Chapter 03

Probability

Biostatistics For the Health Sciences

Dr. Moustafa Omar Ahmed Abu-Shawiesh
Professor of Statistics

3.1 Introduction

In this chapter, probability is defined and some rules for working with probabilities are introduced. Understanding probability is essential in calculating and interpreting p-values in the statistical tests of subsequent chapters. It also permits the discussion of sensitivity, specificity, and predictive values of screening tests.

3.2 Definition of Probability

Question: What is Probability?

Answer

Probability, is a mathematical language or framework that allows us to describe and analyze random phenomena (الظواهر العشوائية) in the world around us as well as in every discipline in science, engineering, technology, medicine and health sciences.

Question: What is a random phenomena?

Answer

By a random phenomena, we mean events or experiments whose outcomes (*results*) (النواتج) we can't predict with certainty.

Probability in Health Sciences

The concept of probability is not foreign to health sciences workers and is frequently encountered in everyday communication.

Example

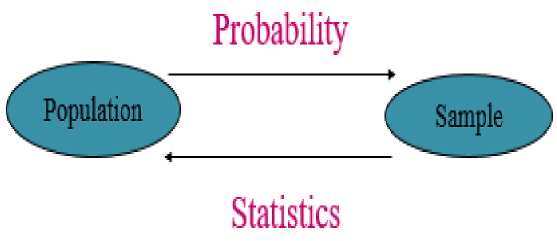
- A physician say that a patient has a 50–50 chance of surviving a certain operation.
- A physician may say that she is 95 percent certain that a patient has a particular disease.
- A public health nurse may say that nine times out of ten a certain client will break an appointment.

Notation (1): As these examples suggest, most people express probabilities in terms of percentages.

Notation (2): In dealing with probabilities mathematically, it is more convenient to express probabilities as fractions. Thus:

- We measure the probability of the occurrence of some event by a number between zero and one.
- The more likely the event, the closer the number is to one.
- The more unlikely the event, the closer the number is to zero.
- An event that cannot occur has a probability of zero.
- An event that is certain to occur has a probability of one.

Notation (3): Probability provides a bridge between descriptive statistics and inferential statistics. That is, the theory of probability provides the foundation for statistical inference.



<u>Definition</u>: **Experiment**

Any activity that yields a result or an outcome is called an experiment. The experiment results in something. The possible results (outcomes) of an experiment may be one or more. Based on the number of possible results in an experiment, we classify the experiments into two types as follows:

- Deterministic (Non-random).
- Random (Nondeterministic).

<u>Examples</u>

Here are some examples of experiments:

- The rolling of a die.
- The tossing of a coin.
- PCR test for COVID-19.
- The selection of a numbered ball (1 - 50) in an urn.
- The throwing of a stone to up.

Definition: Deterministic (Non-Random) Experiment

It is the experiment which have only one possible result or outcome, that is, whose result is certain or unique. The result of this type of experiments is predictable with certainty and is known prior to its conduct (determine the outcome with 100% certainty).

Notation: Probability theory does not based on the paradigm of a non-random (deterministic) experiments.

Example

Here are some examples of deterministic (non-random) experiments:

(1) Combining Hydrogen and Oxygen to make water (H_2O).

(2) The rises of sun tomorrow.

(3) The throwing of a stone to up.

Definition: Random (Non-Deterministic) Experiment

A Random Experiment is an experiment for which we know the set of all possible results (outcomes) for it before it performed but we can't predict which one of the results will occur until it performed.

Notations

(1) Probability theory is based on the paradigm of random experiments.

(2) An outcome is a result of a random experiment.

Example

Here are some examples of random experiments (non-deterministic):

(1) Tossing of a coin.

(2) Rolling a die.

(3) PCR test for COVID-19.

(4) Testing the blood group for a patient in the hospital.

The Sample Space (S)

Definition: Sample Space (S)

A **sample space** is the set of all possible outcomes for a random experiment and is denoted by **S** or Ω. The outcomes are mutually exclusive (disjoint) in the sense that they cannot occur simultaneously.

An Event

Definition : Event

An **event** is a subset of the sample space (S) to which we assign **a probability** and is denoted by capital letters $A, B, C, D, E, F, \ldots$. We say that an event occurs if and only if the outcome of the random experiment is an element of the event.

Types of Events

There are four types of events as follows:

1- Simple Event.

2- Compound Event.

3- Null (Impossible) Event (ϕ): $\phi = \{ \}$ = the event that contains no outcomes.

4- Sure or certain Event (S): the event that contains all outcomes.

Notation: The empty event, ϕ , never occurs while the sure event, S, always occurs.

Important Rules : Counting the Number of Outcomes (n(S)) for a Sample Space (S):

Rule (1): If the number of outcomes for a random experiment is n and the experiment was repeated r times, then the number of outcomes in the sample space (S) is denoted by n(S) and can be calculated by:

$$n(S) = n^r$$

Rule (2): If we have more than one different random experiments combined together, then we calculate $n(S) = n^r$ for each one of experiments and multiple the results for all experiments as follows:

$$n(S) = n_1^{r_1} \times n_2^{r_2} \times \dots \times n_m^{r_m}$$

where m is the number of different random experiments combined together.

Examples of a Sample Space (S)

<u>Example</u>: Coin Tossing

Determine the sample space (S) for each one of the following random experiments:

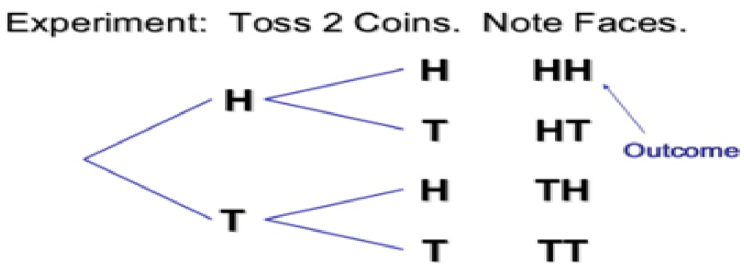
(1) Suppose that we toss a coin one time and observe the result (Head (**H**) or Tail (**T**)) comes up as shown in the figure below:

Number of Outcomes in the Sample Space: $n(S) = 2^1 = 2.$

Sample Space: $S = \{Head, Tail\} = \{H, T\}.$ (Finite & Discrete)

(2) Suppose that we toss a coin two times (two coins once) and observe the result (Head (**H**) or Tail (**T**)) comes up as shown in the figure below:

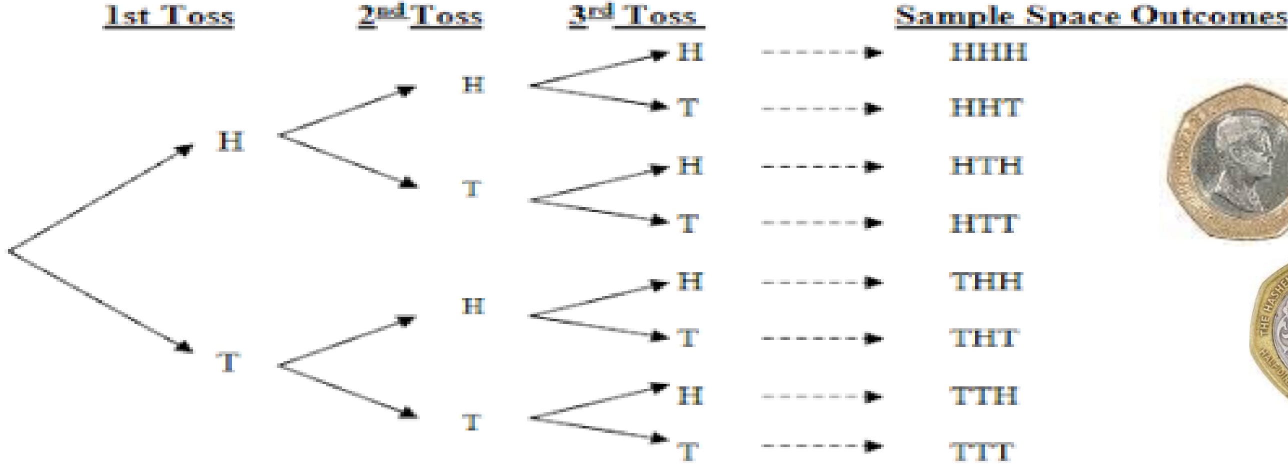
Tree Diagram



Number of Outcomes in the Sample Space: $n(S) = 2^2 = 2 \times 2 = 4.$

Sample Space: $S = \{HH, HT, TH, TT\}.$ (Finite & Discrete)

(3) Suppose that we toss a coin three times (three coins once) and observe the result (Head (**H**) or Tail (**T**)) comes up as shown in the figure below:



Number of Outcomes in the Sample Space: $n(S) = 2^3 = 2 \times 2 \times 2 = 8$.

