CHAPTER THREE

Elementary Probability

"Life is a school of probability."

Walter Bagehot

The notion that chance, or probability, can be treated numerically is relatively recent. Indeed, for most of recorded history it was felt that what occurred in life was determined by forces that were beyond one's ability to understand. It was only during the first half of the 17th century, near the end of Renaissance, that people become curious about the world and the laws governing its operation. Among the curious were the gamblers.

A cynical person once said, "*The only two sure things are death and taxes*." This philosophy no doubt arose because so much in people's lives is affected by chance.

5.1 Introduction

Probability as a general concept can be defined as the chance of an event occurring. Most people are familiar with probability from observing or playing games of chance, such as card games or lotteries. Probability is the basis of *inferential statistics.*

The basic concepts of probability are explained in this chapter. These concepts include *probability experiments*, *sample spaces*, the *addition* and *multiplication rules*, and the *probabilities of complementary events.* Also in this chapter, you will learn the rule for counting, the differences between permutations and combinations, and how to figure out how many different combinations for specific situations exist. Section 4–5 explains how the counting rules and the probability rules can be used together to solve a wide variety of problems. Finally in section six, the concept of probability is extended to conditional probability and independence.

At the end of this chapter students are expected to:

- ✓ *Know what is meant by sample space, event, relative frequency, probability, conditional probability, independence.*

5.2 Definitions of Some concepts of Probability Terms

Terms that are most frequently used and cornerstone of probability are defined as follows:

- ***Probability experiment:*** It is a process that leads to well-defined results called *outcomes.* For example, flipping a coin once, rolling one die once, or the like.
- ***Outcomes:*** It is the result of a single trial of probability experiment. It is sometimes called *sample point. Example 5.1:* when a coin is tossed once, there are two possible outcomes: **head** or **tail**. In the roll of a single die, there are six possible outcomes: 1**, 2, 3, 4, 5, or 6**
- ***Sample Space:*** It is the set of all possible outcomes of a probability experiment and denoted by$S$ *or* Ω. *Example 5.2:* consider example 5.1, S={H, T}, S={1, 2, 3, 4, 5, 6}
- ***Event:*** It is a subset of sample space (contains one or more outcomes which are in the sample space) and is defined for a particular purpose. An event can be one outcome or more than one outcome. *Simple event* is an event having only single outcome. *Compound event* consisting of one or more outcomes or simple events. Event is denoted by capital letters such as A, B, F etc. *Example 5*.3: let A be the event of odd number in tossing a die experiment, then A={1, 3, 5}
- ***Mutually exclusive events:*** Suppose you have two events, say A and B. if these events have no common sample point(s) or do not occur simultaneously, then the two events are called *mutually exclusive events. Example 5.4:* consider experiment o tossing a die. Let A be the event of odd numbers and B be the event of even numbers, A={1, 3, 5}, B={2, 4, 6}, then A and B are mutually exclusive events.
- ***Exhaustive events:*** It is a satiation where the events contain all elements based on the definition of the events. For example S={Head, Tail} is exhaustive for tossing a coin experiment.
- ***Union of events:*** The union of two events A and B, denoted by$A \cup B$, consists of all outcomes that are in A or in B or both A and B. *Example 5.5:* let A={1, 3, 5}, B={2, 4, 5, 6} then AUB={1, 2, 3, 4, 5, 6}
- ***Intersection of events:*** The intersection of event A and B, denoted by$A \cap B$, consists of all outcomes that are in both A and B. *Example 5.6:* A={1, 3, 5}, B={2, 4, 5, 6} then A∩B={5}

- ***Compliment of an event:*** The compliment of event A, denoted by A^c or A', consists of all outcomes that are not in A. *Example 5.7*: Let Sample space S={1, 2, 3, 4, 5, 6}and event A={1, 3, 5}then A^c={2, 4, 6}
- ***Null event:*** The event containing no outcomes. It is the compliment of the sample space.
- ***Probability of an event:*** The probability of event A, denoted by $P(A)$, is the probability the outcome of the experiment is contained in A.
- ***Equally-likely events:*** It is a situation where the probability of the occurrence of one event as likely as the other event. That is, they must have equal probability of occurrence. *Example 5.8:* In example 5.1, outcomes: 1, 2, 3, 4, 5, and 6 are equally likely.
- ***Independent events:*** Two events said to be independent if knowing whether a specific one has occurred does not change the probability that the other occurs. (Example is explained in section 5.6).

5.3 Counting Rules

In order to probabilities, we have to know

- ✓ The elements of an event.
- ✓ The number of elements of the sample space.

That is in order to judge what is probable, we have to know what is possible.

In order to determine the number of outcomes one can use several rules of counting.

