a car accident ( $B$ )

## Solution

1.  $P(B|A) = \frac{4}{51}$  and  $P(B) = \frac{4}{51}$ . The occurrence of A changes the probability of the occurrence of B, so the events are dependent.

2.  $P(B|A) = \frac{1}{6}$  and  $P(B) = \frac{1}{6}$ . The occurrence of A does not change the probability of the occurrence of B, so the events are independent.

3. Driving over 85 miles per hour increases the chances of getting in an accident, so these events are dependent.

# Independent and Dependent Events [4]

Two events A and B are **independent** if and only if  **$P(B|A) = P(B)$**  or  **$P(A|B) = P(A)$** , assuming the existences of the conditional probabilities. Otherwise, A and B are **dependent**

OR

When the occurrence of the first event in some way changes the probability of the occurrence of the second event, the two events are said to be **dependent**.

# Independent and Dependent Events [5]

**Example 1:** Suppose a card is selected from a deck and not replaced, and a second card is selected. In this case, the probability of selecting any specific card on the first draw is **52**, but since this card is not replaced, the probability of selecting any other specific card on the second draw is **51**, since there are only 51 cards left.

**Example 2:** Drawing a ball from an urn, not replacing it, and then drawing a second ball.



# First Multiplication Rule [1]

- Before explaining the first multiplication rule, consider the Example of tossing two coins. The sample space is **HH, HT, TH, TT**. From classical probability theory, it can be determined that the probability of getting two heads is  $\frac{1}{4}$ .
- However, there is another way to determine the probability of getting two heads. In this case, the probability of getting a head on the first toss is  $\frac{1}{2}$ , and the probability of getting a head on the second toss is also  $\frac{1}{2}$ .

# First Multiplication Rule [2]

□ So the probability of getting two heads can be determined by multiplying  $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ .

# Multiplication Rule I [1]

**Multiplication Rule I:** For two **independent** events A and B,

$$P(A \text{ and } B) = P(A) \times P(B).$$

**In other words**, when two independent events occur in sequence, the probability that both events will occur can be found by multiplying the probabilities of each individual event.

The word **“and”** is the key word and means that both events occur in sequence and to multiply.

# Multiplication Rule I [2]

**Example:** A coin is tossed and a die is rolled. Find the probability of getting a **tail on the coin** and a **5 on the die**.

## Solution:

Let A be the event of getting a tail on the coin

$$P(A) = 1/2 = 0.5 \text{ ( or 50\% )}$$

Let B be the event of getting a 5 on the die

$$P(B) = 1/6 = 0.1667 \text{ (or 16.67 \% )}$$

Since A and B are **independent** events, so

$$\begin{aligned} P(A \text{ and } B) &= P(A) \times P(B) \\ &= 1/2 \times 1/6 = 1/12 \\ &= 0.0833 \text{ (or 8.33 \% )} \end{aligned}$$

# Multiplication Rule 1 [3]

The previous Example can also be solved using **classical probability**. Recall that the sample space for tossing a coin and rolling a die is

$$S = \{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6\}$$

$$n(S) = 12$$

Let **A** be the event of getting a “**T5**”

$$A = \{T5\}$$

$$n(A) = 1$$

$$\begin{aligned} P(A) &= \frac{1}{12} \\ &= 0.0833 \text{ (or 8.33 \%)} \end{aligned}$$

# Multiplication Rule 1 [4]

**Example:** An urn contains **2 red balls**, **3 green balls**, and **5 blue balls**. A ball is selected at random and its color is noted. Then it is **replaced** and **another ball** is selected and its color is noted. Find the probability of each of these:

- a. Selecting **2 blue balls**
- b. Selecting a **blue ball** and then a **red ball**
- c. Selecting a **green ball** and then a **blue ball**

## Solution

Let **R** be an event of getting a red ball

Let **G** be an event of getting a green ball

Let **B** be an event of getting a blue ball

$$P(\mathbf{R}) = 2/10, P(\mathbf{G}) = 3/10, P(\mathbf{B}) = 5/10$$

**Since events are independent, so**

$$\begin{aligned} \text{a. } P(\mathbf{B} \text{ and } \mathbf{B}) &= P(\mathbf{BB}) = P(\mathbf{B}) \times P(\mathbf{B}) = 5/10 \times 5/10 = 1/4 \\ &= \mathbf{0.25 \text{ (or 25\%)}} \end{aligned}$$

$$\begin{aligned} \text{b. } P(\mathbf{B} \text{ and } \mathbf{R}) &= P(\mathbf{B}) \times P(\mathbf{R}) = 5/10 \times 2/10 = 1/10 \\ &= \mathbf{0.10 \text{ (or 10 \%)}} \end{aligned}$$

$$\begin{aligned} \text{c. } P(\mathbf{G} \text{ and } \mathbf{B}) &= P(\mathbf{G}) \times P(\mathbf{B}) = 3/10 \times 5/10 = 3/20 \\ &= \mathbf{0.15 \text{ (or 15 \%)}} \end{aligned}$$