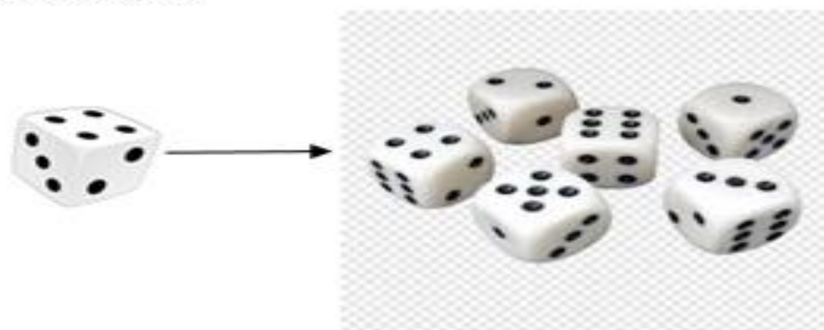
Sample Space: $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.

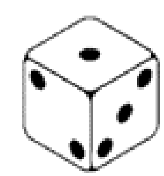
(Finite & Discrete)

<u>Example</u>: Rolling Dice

Determine the sample space (S) for each one of the following random experiments:

(1) Suppose that we roll a die one time and observe the number of dots comes up as shown in the figure below:





Number of Outcomes in the Sample Space: $n(S) = 6^1 = 6$.

Sample Space: $S = \{1, 2, 3, 4, 5, 6\}$. (Finite & Discrete)

(2) Suppose that we roll two dice once (one die twice) and observe the number of dots comes up as shown in the figure below:

Number of Outcomes in the Sample Space: $n(S) = 6^2 = 6 \times 6 = 36$.

Sample Space: $$S = \begin{Bmatrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{Bmatrix}$$

Question: Determine the outcomes and type of the following events defined on the Sample Space (S):

a) The event A is that the sum of the two numbers comes up is exactly 2?
   Solution: $A = \{(1,1)\}$. (elementary or simple event)

b) The event B is that the sum of the two numbers comes up is at least 10?
   Solution: $B = \{(4,6), (5,5), (5,6), (6,4), (6,5), (6,6)\}$. (compound event)

c) The event C is that the sum of the two numbers comes up is exactly 13?

   Solution: $C = \{\,\} = \emptyset$. (impossible event)

d) The event D is that the sum of the two numbers comes up is at least 2?

   Solution: $$D = \begin{Bmatrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{Bmatrix} = S$$

(certain or sure event)

Example

A hospital administrator codes incoming patients according to whether they have health insurance or not as follows:

$$\text{Code} = \begin{cases} 0 \quad , & \text{if the patient don't have health insurance} \\ 1 \quad , & \text{if the patient do have health insurance} \end{cases}$$

and their condition rated as follows:

$$\text{Patient Condition} = \begin{cases} \text{Good (G)} \\ \text{Fair (F)} \\ \text{Serious (R)} \end{cases}$$

if the random experiment is to writ the coding for the patients, then then answer the following:

a) What is the sample space (S) for this random experiment?

Solution: The sample space (S) is discrete and finite consists of all outcomes given as follows:

$$S = \{(0, G), (0, F), (0, R), (1, G), (1, F), (1, R)\} \quad , \quad n(S) = 2 \times 3 = 6$$

b) If the event A is that <u>the patient is in serious condition</u>, then determine the outcomes of A?

Solution: $A = \{(0, R), (1, R)\}$. (compound event)

c) If the event B is that <u>the patient is uninsured</u>, then determine the outcomes of B?

Solution: $B = \{(0, G), (0, F), (0, R)\}$. (compound event)

d) If the event C is that <u>the patient is uninsured and is in serious condition</u>, then determine the outcomes of C?

Solution: $C = \{(0, R)\}$. (elementary or simple event)

DEFINITION 3.1

The probability of an event is the relative frequency of this set of outcomes over an indefinitely large (or infinite) number of trials.

Definition: Definition of Probability

Let S be the finite sample space for a random experiment with equally likely outcomes, then for each event A, we have:

$$P(A) = \frac{n(A)}{n(S)}$$

where:

a) n(A) is the number of outcomes in the event A.

b) n(S) is the total number of outcomes in the sample space (S).

Notation (Important): The probability of an event A, denoted by P(A), always satisfies: $0 \leq P(A) \leq 1$

Example

Let us assume that there were 3,000,000,skydiving jumps in a recent year and 21 of them resulted in deaths. What is the probability of dying when making a skydiving jump?

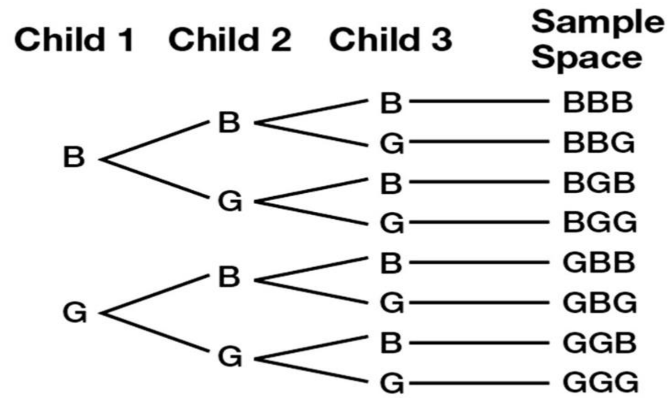
Solution

Let A = Skydiving Death, then we need to find P(A) as follows:

$$P(A) = \frac{\text{Number of Skydiving Deaths}}{\text{Number of Skydiving Jumps}} = \frac{n(A)}{n(S)} = \frac{21}{3{,}000{,}000} = 0.000007$$

Example

When three children are born, if we refer to the boy by B and to the girl by G, then for this random experiment answer the following:



a) Write the outcomes of the sample space (S)?

Answer

$n(S) = 2^3 = 8$ outcomes

S = {BBB, BBG, BGB, BGG, GBB, GBG, GGB, GGG}

b) Find the probability of getting 2 girls and 1 boy?

Answer

Let E = Getting 2 girls and 1 boy = {GGB, BGG, GBG}

$$P(E) = \frac{n(E)}{n(S)} = \frac{3}{8} = 0.375$$

Answer

c) Find the probability of getting 2 girls followed by 1 boy?

Let D = Getting 2 girls followed by 1 boy = {GGB}

$$P(D) = \frac{n(D)}{n(S)} = \frac{1}{8} = 0.125$$

d) Find the probability of getting 3 children all of the same sex?

Answer

Let F = Getting 3 children all of the same sex = {BBB, GGG}

$$P(F) = \frac{n(F)}{n(S)} = \frac{2}{8} = 0.25$$

DEFINITION 3.2 Two events A and B are **mutually exclusive** if they cannot both happen at the same time.

EQUATION 3.1 If outcomes A and B are two events that cannot both happen at the same time, then:

$$P(A \text{ or } B \text{ occurs}) = P(A) + P(B).$$

EXAMPLE 3.6

Hypertension Let A be the event that a person has normotensive diastolic blood-pressure (DBP) readings ($DBP < 90$), and let B be the event that a person has borderline DBP readings ($90 \le DBP < 95$). Suppose that $Pr(A) = .7$, and $Pr(B) = .1$. Let Z be the event that a person has a $DBP < 95$. Then

$$Pr(Z) = Pr(A) + Pr(B) = .8$$

because the events A and B cannot occur at the same time.

Thus the events A and B in Example 3.6 are mutually exclusive.

EXAMPLE 3.7 **Hypertension** Let X be DBP, C be the event $X \ge 90$, and D be the event $75 \le X \le 100$. Events C and D are *not* mutually exclusive, because they both occur when $90 \le X \le 100$.

3.3 Some Useful Probabilistic Notation

DEFINITION 3.3 The symbol **{ }** is used as shorthand for the phrase "the event."