1. Addition rule
2. Multiplication rule
3. Permutation rule
4. Combination rule

5.3.1 Addition Rule

If the choices can't be performed together then the number of ways in which you can make a choice in $n_1 + n_2 + n_3 + \cdots + n_k$ different ways.

Example 5.9: If there are two way of bus to voyage Debre Derhan from Addis Ababa and three railways, then collectively we have $2 + 3 = 5$ different way to arrive Debre Berhan.

5.3.2 Multiplication (Fundamental) Rule

In sequence of n events in which the first one has k_1 possibilities and the second event has k_2 and the third has k_3, and so forth, the total number of possibilities of the sequence will be $k_1.k_2.k_3 \dots k_n$.

Example 5.10: How many different 7-place license plates are possible if the first 3 places are to be occupied by letters and the final 4 by numbers?

Solution: By multiplication rule the answer is $26.26.26.10.10.10.10 = 175{,}760{,}000$

Example 5.11: In the above example, how many license plates would be possible if repetition among letters or numbers were prohibited?

Solution: In this case there would be $26.25.24.10.9.8.7 = 78,624,000$ possible plates.

5.3.3 Permutation Rule

How many different ordered arrangements of letters a, b, c are possible? By direct enumeration we see that there are 6: namely, abc, acb, bac, bca, cab and cba. Each arrangement is known as a *permutation*. That is a **permutation** is an arrangement of n objects in a specific order. Thus, there are six possible permutations of a set of 3 objects. This result could also have been obtained from the basic principle, since the first object in the permutation can be any of the 3, the second object in the permutation can then be chosen any of the remaining 2, and the third object in the permutation is then chosen the remaining one. Thus there are $3.2.1 = 6$ possible permutations.

Permutation Rule 1: Suppose now that we have n objects. Reasoning, similar to that we have just used for the 3 letter shows that there are

$$n.(n-1).(n-2)\dots 3.2.1 = n!$$

Different permutations of the n objects

Example 5.12: A class of stat 173 consists of 6 men and 4 women. An examination is given, and the students are ranked according to their performance. Assume that no two students obtain the same score.

A. How many different rankings are possible?
B. If the men are ranked just among themselves and women among themselves, how many different rankings are possible?

Solution:

A. As each ranking corresponds to a particular ordered arrangement of the 10 people, we see that the answer to this part is $10! = 3,628,800$

B. As there are 6! possible rankings of the men among themselves and 4! possible rankings of the women among themselves, it follows from the basic principle that the two groups arrange themselves; it follows the basic principle that the two groups arrange themselves in 2! way so that we have a total of $6!.4!.2! = 34560$ possible rankings.

Permutation Rule 2: We shall now determine the number of permutations of a set of n objects when certain of the objects are *indistinguishable* from each other. Then the formula is:

$$\frac{n!}{n_1!.n_2!\ldots n_r!}$$

Different permutations of n objects, of which n_1 are alike n_2 are alike, …, n_r are alike.

Example 5.13: How many different letter arrangements can be formed using the letter PEPPER?

Solution: $\frac{6!}{3!.2!.1!} = 60$ possible teller arrangements.

Permutation Rule 3: Generally, if we are asked to arrange r objects among n objects, then we will have the following total arrangements

$$nP_r = \frac{n!}{(n-r)!}$$

Example 5.14: Suppose a business man has a choice of five locations in which to establish his business. He wishes to arrange only the top three locations. How many different ways can he arrange them?

Solution:

$$5P_3 = \frac{5!}{(5-3)!} = 60\,ways$$

5.3.4 Combination Rule

We are often interested in determining the number of different groups of r objects that could be formed from a total of n objects. A selection of objects without regard to order is called a **combination**. That is, combinations are used when the order or arrangement is not important. The number of combinations of r objects selected from n objects is denoted by nC_r and is given by the formula $nC_r = \frac{n!}{r!(n-r)!} = \binom{n}{r}$

Example 5.15: From a group of 5 women and 7 men, how many different committees consisting of 2 women and 3 men can be performed? What if 2 of the men are feuding and refuse to serve on the committee together?

Solution: As there are $\binom{5}{2}$ possible groups of 2 women, and $\binom{7}{3}$ possible groups of 3 me, it follows from the basic principle that there are

$$\binom{5}{2}\binom{7}{3} = \left(\frac{5.4}{2.1}\right)\frac{7.6.5}{3.2.1} = 350$$

Possible committees consisting of 2 women and 3 men. On the other hand, if 2 of the men refuse to serve on the committee together, then, as there are $\binom{2}{0}\binom{5}{3}$ possible group of 3 men not containing either of the 2 feuding men and $\binom{2}{1}\binom{5}{2}$ groups of 3 men containing exactly 1 of the feuding men, it follows that there are $\binom{2}{0}\binom{5}{3} + \binom{2}{1}\binom{5}{2} = 30$ groups of 3 men not containing both of the feuding men. Since there are $\binom{5}{2}$ ways to choose the 2 women, it follows that in this case there are $30\binom{5}{2} = 300$ possible committees.