# Multiplication Rule 1 [5]

**Example:** A die is tossed **3 times**. Find the probability of getting **three 6s**.

# Solution

Let **A** be the event of getting a '6'

$$P(A) = 1/6$$

Since events are independent, so

$$\begin{aligned} P(A \text{ and } A \text{ and } A) &= P(A) \times P(A) \times P(A) \\ &= 1/6 \times 1/6 \times 1/6 \\ &= 1/216 \quad (= 0.0046 \text{ or } 0.4600 \%) \end{aligned}$$

OR

$$\begin{aligned} P(AAA) &= 1/6 \times 1/6 \times 1/6 \\ &= 1/216 \\ &= 0.0046 \quad (\text{or } 0.4600 \%) \end{aligned}$$

# Multiplication Rule 1 [6]

**Example:** A box of transistors has **four good transistors** mixed up with **two bad transistors**. A production worker, in order to sample the product, chooses two transistors at random, the first chosen transistor being **replaced** before the second transistor is chosen. What is the probability that **both transistors** are good?

# Solution

If the events are

$A$  : the first transistor chosen is good

$B$  : the second transistor chosen is good

$$P(A) = \frac{4}{6}$$

$$P(B) = \frac{4}{6}$$

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

$$= \left(\frac{4}{6}\right) \left(\frac{4}{6}\right)$$

$$= \frac{4}{9}$$

$$= \mathbf{0.4444 \text{ (or 44.44\%)}}$$

# Multiplication Rule II [1]

- ❑ When two sequential events are **dependent**, a slight variation of the multiplication rule is used to find the probability of both events occurring.
- ❑ For Example, when a card is selected from an ordinary deck of **52 cards** the probability of getting a specific card is  $\frac{1}{52}$ , but the probability of getting a specific card on the second draw is  $\frac{1}{51}$  since 51 cards remain.

**Example:** Two cards are selected from a deck and the first card is **not replaced**. Find the probability of getting **two kings**.

## Solution

$$P(\text{two kings}) = \frac{4}{52} \times \frac{3}{51}$$

$$= \frac{12}{2652}$$

$$= \frac{1}{221}$$

$$= \mathbf{0.0045 \text{ (or 0.45 \% )}}$$

## Multiplication Rule II [2]

- When the two events A and B are **dependent**, the **probability that the second event B occurs after the first event A has already occurred** is written as  **$P(B|A)$** .
- This does not mean that B is divided by A; rather, it means and is read as **“the probability that event B occurs given that event A has already occurred.”**
- **$P(B|A)$**  also means the **conditional probability** that event B occurs given event A has occurred.



# Multiplication Rule II [3]

□ The probability of an event  **$B$**  occurring when it is known that some event  **$A$**  has occurred is called a **conditional probability** and is denoted by  **$P(B/A)$** .

□ The symbol  **$P(B/A)$**  is usually read “the probability that  **$B$  occurs** given that  **$A$  occurs**”

OR

□ simply “the probability of  $B$ , given  $A$ .”

# Multiplication Rule II [4]:

When two events are **dependent**, the probability of both events occurring is  **$P(A \text{ and } B) = P(A) \times P(B|A)$**

$$P(A) \times P(B)$$

As B is dependent so,  $\Rightarrow P(A) \times P(B|A)$

**Example:** A box contains **24 toasters**, **3** of which are **defective**. If **two toasters** are selected and tested, find the probability that **both are defective** (assume toasters are not replaced).

## Solution

Let  $D_1$  be the event that **first toaster is defective**.

Let  $D_2$  be the event that **second toaster is defective**.

$$P(D_1 \text{ and } D_2) = P(D_1) \times P(D_2 | D_1)$$

$$= \frac{3}{24} \times \frac{2}{23}$$

$$= \frac{1}{8} \times \frac{2}{23}$$

$$= \frac{1}{92}$$

$$= \mathbf{0.0109 \text{ (or 1.0870 \%)}}$$

# Multiplication Rule II [5]:

When two events are **dependent**, the probability of both events occurring is  **$P(A \text{ and } B) = P(A) \times P(B | A)$**

## Multiplication Rule II [6]:

**Example:** Two cards are drawn **without replacement** from a deck of 52 cards. Find the probability that **both are queens**.

