



Daffodil International University

Department of Computer Science and Engineering (CSE)
Faculty of Science and Information Technology (FSIT)

Bayes' Theorem Mathematical Examples Lecture Sheet

Semester: Spring 2024

Course Code and Title: CSE315 – Introduction to Data Science

Course Teacher and Initial: Fahim Faisal (FF)

Section: 61_L

Question 1: Suppose you plan to go on a picnic somewhere with your family and friends on a Friday in January of 2024. When you wake up on that particular Friday morning, you see that the sky is cloudy and you become very worried. If it rains, all of your plans will get canceled! Being a smart, tech-savvy individual, you immediately search on Google for local weather data for your area for determining the probability of rain. The dataset shows the following,

The sky is cloudy in the morning 40% of the days in January.

It is cloudy in the morning half the time when it rains.

In January, it usually does not rain for more than 3 days.

Now, find out the probability based on the given data.

Solution 1:

From the given data we get,

The sky is cloudy in the morning 40% of the days in January. So, $P(C) = 0.4$

It is cloudy in the morning half the time, or 50% of the time, when it rains. So, $P(C | R) = 0.5$

In January, it usually does not rain for more than 3 days, i.e., probability of rain in January is 10%. So, $P(R) = 0.1$

From Bayes' theorem we know,

$$P(R | C) = \frac{P(R) \times P(C | R)}{P(C)}$$

$$\text{Or, } P(R | C) = \frac{0.1 \times 0.5}{0.4}$$

$$\text{Or, } P(R | C) = 0.125$$

So, if there is cloud in the sky in the morning in January, the probability of rain is 0.125 or 12.5%.

Question 2: Suppose you are a medical researcher working in the field of public health. You are interested in finding out a patient's probability of having liver disease if they consume alcohol regularly. For determining the probability, you start collecting data from hospitals, clinics and health research organizations. After a couple of days of rigorous data collection, you manage to compile a big enough database. In your database, a particular dataset from a renowned hospital suddenly draws your attention. The dataset shows that, among all of the hospital's patients, 10% have various liver diseases. Additionally, 5% of the patients identified

themselves as alcoholics and among the liver patients, 7% were tested to be alcoholics. Now, find out the probability based on the given data.

Solution 2:

From the given data we get,

10% of the patients in the hospital have liver disease. So, $P(A) = 0.10$

5% of the hospital's patients are alcoholics. So, $P(B) = 0.05$

The probability that a patient is alcoholic, given that they have liver disease is 7%.

So, $P(B | A) = 0.07$

From Bayes' theorem we know,

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

$$\text{Or, } P(A | B) = \frac{0.10 \times 0.07}{0.05}$$

$$\text{Or, } P(A | B) = 0.14$$

In other words, if a patient is an alcoholic, their chances of having liver disease is 0.14 or 14%.