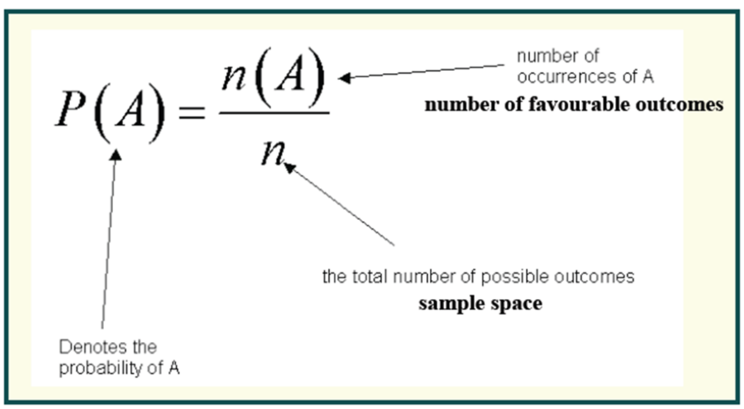
Basic Concept of Probability

For

First Stage

2019-2020



Probability: the measure of the chance an event will occur. Probability is a way of expressing knowledge or belief that an event will occur or has occurred.

Experimental Probability: probability based on an experiment written as a ratio comparing the number of times the event occurred to the number of trials.

$$\frac{\text{Number of times the event occured}}{\text{Number of trials}}$$

Theoretical Probability: probability based on reasoning written as a ratio of the number of favorable outcomes to the number of possible outcomes.

$$\frac{\text{Number of favorable outcomes}}{\text{Number of possible outcomes}}$$

In general:

Probability of an event happening $= \frac{Number\ of\ ways\ it\ can\ happen}{Total\ number\ of\ outcomes}$

Elementary probability concept:

Probability is a measure or estimation of likelihood of occurrence of an event. **Probabilities are given a value between 0 (0% chance or *will not happen*) and 1 (100% chance or *will happen*)**.

The higher the degree of probability, the more likely the event is to happen, or, in a longer series of samples, the greater the number of times such event is expected to happen.

Example

We will use common biomedical situation for instruction in the assessment and use of measures of probability.

Table 1: shows a frequency table for Serum cholesterol level in normal men 40-59 years old.

Serum cholesterol	***Frequency***	***R.F%***	***C.F%***
120-139.5	10	**1.0**	**1.0**
140-159.5	21	2.0	3.0
160-179.5	37	3.5	6.5
180-199.5	97	9.3	15.8
200-219.5	152	14.5	30.3
220-239.5	**206**	19.7	**50**
240-259.5	195	18.6	68.6
260-279.5	131	12.5	81.1
280-299.5	96	9.2	90.3
300-319.5	47	4.5	94.8
320-339.5	30	2.9	97.7
340-359.5	13	1.2	98.9
360-379.5	6	0.6	99.5
380-399.5	4	0.4	99.9
400-419.5	0	0	99.9
420-439.5	1	0.1	100.0
440-459.5	0	0	100.0
460-479.5	1	0.1	100.1
Total	1047	100.1	

- The distribution is symmetric, with the peak interval 220-239.5 mg/100 ml.
- From cumulative percent column we see that 50% of these normal men had cholesterol level measuring less than 240 mg/100 ml.

What would be the probability that his serum cholesterol measurement was in the range of 160-179?

Solution: from table 1 (37 of the 1047 have levels in this range).
The chance of selecting one of the 37 is:
(37/1047) × 100 = 3.5% as in R.F % column. This demonstrates the simple definition of probability.

If you were to select a normal male aged 40-59 at random from the general population from these 1047 drawn, **what is the probability that his cholesterol value would be less than 200**?
Solution:

$$P = \frac{10+21+37+97}{1047} = 15.8\% \quad \text{(or directly from } C.F \text{ column).}$$

Probability Rule One

The probability of an event, which informs us of the likelihood of it occurring, can range anywhere from 0 (indicating that the event will never occur) to 1 (indicating that the event is certain).