# Solution

Let  $Q_1$  be the event that the **first card is a queen**.

Let  $Q_2$  be the event that the **second card is a queen**.

$$P(Q_1 \text{ and } Q_2) = P(Q_1) \times P(Q_2 | Q_1)$$

$$= \frac{4}{52} \times \frac{3}{51}$$

$$= \frac{1}{221}$$

$$= \mathbf{0.0045 \text{ (0.4525\%)}}$$

# Multiplication Rule II [7]:

**Example:** A box contains **3 orange balls**, **3 yellow balls**, and **2 white balls**. **Three balls** are selected **without replacement**. Find the probability of selecting **2 yellow balls** and a **white ball**.



# Solution

Orange balls	Yellow	White balls	Total balls
3	3	2	8

Let  $Y_1$  be the event that the **first ball is yellow**.

Let  $Y_2$  be the event that the **second ball is yellow**.

Let  $W_3$  be the event that the **third ball is white**.

$$P(Y_1 \text{ and } Y_2 \text{ and } W_3) \text{ or } P(Y_1 Y_2 W_3) = \frac{3}{8} \times \frac{2}{7} \times \frac{2}{6}$$

$$= \frac{12}{336}$$

$$= \mathbf{0.0357(\text{or } 3.5714 \%)}$$

**Note:** The key word for the multiplication rule is and. It means to multiply.

# Multiplication Rule II [1]:

**Example:** A box contains 3 orange balls, 3 yellow balls, and 2 white balls. Three balls are selected **without replacement**. Find the probability of selecting **a white ball** and **2 yellow balls**.

## Solution

Orange balls	Yellow balls	White balls	Total balls
3	3	2	8

Let  $W_1$  be the event that the **first ball is white**.

Let  $Y_2$  be the event that the **second ball is yellow**.

Let  $Y_3$  be the event that the **third ball is yellow**.

$$P(W_1 \text{ and } Y_2 \text{ and } Y_3) \text{ or } P(W_1 Y_2 Y_3) = \frac{2}{8} \times \frac{3}{7} \times \frac{2}{6}$$

$$= \frac{12}{336}$$

$$= \mathbf{0.0357 \text{ (or 3.5714 \% )}}$$

**Note:** The key word for the multiplication rule is **and**. It means to multiply.

# Conditional Probability [1]

- ❑ Previously, conditional probability was used to find the probability of sequential events occurring when they were **dependent**.
- ❑ Recall that  $P(B|A)$  means the probability of **event B** occurring given that **event A** has already occurred.
- ❑ Another situation where **conditional probability** can be used is when **additional information** about an event is known.
- ❑ Sometimes it might be known that **some outcomes** in the sample space have **occurred** or that some **outcomes cannot occur**.

# Conditional Probability [2]

When conditions are imposed or known on events, there is a possibility that the probability of the certain **event occurring may change.**

**Example:** A die is rolled; find the probability of getting a 4 if it is known that an even number occurred when the die was rolled.

# Alternative Approach: Conditional Probability [1]

**Solution:**

If it is known that an even number has occurred, the sample space is

**Reduced sample space** = {2, 4, 6}

$$n(S') = 3$$

Let **A** be the event of getting a '**4**'

$$A = \{4\}$$

$$n(A) = 1$$

$$P(A) = \frac{1}{3} = 0.3333 \text{ (33.33\%)}$$

# Sample space of two dice using 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)

# Alternative Approach: Conditional Probability [2]

**Example:** Two dice are rolled. Find the probability of getting a **sum of 3** **if it** is known that the sum of the spots on the dice was **less than six**.



## Solution

**Reduced sample space** = {(1, 1), (1, 2), (2, 1), (3, 1), (2, 2), (1, 3), (1, 4), (2, 3), (3, 2), and (4, 1)}

$$n(S') = 10$$

Let **A** be the event of getting a '**sum of 3**'

$$A = \{(1, 2), (2, 1)\}, n(A) = 2$$

$$P(A) = \frac{2}{10} = \frac{1}{5}$$

or

$$\begin{aligned} P(\text{sum of 3} | \text{sum less than 6}) &= \frac{2}{10} \\ &= \frac{1}{5} = 0.20 \text{ (or 20\%)} \end{aligned}$$

# Alternative Approach: Conditional Probability [3]

The two previous Examples of conditional probability were solved using **classical probability and reduced sample spaces**; however, they can be solved by using the following formula for conditional probability.