DEFINITION 3.4 **Union or A ∪ B** is the event that either A or B occurs, or they both occur.

Example Hypertension

Let events A and B be defined as: A = { X < 90 }, B = { 90 ≤ X < 95 }, where X = DBP. Then A ∪ B = { X < 95 }.

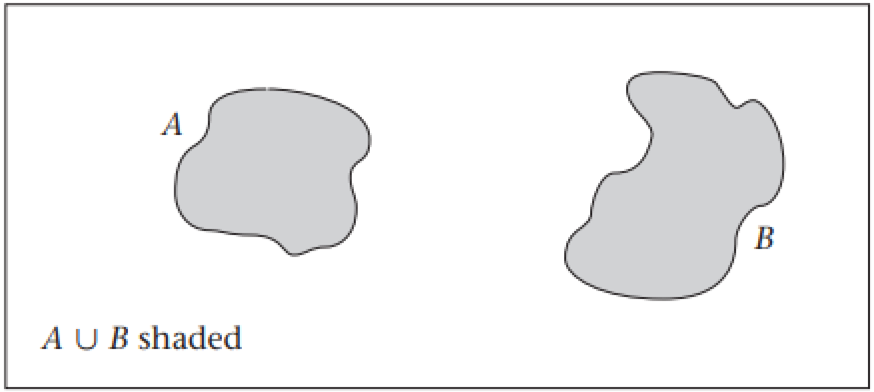
Example Hypertension

Let events C and D be defined as:

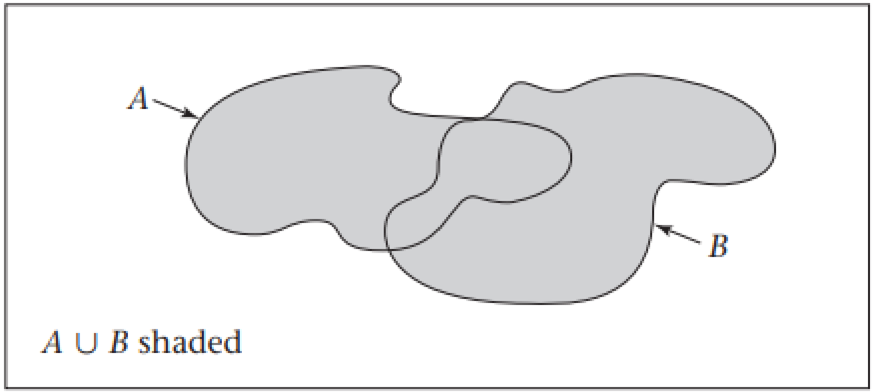
C = { X ≥ 90 } and D = { 75 ≤ X ≤ 100 }. Then C ∪ D = { X ≥ 75 }.

Figure 3.1 Diagrammatic representation of A ∪ B:

(a) A , B Mutually Exclusive

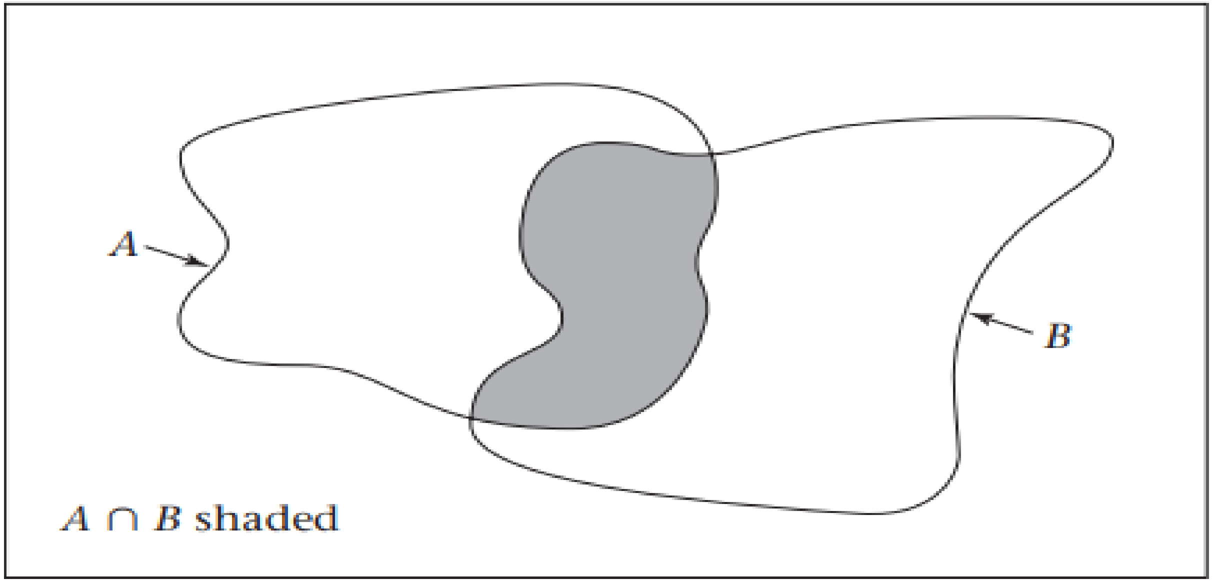


(b) A , B Not Mutually Exclusive



DEFINITION 3.5 **Intersection or A ∩ B** is the event that both A and B occur simultaneously (common outcomes between A and B). A ∩ B is shown in Figure 3.2.

Figure 3.2 Diagrammatic representation of A ∩ B:



Example Hypertension

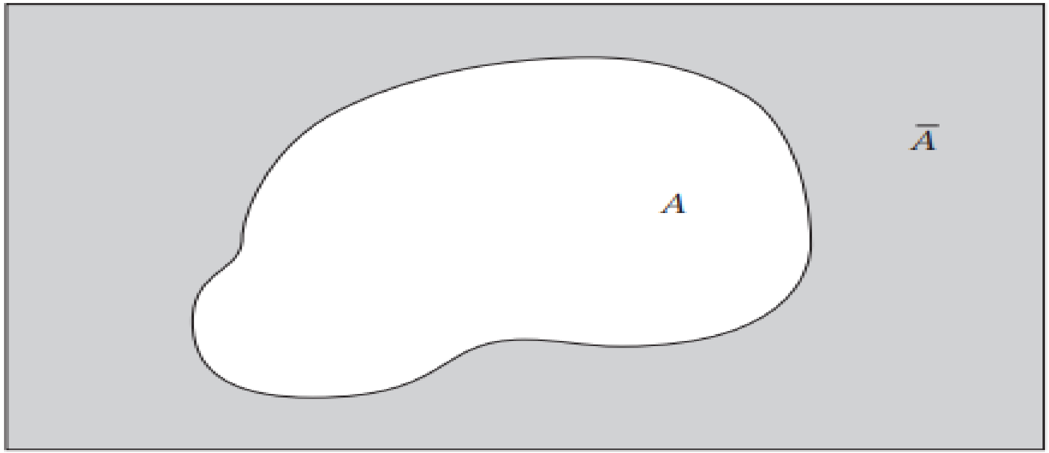
Let events C and D be defined as:

C = { X ≥ 90 } and D = { 75 ≤ X ≤ 100 }.

Then C ∩ D = { 90 ≤ X ≤ 100 }.

DEFINITION 3.6 **Complement of A or $\bar{A}$** is the event that A does not occur (Not A). It is called the complement of A. Notice that $P(\bar{A}) = 1 - P(A)$, because $\bar{A}$ occurs only when A does not occur. Event $\bar{A}$ is diagrammed in Figure 3.3.

Figure 3.3 Diagrammatic representation of $\bar{A}$



Example Hypertension Let events A and C be defined as follows:

$A = \{ X < 90 \}$ and $C = \{ X \geq 90 \}$

Then $C = \bar{A}$, because C can only occur when A does not occur.

Notice that

$P(\bar{A}) = 1 - P(A)$ and thus if 70% of people have $DBP < 90$, then 30% of people must have $DBP \geq 90$ because $P(\bar{A}) = 1 - P(A) = 1 - 0.7 = 0.3$.