5.4 Approaches in Probability Definition

The probability of an event is denoted by $P(.)$ where *P* stands for probability and the *dot* stands for any event, say A, B, G etc.

Generally approaches to probability can be divided into two, namely subjective approach and objective approach.

5.4.1 **Subjective approach:**

A probability derived from an individual's personal judgment about whether a specific outcome is likely to occur. Subjective probabilities contain no formal calculations and only reflect the subject's opinions and past experience.

Subjective probabilities differ from person to person. Because the probability is subjective, it contains a high degree of personal bias. An example of subjective probability could be asking Arsenal fan, before the football season starts, the chances of Arsenal winning the world champions. While there is no absolute mathematical proof behind the answer to the example, fans might still reply in actual percentage terms, such as the Arsenal having a 95% chance of winning the world champions.

5.4.2 Objective approach:

The probability of an event in a certain experiment based on an experimental evidence or random process. In this approach to study probability theory there are three sub approaches.

These are

- The classical approach
- The frequentist approach
- The axiomatic approach and

5.4.3.1 The Classical Approach

If a procedure has n different simple events, each with an equal chance of occurring, and event A can occur in s of these ways, then

$$P(A) = \frac{n(A)}{n(S)} = \frac{number\ of\ elements\ in\ A}{number\ of\ elements\ in\ a\ sample\ space}$$

Assumptions in classical approach

- The outcomes must be equally-likely
- The experiment should never be repeated more than once
- The sample space should be finite

Example 5.16: Toss a fair coin once and find the probability of the occurrence of head

Solution: Since the sample space is finite i.e., either head or tail and the outcomes are equally-likely

$$P(Head) = \frac{n(head)}{n(Sample\ space)} = \frac{1}{2} = 0.5$$

Example 5.17: For a card drawn from an ordinary deck, find the probability of getting a queen.

Solution: Since there are 4 queens and 52 cards, $P(queen) = \frac{4}{52}$

If one of the assumptions stated above is violated, the classical approach no longer valid

5.4.3.2 Frequentist (empirical) Approach

If after n repetition of an experiment, where n is very large, an event is observed to occur in h of these, then the probability of an event is $\frac{h}{n}$ or conduct an experiment a large number of times, and count the number of times event A actually occurs, then an estimate of $P(A)$ is

$$P(A) \approx \frac{number\ of\ times\ A\ ocuured}{number\ of\ times\ trial\ was\ repeated}$$

Example 5.18: Suppose a coin was tossed 1000 times and the result was 587 tails. The relative frequency of tails is $\frac{587}{1000}$. Another 1000 tosses lead to 511 tails. Then the relative frequency of tails is $\frac{587+511}{1000+1000} = \frac{1098}{2000}$. Proceeding, in this manner we obtain a sequence of numbers, which gets closer and closer to the number defined as the probability of a trial in a single toss.

Therefore,

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

5.4.3.3 Axiomatic Approach

Both the classical and frequentist approaches have serious drawbacks, the first because the words *"equally likely"* are vague and the second because the "*large number*" involved is vague. Because of these difficulties, statisticians have been led to an *axiomatic approach* of probability.

Axiom 1: For every event A, $P(A) \geq 0$

Axiom 2: For the sure or certain event, $P(S) = 1$

Axiom 3: For any number of mutually exclusive events A_1, A_2, A_3 ...

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = P(A_1) + P(A_2) + P(A_3) + \cdots$$

In particular, for two mutually exclusive events A_1 *and* A_2

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

5.5 Some Probability Rules

Rule 1: If $A_1 \subseteq A_2$, then $P(A_{1)} \leq P(A_2)$

Rule 2: For every event A, $0 \leq P(A) \leq 1$ i.e. a probability between 0 and 1.

Rule 3: For ϕ, the empty set, $P(\phi) = 0$ i.e. the impossible event has probability zero.

Rule 4: If A' is the complement of A, then $P(A') = 1 - P(A)$

Rule 5: If A and B are any two events, then $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$

More generally, if A_1, A_2, A_3 are three events, then

$$P(A_1 \bigcup A_2 \bigcup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \bigcap A_2) - P(A_2 \bigcap A_3) - P(A_3 \bigcap A_1) + P(A_1 \bigcap A_2 \bigcap A_3)$$

Rule 6: $P(A \bigcap B) = P(A)\ P(B \mid A)\ \ or\ P(A \bigcap B) = P(B)\ P(A \mid B)$

$P(A \bigcap B) = P(A)P(B)$, for independent event

Example 5.11: Suppose we toss two coins and suppose that each of the four points in the sample space is $S = \{HH, HT, TH, TT\}$ equally likely and hence has probability $\frac{1}{4}$. Let E is the event that the first coin falls head, and F is the event that the second coin falls heads.