# Alternative Approach: Conditional Probability [4]

The conditional probability of two events A and B is

$$P(A|B) = P(A \text{ and } B)/P(B)$$

OR

$$= \frac{P(A \text{ and } B)}{P(B)}$$

This formula is generally used when 'if it is known' keyword is used.

**P(A and B)** means the probability of the outcomes that events **A and B have in common.**

# Conditional Probability without reducing the sample space [1]

**Example:** A die is rolled; find the probability of getting **a**  $A < 4$ , if it is known that an **even number** occurred when  $\rightarrow B$  the die was rolled.

# Solution

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

Let events are defined as:

**A:** Getting a **4** on a die

**B:** An **even number** occur on a die

$$\therefore P(A|B) = P(A \text{ and } B)/P(B)$$

$$P(A \text{ and } B) = 1/6$$

$$P(B) = 3/6$$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = 1/6 \times 6/3$$

$$= 1/3 = 0.3333(\text{or } 33.33\%)$$

# Sample space of two dice using 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)

# Conditional Probability without reducing the sample space [2]

**Example:** Two dice are rolled. Find the probability of getting a sum of 3 if it is known that the sum of the spots on the dice **was less than 6**.

## Solution [1]:

$$\therefore P(A|B) = P(A \cap B)/P(B)$$

**Let events are defined as:**

**$A \cap B$ :** Getting a sum 3 **and** sum of the spots on the dice was less than 6

**A:** Getting sum of the spots on the dice was 3

**B:** Getting sum of the spots on the dice was less than 6

$$A \cap B = \{(2, 1), (1, 2)\}$$

$$n(A \cap B) = 2$$

$$P(A \cap B) = \frac{2}{36} = \frac{1}{18}$$

$$= 0.0555 \text{ (or 5.55 \%)}$$



## Solution [2]:

Let **B** be the event of getting sum of the spots on the dice was **less than 6**

$$B = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (4, 1)\}$$

$$P(B) = \frac{10}{36} = \frac{5}{18} = \mathbf{0.2777 \text{ (or 27.78 \%)}}$$

$$\begin{aligned} P(A|B) &= P(A \text{ and } B)/P(B) \\ &= \frac{1}{18} \times \frac{18}{5} = \frac{1}{5} = \mathbf{0.2 \text{ (or 20 \%)}} \end{aligned}$$

# Alternative approach: Conditional Probability with reducing the sample space

**Example:** Two dice are rolled. Find the probability of getting a **sum of 3** if it is known that the sum of the spots on the dice **was less than 6**.

# Solution

If it is known that the sum of the spots on the dice **was less than 6**

**Let reduced sample space =  $S'$**

**$\Rightarrow S' = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (4, 1)\}$**

**$n(S') = 10$**

Let **A** be the event of getting a sum of 3

**$A = \{(2, 1), (1, 2)\}$**

**$P(A) = \frac{2}{10} = \frac{1}{5} = 0.2 \text{ (or 20 \%)}$**

# Alternative approach: Conditional Probability with reducing the sample space

**Example:** When two dice were rolled, it is known that the sum was an **even number**. In this case, find the probability that the **sum was 8**.

# Solution:

**Reduced sample space =  $S'$**

$\{(1, 1), (1, 3), (1, 5), (2, 2), (2, 4), (2, 6), (3, 1), (3, 3), (3, 5), (4, 2), (4, 4), (4, 6), (5, 1), (5, 3), (5, 5), (6, 2), (6, 4), (6, 6)\}$

**$n(S') = 18$**

Let **A** be the event of getting a sum of '**8**'

**A** =  $\{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\}$

$n(A) = 5$

$P(A) = \frac{5}{18} = \mathbf{0.2777 (27.78\%)}$

**Example:** In a large housing plan, 35% of the homes have a deck and a two-car garage, and 80% of the houses have a two-car garage. Find the probability that a house has a deck given that it has a two-car garage.

# Solution

Let **D** be the event of getting **deck**

Let **G** be the event of getting **two-car garage**

Given

$$P(D \cap G) = 0.35$$

$$P(G) = 0.80$$

$$\begin{aligned} P(\text{deck} | \text{two-car garage}) &= \frac{P(D \cap G)}{P(G)} \\ &= 0.35 / 0.80 = 7/16 \\ &= 0.4375 \text{ (or 43.75 \%)} \end{aligned}$$