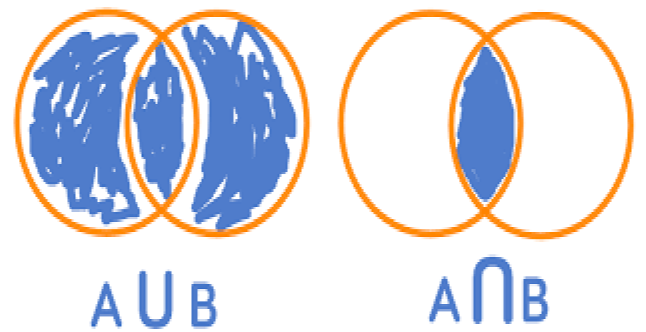
Example

Given the following information:

$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$

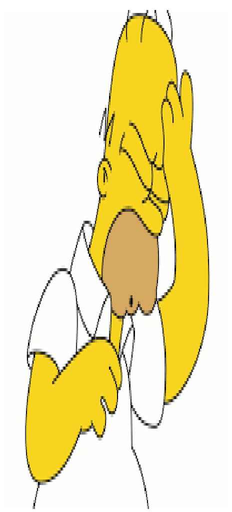
$A = \{1, 2, 5, 9, 13\}$

$B = \{2, 4, 6, 9\}$



Then:

$$A \cap B = \{1, 2, 5, 9, 13\} \cap \{2, 4, 6, 9\}$$
$$= \{2, 9\}$$



$$A \cup B = \{1, 2, 5, 9, 13\} \cup \{2, 4, 6, 9\}$$
$$= \{1, 2, 4, 5, 6, 9, 13\}.$$

$$= \{3, 4, 6, 7, 8, 10, 11, 12, 14, 15\}.$$

3.4 The Multiplication Law of Probability

In this section, certain specific types of events are discussed.

DEFINITION 3.7 Two events A and B are called independent events if

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

DEFINITION 3.8 Two events A, B are dependent if

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

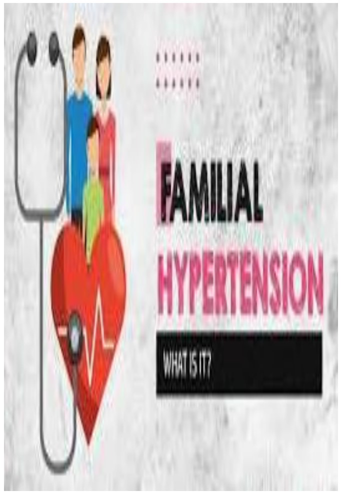
Example Hypertension, Genetics

Suppose we are conducting a hypertension-screening program in the home. In particular, we might be interested in whether the mother or father is hypertensive, which is described, respectively, by events:

A = {mother's DBP ≥ 90}

B = {father's DBP ≥ 90}.

A ∩ B = {mother's DBP ≥ 90 and father's DBP ≥ 90}.



Suppose we know that $P(A) = 0.1$, $P(B) = 0.2$ and $P(A \cap B) = 0.02$. Are the two events A and B independent?

Solution

$P(A \cap B) = 0.02$ AND $P(A) \times P(B) = 0.1 \times 0.2 = 0.02$

Then $P(A \cap B) = 0.02 = P(A) \times P(B)$

Therefore A and B are independent events. The hypertensive status of the mother does not depend at all on the hypertensive status of the father.

Example Hypertension, Genetics

Suppose we are conducting a hypertension-screening program in the home. In particular, we might be interested in whether the mother and mother first-born child's is hypertensive, which is described, respectively, by events:

A = {mother's DBP ≥ 90}

B = {mother first-born child's DBP ≥ 80}.

A ∩ B = {mother's DBP ≥ 90 and mother first-born child's DBP ≥ 80}.

Suppose we know that P(A) = 0.1 , P(B) = 0.2 and P(A ∩ B) = 0.05. Are the two events A and B independent?

Solution

$P(A \cap B) = 0.05$ AND $P(A) \times P(B) = 0.1 \times 0.2 = 0.02$

Then $P(A \cap B) = 0.02 \neq 0.05 = P(A) \times P(B)$

Conclusion: Therefore A and B dependent events. This outcome would be expected because the mother and her first-born child both share the same environment and are genetically related.

Notation: If two events are not independent, then they are said to be dependent.

*** Example Sexually Transmitted Disease Suppose two doctors, A and B, test all patients coming into a clinic for syphilis. Let us define the following two events:

A^+={doctor A makes a positive diagnosis} ; B^+={doctor B makes a positive diagnosis}

Suppose doctor A diagnoses 10% of all patients as positive, doctor B diagnoses 17% of all patients as positive, and both doctors diagnose 8% of all patients as positive. Are the two events A^+, B^+ independent?

Solution

We are given that $P(A^+) = 0.1$, $P(B^+) = 0.17$ and $P(A^+ \cap B^+) = 0.08$

$P(A^+ \cap B^+) = 0.08$ AND $P(A^+) \times P(B^+) = 0.1 \times 0.17 = 0.017$

Then $P(A^+ \cap B^+) = 0.08 \neq 0.017 = P(A^+) \times P(B^+)$

Conclusion: The two events are dependent. This result would be expected because there should be a similarity between how two doctors diagnose patients for syphilis.

Now Definition 3.7 can be generalized to the case of k (> 2) independent events. This is often called the multiplication law of probability given as follows:

EQUATION 3.2 Multiplication Law of Probability

If $A_1, A_2, \ldots, A_k$ are mutually independent events, then

$$P(A_1 \cap A_2 \cap \ldots \cap A_k) = P(A_1) \times P(A_2) \times \ldots \times P(A_k)$$

3.5 T he Addition Law of Probability

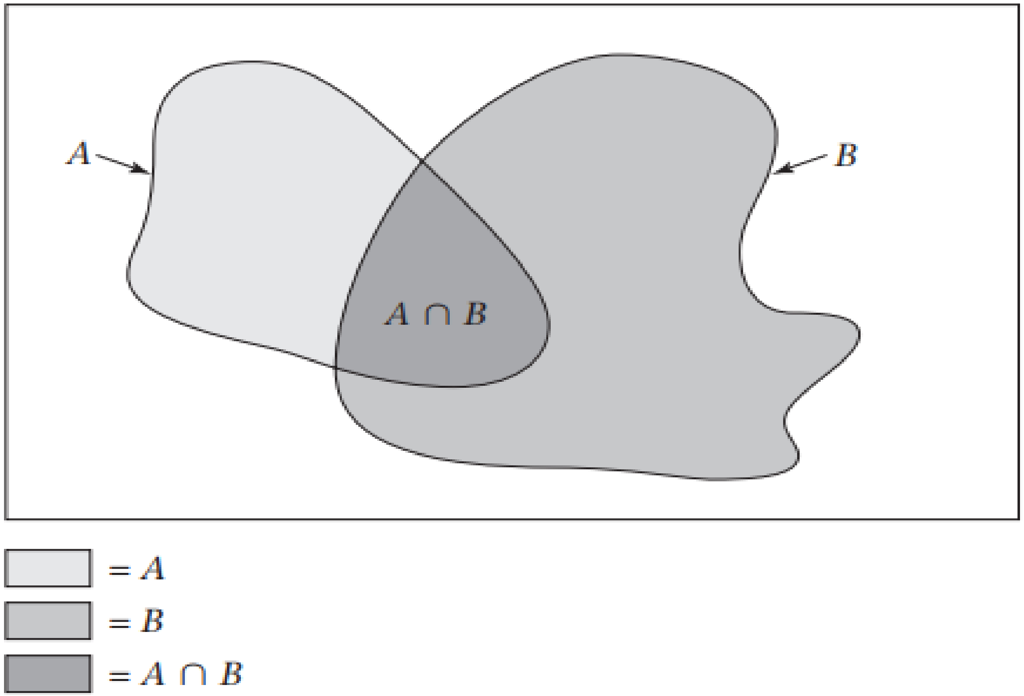
EQUATION 3.3 Addition Law of Probability

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

Notice that if A and B are mutually exclusive events, that is $A \cap B = \phi$, then $P(A \cap B) = P(\phi) = 0$ and therefore $P(A \cup B) = P(A) + P(B)$.

The Addition Law of Probability principle is diagrammed in Figure 3.5.