Probability Rule One:

- ***For any event A, 0 ≤ P (A) ≤ 1.***

Probability Rule Two:

- **The sum of the probabilities of all possible outcomes is 1**.

Note:-space that are not part of Event A and is denoted as A′.

P [A] + P [A′] = 1 or P [A] = 1 – P [A′]

Example: Blood Types

All human blood can be typed as O, A, B or AB. In addition, the frequency of the occurrence of these blood types varies by ethnic and racial groups, these are the probabilities of human blood types in the United States (the probability for type A has been omitted on purpose):

Blood type	O	A	B	AB
Probability	0.44	?	0.10	0.04

What is the probability of the person having blood type A?

Answer: Our intuition tells us that since the four blood types O, A, B, and AB exhaust all the possibilities, their probabilities together must sum to 1.Since the probabilities of O, B, and AB together sum to 0.44 + 0.1 + 0.04 = **0.58**, the probability of type A must be the remaining **0.42** (1 – 0.58 = 0.42):

Blood type	O	A	B	AB
Probability	0.44	0.42	0.10	0.04

Probability Rule Three (The Complement Rule):

- **P(not A) = A′= 1 – P(A)**
- That is, the probability that an event does not occur is 1 minus the probability that it does occur.

Example: Blood Types

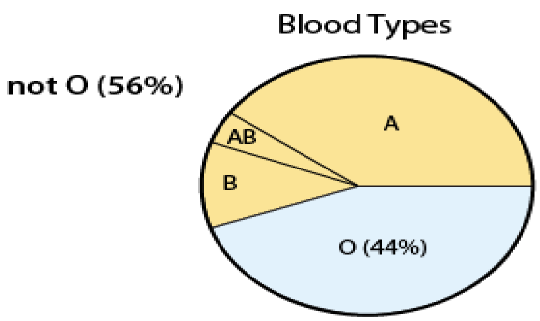
Blood type	O	A	B	AB
Probability	0.44	0.42	0.10	0.04

Here is some additional information:

- A person with type **A** can donate blood to a person with type **A** or **AB**.
- A person with type **B** can donate blood to a person with type **B** or **AB**.
- A person with type **AB** can donate blood to a person with type **AB** only.
- A person with type **O** blood can donate to anyone.

What is the probability that a randomly chosen person does not have blood type O?

We need to find P(not O). Using the Complement Rule, P(not O) = 1 – P(O) = 1 – 0.44 = 0.56. In other words, 56% of the population does not have blood type O:



Clearly, we could also find P (not O) directly by adding the probabilities of B, AB, and A.