Solution: $E = \{HH, HT\} and\ F = \{HH, TH\}$. Then the probability of either the first or the second coin falls head is

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) = \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{4} = \tfrac{3}{4}$$

5.5 Conditional Probability and Independence

Let A and B two events such that $P(A) > 0$. Denote $P(B \mid A)$ the probability of B given that A has occurred since A is know to have occurred; it becomes the new sample replacing the original S. From this we are led to the definition

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0 \ \ or$$

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

In words, this is saying that the probability that both A and B occur is equal to the probability that A occurs times the probability that B occurs given that has occurred. We call $P(B \mid A)$ the *conditional probability of B given A*, i.e. the probability that B will occur given that A has occurred.

Example 5.12: A jar contains black and white marbles. Two marbles are chosen without replacement. The probability of selecting a black marble and then a white marble is 0.34, and the probability of selecting a black marble on the first draw is 0.47. What is the probability of selecting white marble on the second draw, given that the first marble drawn was black?

Solution: $P(White \mid Black) = \dfrac{P(Black\ and\ White)}{P(Black)} = \dfrac{0.34}{0.47} = 0.72$

Example 5.13: The probability that it is Friday and that a student is absent is 0.03. Since there are 5 schooldays in a week, the probability that it is Friday is 0.2. What is the probability that a student is absent given that today is Friday?

Solution: $P(Absent \mid Friday) = \dfrac{P(Friday\ and\ Absent)}{P(Firday)} = \dfrac{0.03}{0.2} = 0.15$

It often happens that the knowledge that a certain event E has occurred has no effect on the probability that some other event F has occurred, that is, that $P(E \mid F) = P(E)$. One would expect that in this case, the equation $P(F \mid E) = P(F)$ would also be true. If these equations are true, we might say the F is *independent* of E.

Definition: Two events E and F are independent if both E and F have positive probability and if

$$P(E \mid F) = P(E) \quad and \quad P(F \mid E) = P(F)$$

Note that: If $P(E) > 0 \ \ and \ P(F) > 0,$ then E and F are independent if and only if

$$P(E \cap F) = P(E)P(F)$$

Example 5.14: Suppose that we roll a pair of fail dice, so each of the 36 possible out come is equally likely. Let A denotes the event that the first die lands on 3, let C be the event that the sum of the dice is 7

A. Are A and B independent?

B. Are A and C independent

Solution:

A. Since $A \cap B$ is the event that the first die lands on 3 and the second on 5, we see that

$$P(A \cap B) = P(\{(3,5)\}) = \frac{1}{36}$$

On the other hand

$$P(A) = P(\{(3,1),(3,2),(3,3),(3,4),(3,4),(3,6)\}) = \frac{6}{36} \quad \text{and}$$

$$P(B) = P(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = {}^{5}\!/_{36}$$

Therefore, since ${}^{1}\!/_{36} \neq ({}^{6}\!/_{36}).({}^{5}\!/_{36}),$ we see that $P(A \cap B) \neq P(A)P(B)$ and so events A and B are not independent

B. Events A and C are independent. This is seen by noting that

$$P(A \cap C) = P(\{3,4\}) = \frac{1}{36}$$

While $P(A) = \dfrac{1}{6}$ and $P(C) = P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \dfrac{6}{36}$ .Therefore, $P(A \cap C) = P(A).P(C)$ and so events A and C are independent.

CHAPTER FOUR

Probability Distribution

Before probability distribution is defined formally, the definition of reviewed. In the first chapter, a variable was defined as a characteristic or attribute that can assume different values various letter of the alphabet are used to represent the variables.

At the end of this chapter students are expected to:

- ✓ Know what meant by random variable, probability distribution, probability density function, expected value and variance;
- ✓ Be familiar with some standard discrete and continuous probability distributions;
- ✓ Be able to use standard statistical tables for Normal, t, Chi-square distributions.

4.1 The Definition of Random Variable and Probability Distribution

Definition: Let S be a sample space of an experiment and X is a real valued function defined over the sample space S, then X is called a random variable (or *stochastic variable*).

A random variable, usually shortened to r.v. (rv), is a function defined on a sample space S and taking values in the real line $\Re$, and denoted by capital letters, such as X, Y, Z. Thus, the value of the r.v. X at the sample point s is X(s), and the set of all values of X, that is, the range of X, is usually denoted by X(S) or R_X.