FIGURE 3.5 Diagrammatic representation of the addition law of probability



Example **Sexually Transmitted Disease** Suppose two doctors, A and B, test all patients coming into a clinic for syphilis. Let us define the following two events:

A^+={doctor A makes a positive diagnosis} ; B^+={doctor B makes a positive diagnosis}

A diagnoses 10% of all patients as positive, doctor B diagnoses 17% of all patients as positive, and both doctors diagnose 8% of all patients as positive, that is, we are given the following probabilities:

$$P(A^+) = 0.10 \,,\; P(B^+) = 0.17 \text{ and } P(A^+ \cap B^+) = 0.08$$

Suppose a patient is referred for further lab tests if either doctor A or B makes a positive diagnosis. What is the probability that a patient will be referred for further lab tests?

Solution

The event that either doctor makes a positive diagnosis can be represented by $(A^+ \cup B^+)$. Therefore, from the addition law of probability:

$$P(A^+ \cup B^+) = P(A^+) + P(B^+) - P(A^+ \cap B^+)$$
$$P(A^+ \cup B^+) = 0.10 + 0.17 - 0.08$$
$$= 0.19$$

Thus, 19% of all patients will be referred for further lab tests.

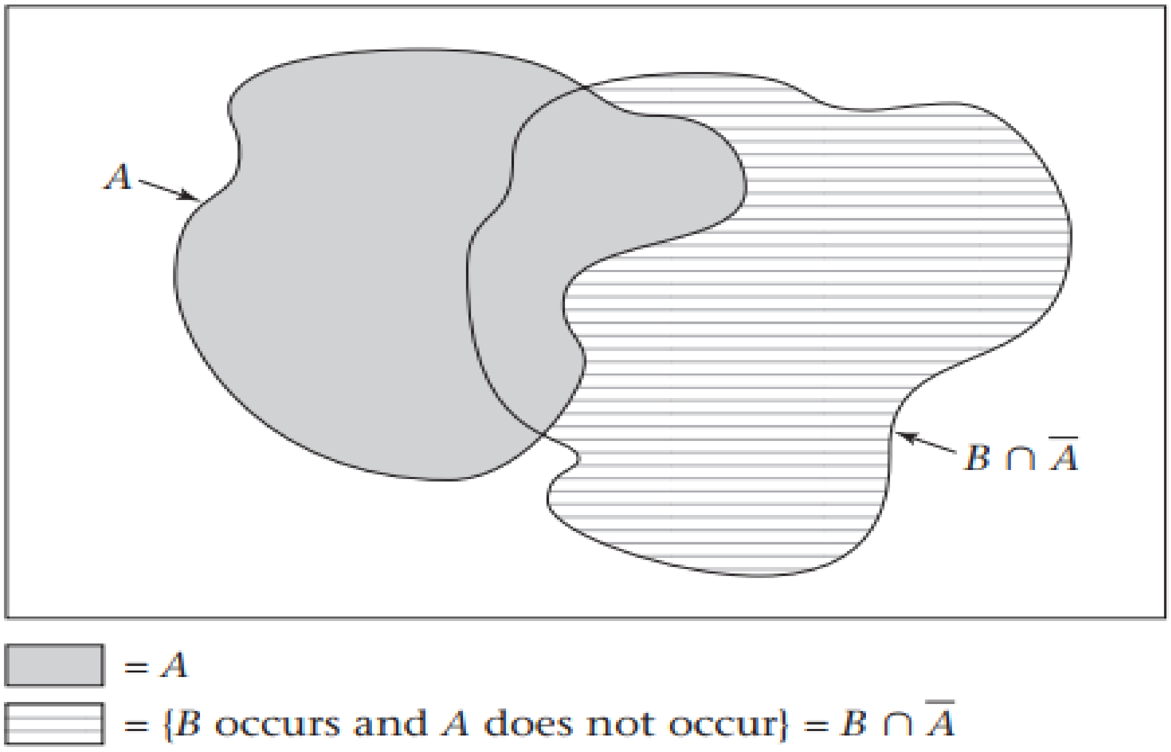
EQUATION 3.4 Special Case of Addition Law of Probability for Independent Events

If the two events A and B are independent, then we have the following rule:

$$P(A \cup B) = P(A) + P(B) \times [1 - P(A)]$$

This probability is diagrammed in Figure 3.6.

FIGURE 3.6 Diagrammatic representation of the addition law of probability for independent events



Example Hypertension

Let A = { mother's DBP ≥ 90 } and B = { father's DBP ≥ 90 }

If P (A) = 0.1 , P(B) = 0.2 and assuming that A and B are independent events, then:

What is the probability of a hypertensive household?

Notation: A "hypertensive household" is defined as one in which either the mother or the father is hypertensive, with hypertension defined for the mother and father, respectively, in terms of events A and B.

Solution

$$
\begin{aligned}
P(\text{hypertensive household}) &= P(A \cup B) = P(A) + P(B) \times [1 - P(A)] \\
&= 0.1 + 0.2 \times (1 - 0.1) \\
&= 0.1 + 0.18 \\
&= 0.28
\end{aligned}
$$

Thus, 28% of all households will be hypertensive.

Important Notation

It is possible to extend the addition law to more than two events. In particular, if there are three events A, B, and C, then

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

This result can be generalized to an arbitrary number of events, although that is beyond the scope of this textbook.

3.6 Conditional Probability

DEFINITION 3.9

The quantity $P(A|B) = \frac{P(A \cap B)}{P(B)}$; $P(B) \neq 0$ is defined as the conditional probability of A given B, which is written as P(A|B). Also, from the definition of conditional probability, we have:

$$P(A \cap B) = P(B)\, P(A|B)$$

Example Breast Cancer

Let A = {breast cancer} , B = {mammogram+}, and suppose we are interested in the probability of breast cancer (A) given that the mammogram is positive (B). This probability can be written P(A|B).

EQUATION 3.5

(1) If A and B are independent events, then $P(A|B) = P(A) = P(A|\overline{B})$.

(2) If two events A, B are dependent, then $P(A|B) \neq P(A) \neq P(A|\overline{B})$ and therefore $P(A \cap B) \neq P(A) \times P(B)$.

DEFINITION 3.10

The relative risk (RR) of A given B is given as follows: $RR = \frac{P(A|B)}{P(A|\overline{B})}$

Notice that if two events A, B are independent, then the RR is 1. If two events A, B are dependent, then the RR is different from 1. Heuristically, the more the dependence between events increases, the further the RR will be from 1.

Example

Sexually Transmitted Disease Suppose two doctors, A and B, test all patients coming into a clinic for syphilis. Let us define the following two events:

A^+={doctor A makes a positive diagnosis} ; B^+={doctor B makes a positive diagnosis}

$$P(A^+) = 0.10 ,\ P(B^+) = 0.17 \text{ and } P(A^+ \cap B^+) = 0.08$$

Answer the following

(a) Find the conditional probability that doctor B makes a positive diagnosis of syphilis given that doctor A makes a positive diagnosis?

Solution

$$P(B^+ | A^+) = \frac{P(A^+ \cap B^+)}{P(A^+)} = \frac{0.08}{0.10} = 0.8$$

Thus, doctor B will confirm doctor A's positive diagnoses 80% of the time.

(b) What is the conditional probability that doctor B makes a positive diagnosis of syphilis given that doctor A makes a negative diagnosis?

Solution

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

$$= \frac{P(B^+) - P(A^+ \cap B^+)}{1 - P(A^+)} = \frac{0.17 - 0.08}{1 - 0.10} = \frac{0.09}{0.9} = 0.1$$

(a) $P(A \cap \bar{B}) = P(A) - P(A \cap B)$

(b) $P(\bar{A} \cap B) = P(B) - P(A \cap B)$

Thus, when doctor A diagnoses a patient as negative, doctor B will contradict the diagnosis 10% of time.

(c) What is the RR of B^+ given A^+?

Solution

The relative risk (RR) of B^+ given A^+ is given as follows: $RR = \frac{P(B^+|A^+)}{P(B^+|A^-)} = \frac{0.8}{0.1} = 8$

This indicates that doctor B is 8 times as likely to diagnose a patient as positive when doctor A diagnoses the patient as positive than when doctor A diagnoses the patient as negative. These results quantify the dependence between the two doctors' diagnoses.

EQUATION 3.6 For any events A and B, we have

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

See Example 3.21 Page 53

DEFINITION 3.11 A set of events $A_1, \ldots, A_k$ is exhaustive if at least one of the events must occur.