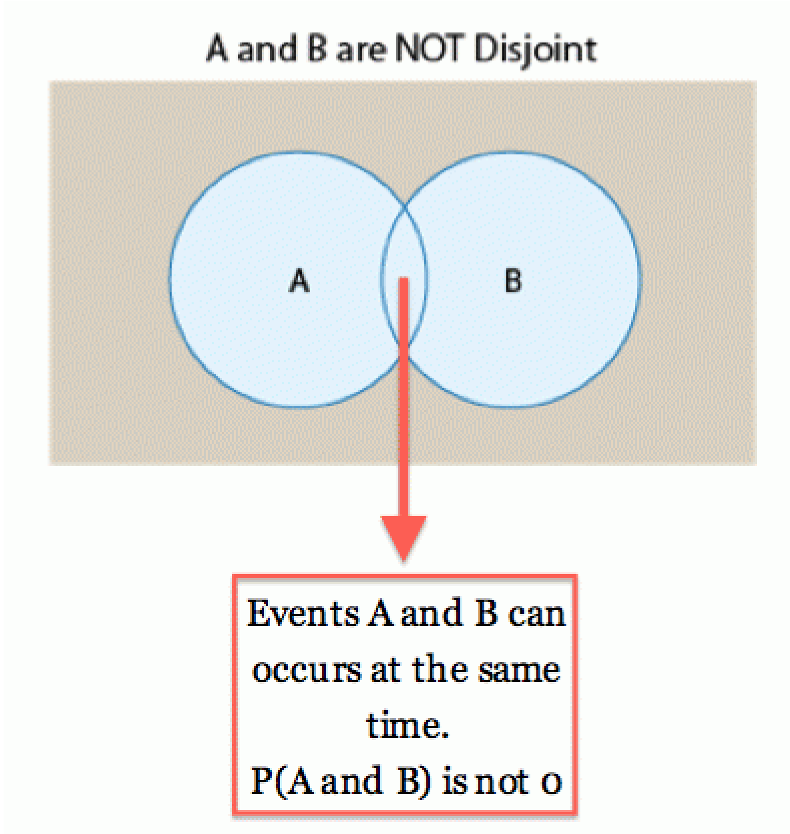
Probabilities Involving Multiple Events

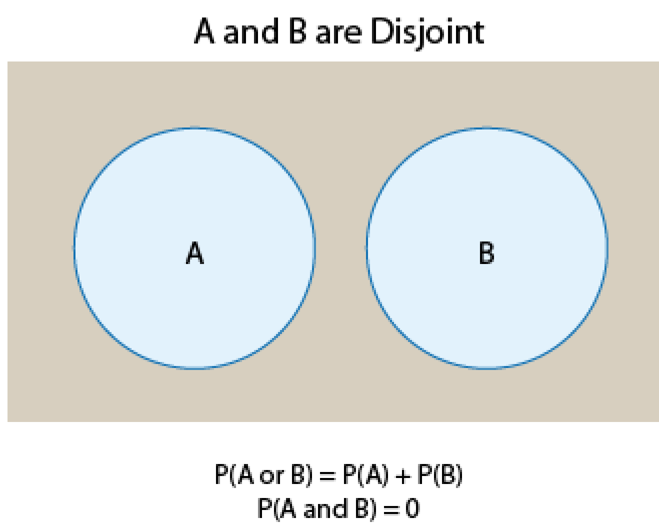
We will often be interested in finding probabilities involving multiple events such as

- **P(A or B) = P(event A occurs or event B occurs or both occur)**
- **P(A and B)= P(both event A occurs and event B occurs)**

The distinction between events that can happen together and those that cannot is an important one.

Mutually Exclusive: Two events that cannot occur at the same time are called disjoint or mutually exclusive. (We will use disjoint.)





It should be clear from the picture that

- in the first case, where the events are **NOT mutually exclusive events, P(A and B) $\neq$ 0**
- In the second case, where the events **mutually exclusive events, P (A and B) = 0.**

<u>Addition rule of probability</u>

If A and B are mutually exclusive events:
The probability that either Event A or Event B will occur.
P (A or B) = P(A) + P(B)

If A and B are **Not mutually exclusive events*:***
Events that can occur at the same time.
P[A or B] = P[A] + P[B] – P[A and B].

Example:

Consider the following two events:

A — a randomly chosen person has blood type A, and

B — a randomly chosen person has blood type B.

We are going to assume that each person can have only one blood type. Therefore, it is impossible for the events A and B to occur together.

- **Events A and B are DISJOINT**

Example:

Consider the following two events:

A — a randomly chosen person has blood type A

B — a randomly chosen person is a woman.

In this case, it **is possible** for events A and B to occur together.

- **Events A and B are NOT DISJOINT.**

Probability Rule Four (The Addition Rule for Disjoint Events):

- **If A and B are disjoint events, then P (A or B) = P (A) + P (B).**

Example: Blood Types

Blood type	O	A	B	AB
Probability	0.44	0.42	0.10	0.04

Here is some additional information

- A person with type **A** can donate blood to a person with type **A** or **AB**.
- A person with type **B** can donate blood to a person with type **B** or **AB**.
- A person with type **AB** can donate blood to a person with type **AB**
- A person with type **O** blood can donate to anyone.