The difference between a r.v. and a function is that, the domain of a r.v. is a sample space S, unlike the usual concept of a function, whose domain is a subset of $\Re$ or of a Euclidean space of higher dimension. The usage of the term "random variable" employed here rather than that of a function may be explained by the fact that a r.v is associated with the outcomes of a random experiment. Of course, on the same sample space, one may define many distinct r.vs.

Example4 .1: Suppose we are about to learn the sexes of the three children of a certain family. The sample space of this experiment consists of the following 8 outcomes.

$$S = \{(b,b,b),(b,b,g),(b,g,b),(b,g,g,),(g,b,b),(g,b,g),(g,g,b),(g,g,g)\}$$

The outcomes (g,b,b) means, for instance that the youngest child is a girl, the next youngest is a boy, and the oldest is a boy. Suppose that each of these 8 possible outcomes is equally likely, and so each has probability 1/8. If we let X denote the number of female children in this family, then the value of X is determined by the outcomes of the experiment. That is, X is a random variable whose value will be $0, 1, 2\ or\ 3$. i.e.

$$X(bbb) = 0, X(gbb) = X(bgb) = X(bbg) = 1,$$
$$X(ggb) = X(gbg) = X(bgg) = 2, X(ggg) = 3$$

Example 4.2: Recording the lifetime of an electronic device, or of an electrical appliance. Here S is the interval (0, T) or for some justifiable reasons, S = (0, ∞), a r.v. X of interest is X(s) = s, s ∈ S.

Example 4.3: Measuring the dosage of a certain medication administered to a patient, until a positive reaction is observed. Here S = (0, D) for some suitable D.

In the examples discussed above we have seen r.v.s with different values. Hence, random variables can be categorized in to two broad categories such as discrete and continuous random variables.

4.1.1 Discrete Random Variable and Probability Distribution (pmf)

Definition 6.2:A random variable X is called discrete (or of the discrete type), if X takes on a finite or countably infinite number of values; that is, either finitely many values such as $x_1, \ldots,$ x_n, or countably infinite many values such as $x_0, x_1, x_2, \ldots$.

Or we can describe discrete random variable as, it

- Take whole numbers (like 0, 1, 2, 3 etc.)
- Take finite or countably infinite number of values
- Jump from one value to the next and cannot take any values in between.

Example4.3:

Experiment	**Random Variable (X)**	**Variable values**
Children of one gender in a family	Number of girls	0, 1, 2, …
Answer 23 questions of an exam	Number of correct	0, 1, 2, ..., 23
Count cars at toll between 11:00 am & 1:00 pm	Number of cars arriving	0, 1, 2, ..., *n*

Definition: If X is a discrete random variable, the function given by *f(x)* = $p(X = x) or\ P\{X = x_i\}$ for each x within the range of ***X*** is said to be *probability distribution* or *probability mass function* of ***X*** if it satisfies the following two conditions:

1. The sum of the probabilities of all the events in the sample space must equal 1; that is, $\sum p(x) = 1$
2. The probability of each event in the sample space must be between or equal to 0 and 1. that is, $0 \leq p(x) \leq 1$.

Example 4.4: Consider r.v. X in *Example 6.1* and construct probability distribution of X.

Solution: Since X will equal 0 if the outcome is (b, b, b), we see that

$$P(X = 0) = P(bbb) = \frac{1}{8}$$

Since X will equal 1 if the outcome is $(gbb), (bgb), (bbg)$ we have $P(X = 1) = P\{(gbb) \; or \; (bgb) \; or \; (bbg)\} = \frac{3}{8}$

Similarly, $P(X = 2) = P\{(bgg) \; or \; (gbg) \; or \; (ggb)\} = \frac{3}{8}$, $P(X = 3) = P(ggg) = \frac{1}{8}$

Therefore,

$$f(x) = \begin{cases} \frac{1}{8} \; if \; x = 0, 3 \\ \frac{3}{8} \; if \; x = 1,2 \\ 0 \; otherwise \end{cases}$$

Example 4.5: Suppose we toss a coin three times, the sample space is represented as $TTT, TTH, THT, HTT, HHT, HTH, THH, HHH$ and if the random variable for the number of heads.

A. Assign a value for a random variable
B. Find the probability distribution for A

Solution:

A. Once a random variable, say X, is defined as the number of heads, $X = 0,1,2, or 3$
B.

Number of heads X	0	1	2	3
Probability $p(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

We can check that $\sum p(X = x) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

Example 6.7: Suppose that X is a random variable that takes on one of the value $0,1,2, or 3$. If $P\{X = 1\} = 0.4$ and $P\{X = 2\} = 0.1$. What is $P\{X = 1\}$?