EQUATION 3.7 Total-Probability Rule

Let A1, . . . , Ak be mutually exclusive and exhaustive events. The unconditional probability of B (P(B)) can then be written as a weighted average of the conditional probabilities of B given Ai (P(B | Ai)) with weights = P(Ai) as follows:

$$P(B) = \sum_{i=1}^{k} P(B|A_i)\, P(A_i)$$

Example

Ophthalmology: We are planning a 5-year study of cataract in a population of 5000 people 60 years of age and older. We know that:

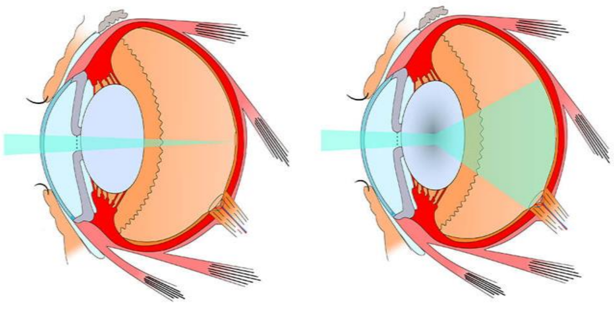
A1 = {ages 60–64} , A2 = {ages 65–69} , A3 = {ages 70–74} , A4 = {ages 75+}

Notice that: The above events are mutually exclusive and exhaustive because each person in our population must be in one and only one age group.

Define the event B = {develop cataract in the next 5 years}

And you know that:

P(A1) = 0.45 , P(B|A1) = 0.024
P(A2) = 0.28 , P(B|A2) = 0.046
P(A3) = 0.20 , P(B|A3) = 0.088
P(A4) = 0.07 , P(B|A4) = 0.153



then find P(B)?

Solution By using the total-probability rule, we have:

$$P(B) = \sum_{i=1}^{k=4} P(B|A_i)\, P(A_i)$$
$$= P(B|A1)\, P(A1) + P(B|A2)\, P(A2) + P(B|A3)\, P(A3) + P(B|A4)\, P(A4)$$
$$= (0.024)(0.45) + (0.046)(0.28) + (0.088)(0.20) + (0.153)(0.07)$$
$$= 0.05199 \cong 0.052$$

Conclusion: Thus 5.2% of this population will develop cataract over the next 5 years, which represents a total of 5000 × 0.052 = 260 people with cataract.

Notation: The definition of conditional probability also allows the multiplication law of probability to be extended to the case of dependent events as follows:

EQUATION 3.8

Generalized Multiplication Law of Probability

If $A_1, \ldots, A_k$ are an arbitrary set of events, then

$$Pr(A_1 \cap A_2 \cap \cdots \cap A_k)$$
$$= Pr(A_1) \times Pr(A_2|A_1) \times Pr(A_3|A_2 \cap A_1) \times \cdots \times Pr(A_k|A_{k-1} \cap \cdots \cap A_2 \cap A_1)$$

If the events are independent, then the conditional probabilities on the right hand side of Equation 3.8 reduce to unconditional probabilities and the generalized multiplication law reduces to the multiplication law for independent events given in Equation 3.2 below:

EQUATION 3.2

Multiplication Law of Probability

If $A_1, \ldots, A_k$ are mutually independent events,

then $Pr(A_1 \cap A_2 \cap \ldots \cap A_k) = Pr(A_1) \times Pr(A_2) \times \ldots \times Pr(A_k)$

3.7 Bayes ' Rule and Screening Tests

(a) Screening Tests

In the health sciences field a widely used application of probability laws and concepts is found in the evaluating of screening tests and diagnostic criteria. Our interest is to predict the presence or absence of a particular disease from a knowledge of a test results (positive or negative). The general concept of the predictive value of a screening test can be defined as follows:

DEFINITION 3.12

(a) Predictive Value Positive (PV+) of a Screening Test

It is the probability that a person has a disease given that the test is positive.

P(Disease | Test+)

(b) Predictive Value Negative (PV−) of a Screening Test

It is the probability that a person does not have a disease given that the test is negative.

P(No Disease | Test −)

Example

Cancer Suppose that in a particular study conducted on a random sample of women to diagnose whether or not they have a breast cancer within two years, we have the following probabilities:

Let

$A = \{\text{mammogram}^+\}$ with $P(A) = 0.07$

$B = \{\text{breast cancer}\}$ with $P(B) = 0.0072$

$B \cap A = \text{breast cancer} \cap \text{mammogram}^+$

with $P(\text{breast cancer} \cap \text{mammogram}^+) = P(B \cap A) = 0.007$

Now:

(a) Suppose that we are interested in the probability of breast cancer (B) given that the mammogram is positive (A), that is:

$$P(B \mid A) = \frac{P(\text{breast cancer} \cap \text{mammogram}^+)}{P(\text{mammogram}^+)} = \frac{P(B \cap A)}{P(A)} = \frac{0.007}{0.07} = 0.10$$

(b) Suppose that we are interested in the probability of no breast cancer ($\bar{B}$) given that the mammogram is negative ($\bar{A}$), that is:

$$P(\bar{B}|\bar{A}) = \frac{P(\text{breast cancer}^- \cap \text{mammogram}^-)}{P(\text{mammogram}^-)} = \frac{P(\bar{B} \cap \bar{A})}{P(\bar{A})}$$

$$= \frac{P(\overline{A \cup B})}{P(\bar{A})} = \frac{1 - P(A \cup B)}{1 - P(A)}$$

$$= \frac{1 - [\, P(A) + P(B) - P(A \cup B)]}{1 - P(A)}$$

$$= \frac{1 - [\, 0.07 + 0.0072 - 0.007]}{1 - 0.07}$$

$$= \frac{1 - 0.0702}{0.93} = \frac{0.9298}{0.93} = 0.99978$$

(c) Find PV^+ and PV^- for mammography by using the given data in this Example?

Solution

(i) $PV^+ = P(\text{breast cancer} \mid \text{mammogram}^+) = P(B \mid A) = 0.1$

(ii) $PV^- = P(\text{breast cancer}^- \mid \text{mammogram}^-) = P(\overline{B} \mid \overline{A}) = 0.99978 \approx 1$

Thus, if the mammogram is negative, the woman is virtually certain not to develop breast cancer over the next 2 years ($PV^- \cong 1$); whereas if the mammogram is positive, the woman has a 10% chance of developing breast cancer ($PV^+ = 0.10$).

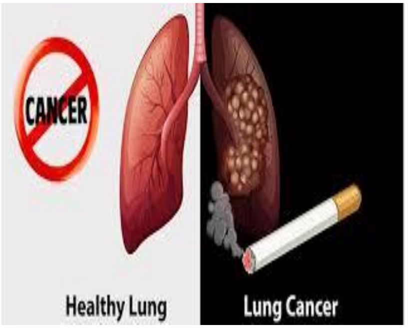
A symptom or a set of symptoms can also be regarded as a screening test for disease. The higher the PV of the screening test or symptoms, the more valuable the test will be. Ideally, we would like to find a set of symptoms such that both PV^+ and PV^- are 1. Then we could accurately diagnose disease for each patient. However, this is usually impossible. Clinicians often cannot directly measure the PV of a set of symptoms. However, they can measure how often specific symptoms occur in diseased and normal people. These measures are defined as follows:

DEFINITION 3.13 The **sensitivity** of a symptom (or set of symptoms or screening test) is the probability that the symptom is present given that the person has a disease.

DEFINITION 3.14 The **specificity** of a symptom (or set of symptoms or screening test) is the probability that the symptom is *not* present given that the person does *not* have a disease.

DEFINITION 3.15 A false negative is defined as a negative test result when the disease or condition being tested for is actually present. A false positive is defined as a positive test result when the disease or condition being tested for is not actually present.

EXAMPLE 3.24



Cancer Suppose the disease is lung cancer and the symptom is cigarette smoking. If we assume that 90% of people with lung cancer and 30% of people without lung cancer (essentially the entire general population) are smokers, then the sensitivity and specificity of smoking as a screening test for lung cancer are .9 and .7, respectively. Obviously, cigarette smoking cannot be used by itself as a screening criterion for predicting lung cancer because there will be too many false positives (people without cancer who are smokers).

EXAMPLE 3.25

Cancer Suppose the disease is breast cancer in women and the symptom is having a family history of breast cancer (either a mother or a sister with breast cancer). If we assume 5% of women with breast cancer have a family history of breast cancer but only 2% of women without breast cancer have such a history, then the sensitivity of a family history of breast cancer as a predictor of breast cancer is .05 and the specificity is .98 = (1 − .02). A family history of breast cancer cannot be used by itself to diagnose breast cancer because there will be too many false negatives (women with breast cancer who do not have a family history).