What is the probability that a randomly chosen person is a potential donor for a person with blood type A?

From the information given, we know that being a potential donor for a person with blood type A means having blood type A or O. We therefore need to find P(A or O). Since the events A and O are disjoint, we can use the addition rule for disjoint events to get:

- **P(A or O) = P(A) + P(O) = 0.42 + 0.44 = 0.86.**

If 42% of the population has blood type A and 44% of the population has blood type O, then 42% + 44% = 86% of the population has either blood type A or O, and thus are potential donors to a person with blood type A.

Example: Periodontal Status and Gender

Consider randomly selecting one individual from those represented in the following table regarding the periodontal status of individuals and their gender. Periodontal status refers to gum disease where individuals are classified as either healthy, have gingivitis, or have periodontal disease.

Count

		periodontal status			
		healthy	gingivitis	perio	Total
GENDER	male	1143	929	937	3009
	female	2607	1490	921	5018
Total		3750	2419	1858	8027

Let's review what we have learned so far. We can calculate any probability in this scenario if we can determine how many individuals satisfy the event or combination of events.

- P(Male) = 3009/8027 = 0.3749
- P(Female) = 5018/8027 = 0.6251
- P(Healthy) = 3750/8027 = 0.4672
- P(Not Healthy) = P(Gingivitis or Perio) = (2419 + 1858)/8027 = 4277/8027 = 0.5328

  We could also, calculate this using the complement rule: 1 – P(Healthy)

  We also previously found that
- P(Male AND Healthy) = 1143/8027 = 0.1424

Count

		healthy	gingivitis	perio	Total
		periodontal status			
GENDER	male	1143	929	937	3009
	female	2607	1490	921	5018
Total		3750	2419	1858	8027

- P(Male or Healthy) = P(Male) + P(Healthy) – P(Male and Healthy)
  = 0.3749 + 0.4672 – 0.1424 = 0.6997 or about 70%

  We solved this question earlier by simply counting how many individuals are either Male or Healthy or both. The picture below illustrates the values we need to combine. We need to count
  - All males
  - All healthy individuals
  - BUT, not count anyone twice!!

Count

		healthy	gingivitis	perio	Total
		periodontal status			
GENDER	male	1143	929	937	3009
	female	2607	1490	921	5018
Total		3750	2419	1858	8027

- P(Male or Healthy) = (1143 + 929 + 937 + 2607)/8027 = 5616/8027 = 0.6996

Independent Events:

- Two events A and B are said to be **independent** if the fact that one event has occurred **does not affect** the probability that the other event will occur.
- If whether or not one event occurs **does affect** the probability that the other event will occur, then the two events are said to be **dependent.**

(The Multiplication Rule for Independent Events):

- **If A and B are two Independent events, then**
  P (A and B) = P(A) * P(B).

Example:

Blood type	O	A	B	AB
Probability	0.44	0.42	0.10	0.04

Two people are selected simultaneously and at random from all people **What is the probability that both have blood type O?**

- Let O1= "person 1 has blood type O" and
- O2= "person 2 has blood type O"

We need to **find P(O1 and O2)** Since they were chosen simultaneously and at random, the blood type of one has no effect on the blood type of the other. Therefore, O1 and O2 are independent, and we may apply Rule:

- **P(O1 and O2) = P(O1) * P(O2) = 0.44 * 0.44 = 0.1936.**

Conditional Probability

- **The conditional probability of event B, given event A, is P(B | A) = P(A and B) / P(A)**

EXAMPLE:

On the "Information for the Patient" label of a certain antidepressant, it is claimed that based on some clinical trials,

- there is a 14% chance of experiencing sleeping problems known as insomnia (**denote this event by I**),
- there is a 26% chance of experiencing headache (**denote this event by H**),
- and there is a 5% chance of experiencing both side effects (**I and H**).

(a) Suppose that the patient experiences insomnia; what is the probability that the patient will also experience headache?