Solution: Since the probability must sum to 1, we have

$$
\begin{aligned}
1 &= P\{X=1\}+P\{X=2\}+P\{X=3\} \\
1 &= 0.4+0.1+P\{X=3\} \\
P\{X=3\} &= 1-0.4-0.1 \\
&= 0.5
\end{aligned}
$$

Example 4.6: A sales women has scheduled two appointments to sell encyclopedias. She feels her first appointments will lead to a sale with probability 0.3. She also feels that the second will lead to a sale with probability 0.6 and that the results from the two appointments are independent. What is the probability distribution of X, the number of sales made?

Solution: The random variable X can take on any of the value $0,1,2$. It will equal 0 if neither appointment leads to a sale, and so

$$
\begin{aligned}
P\{X=0\} &= P\{no\, sale on\ first, no\, sale on\, \mathrm{se}cond\} \\
&= P\{no\, sale on\ first\}P\{no\, sale on\, \mathrm{se}cond\} \\
&= (1-0.3)(1-0.6) \\
&= 0.28
\end{aligned}
$$

The random variable X will equal 1 either if there is a sale on the first and not on the second appointment or if there is no sale on the first and one sale on the second appointment. Since these two events are disjoint, we have

$$
\begin{aligned}
P\{X=1\} &= P\{Sale on\ first, no\, sale on\, \mathrm{se}cond\}+P\{No\ sale on\ first, sale on\, \mathrm{se}cond\} \\
&= P\{Sale on\ first\}P\{no\, sale on\, \mathrm{se}cond\}+\{no\, sale on\ first\}P\{no\, sale on\, \mathrm{se}cond\} \\
&= 0.3(1-0.6)+0.6(1-0.3) \\
&= 0.54
\end{aligned}
$$

Finally, the random variable X will equal 2 if both appointments result in sales; thus

$$
\begin{aligned}
P\{X=2\} &= P\{sale on\ first, sale on\, \mathrm{se}cond\} \\
&= P\{sale on\ first\}P\{Sale on\, \mathrm{se}cond\} \\
&= 0.3x0.6 \\
&= 0.18
\end{aligned}
$$

As check on this result, we note that

$$P\{X=0\}+P\{X=1\}+P\{X=2\}=0.28+0.54+0.18=1$$

Exercise 4.1: Check whether the function given by $f(x)=\frac{x+2}{25}$ for x = 1, 2, 3, 4, 5 is a *p.m.f*?

Definition: If X is a discrete random variable, the function given by

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad \text{for all x in } \Re \text{ and } t \in X.$$

Where *f(t)* is the value of probability distribution or *p.m.f* of X at *t*, is called the distribution function, or the cumulative distribution function of X.

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases}$$

Example 4.7

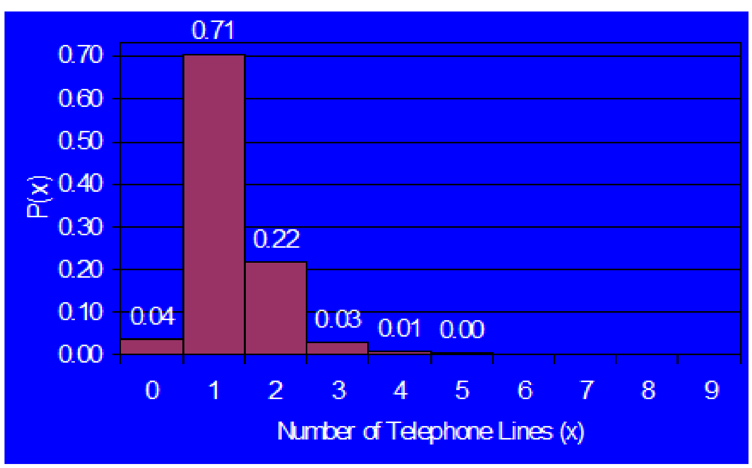
Find the distribution function F of the total number of heads obtained in four tosses of a balanced coin?

The distribution function, or the cumulative distribution function *F(X)* will be the following;

$$F(X) = \begin{cases} 0 \quad for \quad x < 0 \\ \frac{1}{16} \; for \; 0 \leq x < 1 \\ \frac{5}{16} \; for \; 1 \leq x < 2 \\ \frac{11}{16} \; for \; 2 \leq x < 3 \\ \frac{15}{16} \; for \; 3 \leq x < 4 \\ 1 \quad for \quad x \geq 4 \end{cases}$$

Exercise 4.2: A telephone survey of households throughout Washington State is given below:

Number of telephones,x	P(x)
0	0.035
1	0.70553
2	0.21769
3	0.02966
4	0.00775
5	0.00332
6	0.00088
7	0.00002
8	0
9	0.00015
Total	1



a. What is the probability that a household will have no telephone?

b. What is the probability that a household will have 2 or more telephone lines?

c. What is the probability that a household will have 2 to 4 phone lines?

d. What is the probability a household will have no phone lines or more than 4 phone lines?

e. Who do you think is in that 3.5% of the population?