Notation: For a symptom to be effective in predicting disease, it is important that both the sensitivity and specificity be high.

(b) Bayes' Rule

Definition: Cases are participants possessing a condition of interest (disease).

Controls are participants lacking that (those without it) condition (disease).

Example

If we are interested in the effects of exposure to a particular chemical on diagnoses of mesothelioma, a type of cancer. Here, you would compare those exposed (cases) with those not exposed (controls) to the chemical. This would allow us to observe whether the people exposed to the chemical had more instances of mesothelioma than those who weren't exposed.

The PV^+ and PV^- can be directly evaluate from the data provided. Instead, in many screening studies, a random sample of cases and controls is obtained, then one can estimate sensitivity and specificity from such a design. However, because cases are usually oversampled relative to the general population (e.g., if there are an equal number of cases and controls), one cannot directly estimate PV^+ and PV^- from the frequency counts available in a typical screening study. Instead, an indirect method known as Bayes' rule is used for this purpose.

General Question: How can the sensitivity and specificity of a symptom (or set of symptoms or diagnostic test), which are quantities a physician can estimate, be used to compute PVs (PV^+ and PV^-), which are quantities a physician needs to make appropriate diagnoses?

Definition

Let A = symptom and B = disease. Let P(B) = Probability of disease in the reference population. We wish to compute $P(B \mid A)$ and $P(\bar{B} \mid \bar{A})$ in terms of the other quantities. This relationship is known as Bayes' rule . From Definitions 3.12, 3.13, and 3.14, we have the following:

Predictive value positive = $PV^{+} = P(B \mid A)$ → Sensitivity = $P(A \mid B)$

Predictive value negative = $PV^{-} = P(\bar{B} \mid \bar{A})$ → Specificity = $P(\bar{A} \mid \bar{B})$

EQUATION 3.9

Bayes' Rule

Let A = symptom and B = disease.

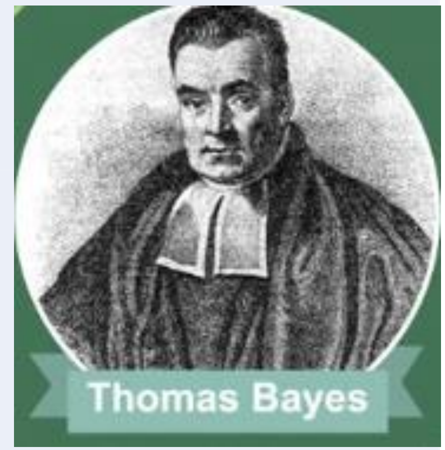
$$PV^{+} = Pr(B|A) = \frac{Pr(A|B) \times Pr(B)}{Pr(A|B) \times Pr(B) + Pr(A|\bar{B}) \times Pr(\bar{B})}$$

In words, this can be written as

$$PV^{+} = \frac{\text{Sensitivity} \times x}{\text{Sensitivity} \times x + (1 - \text{Specificity}) \times (1 - x)}$$

where $x = Pr(B)$ = prevalence of disease in the reference population. Similarly,

$$PV^{-} = \frac{\text{Specificity} \times (1 - x)}{\text{Specificity} \times (1 - x) + (1 - \text{Sensitivity}) \times x}$$



That is, PV^{+} can be expressed as a function of sensitivity, specificity, and the probability of disease in the reference population. A similar expression can be used to obtain PV^{-}.

EXAMPLE 3.26

Hypertension Suppose 84% of hypertensives and 23% of normotensives are classified as hypertensive by an automated blood-pressure machine. What are the PV^+ and PV^- of the machine, assuming 20% of the adult population is hypertensive?

Solution

Let A = symptom and B = hypertensive

P(B) = P(the adult population is hypertensive) = 20% = 0.2 = x

Sensitivity = P(hypertensives classified by an automated blood-pressure machine)

= $P(A \mid B)$ = 84% = 0.84

Specificity = P(normotensives classified by an automated blood-pressure machine)

= $P(\bar{A} \mid \bar{B})$ = 1 - 23% = 1 - 0.23 = 0.77

Thus, from Bayes' rule it follows that:

(a) Predictive Value Positive

$$PV^+ = \frac{\text{Sensitivity} \times x}{\text{Sensitivity} \times x + (1 - \text{Specificity}) \times (1 - x)}$$

$$= \frac{(0.84)(0.20)}{[(0.84 * 0.20) + (0.23 * 0.80)]} = \frac{0.168}{0.352} = 0.48$$

(b) Predictive Value Negative

$$PV^{-} = \frac{\text{Specificity} \times (1 - x)}{\text{Specificity} \times (1 - x) + (1 - \text{Sensitivity}) \times x}$$

$$= \frac{(0.77)(0.80)}{[(0.77 * 0.80) + (0.16 * 0.20)]} = \frac{0.616}{0.648} = 0.95$$

Conclusion

Thus, a negative result from the machine is reasonably predictive because we are 95% sure a person with a negative result from the machine is normotensive. However, a positive result is not very predictive because we are only 48% sure a person with a positive result from the machine is hypertensive.

Notation

Example 3.26 considered only two possible disease states: hypertensive and normotensive. In clinical medicine there are often more than two possible disease states. We would like to be able to predict the most likely disease state given a specific symptom (or set of symptoms). Let's assume that the probability of having these symptoms among people in each disease state (where one of the disease states may be normal) is known from clinical experience, as is the probability of each disease state in the reference population. This leads us to the generalized Bayes' rule.

Equation 3.10 Generalized Bayes' Rule

Let B1 B2 , . . . , Bk be a set of mutually exclusive and exhaustive disease states; that is, at least one disease state must occur and no two disease states can occur at the same time. Let A represent the presence of a symptom or set of symptoms. Then, $P(B_i|A)$ can be calculated as follows:

$$Pr(B_i|A) = \frac{Pr(A|B_i) \times Pr(B_i)}{\sum_{j=1}^{k} Pr(A|B_j) \times Pr(B_j)}$$

EXAMPLE 3.27 Pulmonary Disease

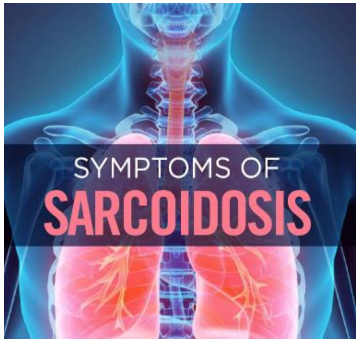
Suppose a 60-year-old man who has never smoked cigarettes presents to a physician with symptoms of a chronic cough and occasional breathlessness. The physician becomes concerned and orders the patient admitted to the hospital for a lung biopsy. Suppose the results of the lung biopsy are consistent either with lung cancer or with sarcoidosis, a fairly common, usually nonfatal lung disease. In this case we have:

A = {chronic cough, results of lung biopsy}

Disease state: B1={normal} ; B2={lung cancer} ; B3={sarcoidosis}

and that in the 60-year-old, never-smoking men, we have:

$P(B_1) = 0.99$; $P(B_2) = 0.001$; $P(B_3) = 0.009$

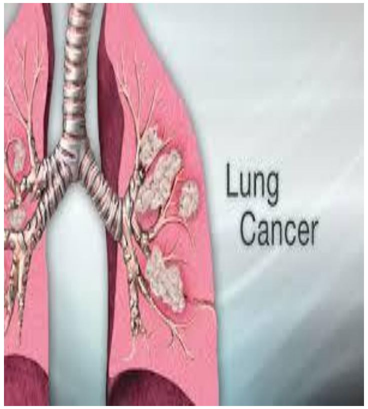


Suppose also that

$$P(A|B_1) = 0.001 \quad ; \quad P(A|B_2) = 0.900 \quad ; \quad P(A|B_3) = 0.900$$

The interesting question now becomes what are the probabilities P(Bi|A) of the three disease states given the previous symptoms?