Since we know (or it is **given**) that the patient experienced **insomnia**, we are looking for **P(H | I)**.

According to the definition of conditional probability:

P(H | I) = P(H and I) / P(I) = 0.05/0.14 = 0.357.

(b) Suppose the drug induces headache in a patient; what is the probability that it also induces insomnia?

Here, we are given that the patient experienced headache, so we are looking for P(I | H).

Using the definition

P(I | H) = P(I and H) / P(H) = 0.05/0.26 = 0.1923.

Bayes Rule

This rule is used to find the probability of one state or condition given that some other state or condition has been observed. The formula is:

$$\mathbf{P(A|B)} = \frac{\boldsymbol{P(A)\,P(B|A)}}{\boldsymbol{P(B)}}$$

Which tells us how often A happens *given that B happens*, written **P(A|B)**, when we know how often B happens *given that A happens*, written **P(B|A)** and how likely A is on its own, written **P(A)**
and how likely B is on its own, written **P(B)**

Example: -

Suppose that the rate of certain disease in given population is 5 percent. Suppose also that 80 percent of those with disease show certain symptoms while 10 percent of those without disease also show the same symptoms. For a person selected at random from the population, what is the Prob. that the disease is present if the symptoms are observed?

Ans.:-

We need to evaluate P(disease | symptoms).

Let D be the event that the disease is present.

Let D′ be the event that the disease is absent.

Let S be the event that the symptoms are present.

Let S′ be the event that the symptoms are absent.

P(D)=0.05 , P(D′)= 1- 0.05 = 0.95

P(S|D) = 0.8 ,P(S′|D) =1- 0.8 = 0.2

P(S| D′)=0.1 , P(S′|D′)=1- 0.1 = 0.9

$$P(D|S) = \frac{P(D)\,P(S|D)}{P(S)}$$

From the data given in the problem we know the value $P(S|D)$ and $P(D)$. what is the value the value of P(S) ? clearly , anyone with the symptoms either has the disease or dose not have the disease . Each person can be classified to

“ S and D “ or “ S and D′)

Are mutually exclusive

P(S) = P(S and D) + P(S
P(S and D) =P(S|D)P(D)
P(S and D′)=P(S| D′)P(D′)

$$P(D|S) = \frac{P(D)P(S|D)}{P(S|D)P(D) + P(S|D')P(D')}$$

$$P(D|S) = \frac{0.05 \times 0.8}{0.8 \times 0.05 + 0.1 \times 0.95}$$

$$= 0.30$$

Example //

If tuberculous meningitis had a case fatality of 20%, (a) Find the probability that this disease would be fatal in two randomly selected patients (the two events are <u>independent)</u> (b) If two patients are selected randomly what is the probability that at least one of them will die?

(a) P(first die and second die) = 20% × 20% = 0.04

(b) P(first die or second die)

= P(first die) + P(second die) - P(both die)

= 20% + 20% - 4%

= 36%

Example

Approximately 1% of women age 40 – 50 have breast cancer. A women with breast cancer has a 90% chance of positive test from a mammograms , while a women without has 10% chance of false positive result what is the probability a women has breast cancer given that she just had positive test .

Let cancer = the women has breast cancer

Let + = a positive test

P(cancer |+) = ???

$$P(cancer|+) = \frac{P(cancer)P(+|cancer)}{P(+)}$$

$$P(cancer|+) = \frac{P(cancer)P(+|cancer)}{P(+|cancer)P(cancer) + P(+|cancer')P(cancer')}$$

P(cancer)=0.01 , P(cancer′)= 1- 0.01 = 0.99

P(+|cancer) = 0.9 ,P(+′|cancer) =1- 0.9 = 0.1

P(+| cancer′)=0.1 , P(+′|cancer′)=1- 0.1 = 0.9

$$P(cancer|+) = \frac{0.01 \times 0.9}{0.9 \times 0.01 + 0.1 \times 0.99} = \frac{9}{108} = 8.3\%$$