4.1.2 Continuous Random Variable and Probability Distribution

Definition : A r.v X is called continuous (or of the continuous type) if X takes all values in a proper interval $I \subseteq \Re$.

Or we can describe continuous random variables as follows:

- Take whole or fractional number.
- Obtained by measuring.
- Take infinite number of values in an interval.
- Too many to list like discrete variable

Example 4.9:

The following examples are continuous r.v.s

Experiment	Random Variable X	Variable values
Weigh 100 People	Weight	45.1, 78, ...
Measure Part Life	Hours	900, 875.9, …
Ask Food Spending	Spending	54.12, 42, ...
Measure Time Between Arrivals	Inter-Arrival time	0, 1.3, 2.78, ...

Definition 4.4: A function with values *f(x)*, defined over the set of all real numbers, is called a probability density function of the continuous random variable X if and only if

$$P\ (a \leq x \leq b) = \int_a^b f(x)dx \quad \text{for any real constant } a \leq b.$$

Probability density function also referred as probability densities (*p.d.f.*), probability function, or simply densities.

Remark:

- The probability density function *f (x)* of the continuous random variable X, has the following properties (satisfy the conditions)

  1. $f(x) \geq 0$ for all x, or for $-\infty < x < \infty$

2. $f(x) = \int_{-\infty}^{\infty} f(x)\,dx = 1$

- If X is a continuous random variable and *a* and *b* are real constants with $a \leq b$, then

$$P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b) = P(a < x < b)$$

Example 6.11: If X is the probability density

$$f(x) = \begin{cases} k.e^{-3x} & for\ x > 0 \\ 0 & elsewhere \end{cases}$$

Find the constant k and $P(0.5 \leq X \leq 1)$?

Solution:

$\int_{\mathbf{0}}^{\infty} f(x)dx = 1$**,** since $f(x)$ is pdf.

$$\int_{\mathbf{0}}^{\infty} k.e^{-3x}\,dx = 1 \Rightarrow -\frac{k}{3}\left[\left(\lim_{x\to\infty} e^{-3x}\right) - 1\right] = 1 \Rightarrow \frac{k}{3} = 1 \Rightarrow k = 3$$

And $P(0.5 \leq X \leq 1) = \int_{\mathbf{0.5}}^{\mathbf{1}} 3.e^{-3x}\,dx = [-e^{-3x}]_{0.5}^{1} = \boldsymbol{-e^{-3} + e^{-1.5}} = \frac{\mathbf{1}}{\boldsymbol{e^{1.5}}} - \frac{\mathbf{1}}{\boldsymbol{e^{3}}}$

Exercise 6.3:

The p.d.f of the random variable X is given by

$$f(x) = \begin{cases} \frac{C}{\sqrt{x}} & for\ 0 < x < 4 \\ 0 & elsewhere \end{cases}$$

Find a. the value of C?

b. $P(X < \frac{1}{4})$ and $P(X > 1)$?

Definition: If X is a continuous random variable and the value of its probability density is *f (t),* then function given by $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$ is called the distribution function, or the cumulative distribution of the continuous r.v. X.

Theorem 4.1: If *f (x)* and *F(x)* are the values of the probability density and the distribution function of X at *x*, then

$$P(a \leq x \leq b) = F(b) - F(a)$$

For any real constant a and b with $a \leq b$, and

$$f(x) = \frac{d\,F(x)}{dx}$$

where the derivative exist.

Exercise 4.4: Find the distribution function of the random variable X and evaluate *P (0.5 ≤ X ≤ 1)*? If is the probability density of X is *f (x)*, $f(x) = \begin{cases} 3e^{-3x} & for\ x > 0 \\ 0 & elsewhere \end{cases}$

Exercise 4.5:

A r.v. X has d.f. F given by:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 2c\left(x^2 - \frac{1}{3}x^3\right), & 0 < x \leq 2 \\ 1, & x > 2. \end{cases}$$

(i) Determine the corresponding *p.d.f.* (*f(x)*).

(ii) Determine the constant c.

6.2 Introduction to Expectation- Mean and Variance of a Random Variable

A key concept in probability is the expected value of a random variable.