Solution Bayes' rule can be used to answer this question follows:



$$P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{\sum_{i=1}^{k=3} P(A|B_i)\,P(B_i)} \quad ; \quad i = 1, 2, 3$$

For $i = 1$, we have $P(B_1|A) = \frac{P(A|B_1)\,P(B_1)}{\sum_{i=1}^{k=3} P(A|B_i)\,P(B_i)}$

$$= \frac{(0.001)(0.99)}{(0.001)(0.99) + (0.900)(0.001) + (0.900)(0.009)} = 0.0991$$

For $i = 2$, we have $P(B_2|A) = \frac{P(A|B_2)\,P(B_2)}{\sum_{i=1}^{k=3} P(A|B_i)\,P(B_i)}$

$$= \frac{(0.900)(0.001)}{(0.001)(0.99) + (0.900)(0.001) + (0.900)(0.009)} = 0.0901$$

For $i = 3$, we have $P(B_3|A) = \frac{P(A|B_3)\,P(B_3)}{\sum_{i=1}^{k=3} P(A|B_i)\,P(B_i)}$

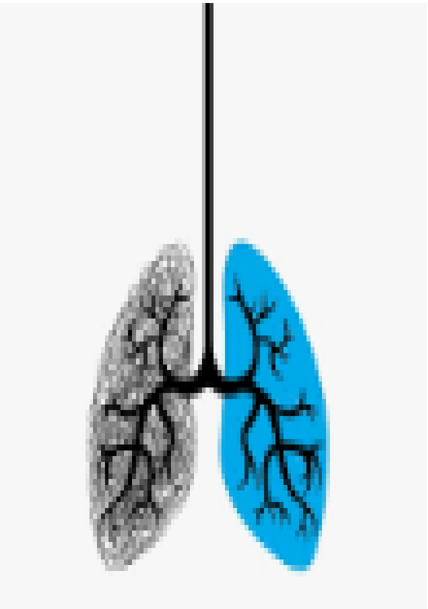
$$= \frac{(0.900)(0.009)}{(0.001)(0.99) + (0.900)(0.001) + (0.900)(0.009)} = 0.8108$$

EXAMPLE 3.28

Pulmonary Disease Now suppose the patient in Example 3.27 smoked two packs of cigarettes per day for 40 years. Then assume $Pr(B_1) = .98$, $Pr(B_2) = .015$, and $Pr(B_3) = .005$ in this type of person. What are the probabilities of the three disease states for this type of patient, given these symptoms?

Solution:

$$Pr(B_1|A) = .001(.98)/[.001(.98) + .9(.015) + .9(.005)]$$
$$= .00098 / .01898 = .052$$
$$Pr(B_2|A) = .9(.015) / .01898 = .01350 / .01898 = .711$$
$$Pr(B_3|A) = .9(.005) / .01898 = .237$$



Thus, in this type of patient (i.e., a heavy-smoking man) lung cancer is the most likely diagnosis.

Problems: 3.1-3.25, 3.32-3.36, 3.68-3.73.

Probability Rules

De Morgan's Laws

1. $\overline{A \cup B} = \bar{A} \cap \bar{B}$ 2. $\overline{A \cap B} = \bar{A} \cup \bar{B}$

$$(a)\ P(\overline{A \cup B}) = P(\bar{A} \cap \bar{B}) = 1 - P(A \cup B)$$

$$(b)\ P(\overline{A \cap B}) = P(\bar{A} \cup \bar{B}) = 1 - P(A \cap B)$$

Rules Involving the Empty Set (φ) and the Entire Event (S)

1. $A \cup \phi = A$ 2. $A \cap \phi = \phi$

3. $A \cup S = S$ 4. $A \cap S = A$

Important Probability Rules

$$P(\bar{A}) = 1 - P(A)$$

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

Special Case

When the two events A and B are **mutually exclusive (disjoint)**, then $P(A \cap B) = P(\phi) = 0$,

and therefore $P(A \cup B) = P(A) + P(B)$.

Additional Rules

$$(a)\ \ P(A \cap \overline{B}) = P(A) - P(A \cap B)$$

$$(b)\ \ P(\overline{A} \cap B) = P(B) - P(A \cap B)$$

$$(c)\ \ P(A \cup \overline{B}) = P(\overline{B}) + P(A \cap B)$$

$$(d)\ \ P(\overline{A} \cup B) = P(\overline{A}) + P(A \cap B)$$

Conditional Probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \text{ if } P(B) \neq 0$$

"given"

(1) $P(B/A)=\dfrac{P(A\cap B)}{P(A)}$ if $P(A)>0 \Rightarrow P(A\cap B)=P(A)\,P(B/A)$.

(2) $P(A/\bar{B})=\dfrac{P(A\cap\bar{B})}{P(\bar{B})}=\dfrac{P(A)-P(A\cap B)}{1-P(B)}$.

(3) $P(\bar{A}/B)=\dfrac{P(\bar{A}\cap B)}{P(B)}=\dfrac{P(B)-P(A\cap B)}{P(B)}=1-P(A/B)$.

(4) $P(\bar{A}/\bar{B})=\dfrac{P(\bar{A}\cap\bar{B})}{P(\bar{B})}=\dfrac{P(\overline{A\cup B})}{1-P(B)}=\dfrac{1-P(A\cup B)}{1-P(B)}=\dfrac{1-\left[P(A)+P(B)-P(A\cap B)\right]}{1-P(B)}$.

Exercises

Exercise (1): If the probability a person weight become normal after joining a health club for 3 months is 0.35, find the probability that a person weight joining this club does not become normal after the 3 months? Answer: 0.65.

Exercise (2): Suppose that a researcher asked 25 people if they liked the taste of a new fruit drink. The responses were classified as “yes”, “no” or “undecided.” The results were categorized in simple frequency table as shown below:

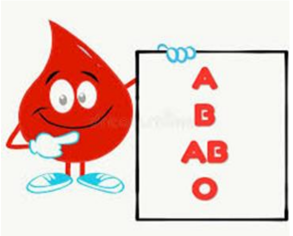
Response	Yes	No	Undecided	Total
Frequency	15	8	2	25

What is the probability of selecting a person who liked the taste? Answer: 0.60.

Exercise (3): In a shipment of 25 treadmill to a health club in Jordan, 2 are defective. If two devices are randomly selected and tested, find the probability that both are defective if the first one is not replaced after it has been tested? Answer: 0.00333.

Exercise (4): In a random sample of 50 people selected from a GYM in Jordan, 21 had type O blood, 22 had type A blood, 5 had type B blood, and 2 had type AB blood. The results were categorized in simple frequency table as shown below:

Blood Type	A	B	AB	O
Frequency	22	5	2	21



Find the following probabilities:

a) What is the probability of selecting a person has type O blood? Answer: 0.42.

b) What is the probability of selecting a person has type A or type B blood? Answer: 0.54.

c) What is the probability of selecting a person has neither type A nor type O blood? Answer: 0.14.

d) What is the probability of selecting a person does not have type AB blood? Answer: 0.96.

Exercise (5): The hospital records from Jordan indicates that maternity patients stayed in the hospital for the number of days shown in the simple frequency table below:

Number of days stayed	3	4	5	6	7	Total
Frequency	15	32	56	19	5	127

Find the following probabilities:

a) What is the probability of selecting a patient stayed exactly 5 days? Answer: 0.441.

b) What is the probability of selecting a patient stayed less than 6 days? Answer: 0.811.

c) What is the probability of selecting a patient stayed at most 4 days? Answer: 0.370.

d) What is the probability of selecting a patient stayed at least 5 days? Answer: 0.630.

Exercise (6): In a hospital unit there are 8 nurses and 5 physicians. Seven nurses and three physicians are females as shown in the table below:

Staff	Sex		Total
	Male	Female	
Nurses	1	7	**8**
Physicians	2	3	**5**
Total	**3**	**10**	**13**

If a staff person is selected, find the probability that the person is a nurse or a male? Answer: 0.76923.

Exercise (7): In a pizza restaurant, 95% of the customers order pizza, 68% of the customers order salad and 65% of the customers order both pizza and salad. Are the two events, order pizza and order salad independent?

Exercise (8): The table below shows the smoking habits for group of cancer patients selected randomly from the King Hussein Cancer Center (KHCC) in Jordan:

Sex	Smoking Habits			Total
	Non-Smoker (N)	Smoker (S)	Heavy Smoker (H)	
Male (M)	384	97	49	530
Female (F)	349	116	38	503
Total	733	213	87	1033

If a patient is selected at random from this group, then the value of $P(F \cap \overline{N}) = ?$ Answer: 0. 149.

THE END