Definition : If X is a discrete random variable that takes on one of the possible values $x_1, x_2, \cdots x_n$ then the expected value of X, denoted by $E(X) or E[X]$, is defined by

$$E(X) = \sum_{i=1}^{n} x_i p(x_i)$$

If X is continuous random variable

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Where $f(x)$ is probability density function in the case of discrete random variable its name will change to *probability mass function (pmf).*

Example 6.12: Find the expected value of the following random variable

X	0	1	2	3	4
$\boldsymbol{P(X)}$	0.18	0.34	0.23	0.21	0.04

Solution: $E(X) = \sum_{x=0}^{4} xP(X = x)$

$$= 0(0.18) + 1(0.34) + 2(0.23) + 3(0.21) + 4(0.04)$$

$$= 1.14$$

Note that: The expected value of a random variable is the same as with the mean of a random variable

$$\bar{X} = E(X) = \begin{cases} \sum xP(x)\,, & if\ X\ is\ discrete \\ \int xf(x)\,dx, & if\ X\ is\ continous. \end{cases}$$

Suppose that we are given random variable random variable along with its probability mass function (pmf) if it is discrete or probability density function (pdf) if it is continuous, and that we want to compute the expected value of some function of X, say $g(X)$. How can we accomplish this? One way is follows: Since $g(x)$ is determined from the pmf/pdf of X. Once we have determined the pmf/pdf of $g(x)$ we can compute $E\,[g(x)]$ by using the definition of expected value.

$$E[g(x)] = \begin{cases} \sum g(x)P(x)\,, & if\ X\ is\ discrete \\ \int g(x)f(x)dx\,, & if\ X\ is\ continous. \end{cases}$$

Example 4.10: Let X denote a random variable that takes on any of the values -1, 0, 1 with respective probability $P\{X = -1\} = 0.2$,$P\{X = 0\} = 0.5$,$p\{\,X = 1\} = 0.3$, then compute $E(X^2)$.

Solution: Letting $Y = g(x)\ =\ X^2$,

$$E[g(x)] = \sum g(x)P(x) = (-1)^2.P(X = -1) + 0^2.P(X = 0) + 1^2.P(X = 1)$$

$$= 1(0.2) + 0(0.5) + 1(0.3) = 0.5$$

The reader should note that $(E\,[X])^2\ =\ 0.01$

$$0.5\ =\ E\,[X^2]\ \neq\ (E[X])^2\ =\ 0.01$$

If a and b are constants then

$$E(aX + b)\ =\ aE[X]\ +\ b$$

The expected value of a random variable X, $E\,[X]$ is also referred to as the mean or the first moment of X. The quantity $E[X^n], n\ \geq\ 1$, is called the n^{th} moment of X. By definition

$$E[X^n] = \begin{cases} \sum X^n P(X = x), if\ X\ is\ discrete \\ \int X^n f(x)dx, if\ X\ is\ continous. \end{cases}$$

Exercise 4.6: The following are the annual income of 7 men and 7 women residents of a certain community.

Annual income (in $ 1000)

Men	Women
33.5	24.2
25.0	19.5
28.6	27.4
41.0	28.6
30.5	32.2
29.6	22.4
32.8	21.6

Suppose that a woman and a man randomly chosen. Find the expected value of the sum of their incomes.

Solution: Let X be the man's income and Y is the woman's income. Since X is equally likely to be any of the values in the men's column, we see that

$$E\,(X) = \frac{1}{7}\,(33.5\ +\ 25\ + \cdots +\ 32.8) =\ 31.571$$

Similarly, $E\,[\,Y]\ =\ \frac{1}{7}\,(24.2\ +\ 19.5\ + \cdots +\ 21.6) =\ 25.129$

Therefore, the expected value of the sum of their incomes is

$$E\,[X + Y] = E\,[X] + E\,[Y]\ \ = 56.7$$

That is, the expected value of the sum of their incomes is approximately $ 56,700.

Definition: If X is a random variable with mean μ, then the variance of x, denoted by $Var\ (x)$, is defined by $Var\ (X)\ =\ E\ [(X - \mu)^2]$

An alternative formula for $Var(X)$ is derived as follows

$$Var(X) = E[X - \mu]^2$$

$$= \sum (X - \mu)^2 P(X)$$

$$= \sum (X^2 - 2\mu X + \mu^2) P(X)$$

$$= E[x2] - 2\mu^2 + \mu^2$$

$$= E(X^2) - \mu^2$$

That is, $Var(X) = E[X^2] - (E[X])^2$

Example 4.14: The return from a certain investment is a random variable X with probability distribution. $P\{X = -1\} = 0.7,\ P\{X = 4\} = 0.2,\ P\{X = 8\} = 0.1$

Find $Var(X)$, the variance of the return.

Solution: Let us first compute that expected return as follows:

$$\mu = E(X) = -1(0.7) + 4(0.2) + 8(0.1) = 0.9$$

To compute $Var(X)$, we use the formula $Var(X) = E(X^2) - \mu^2$

Now, since X^2 will equal $(-1)^2, 4^2, or\ 8^2$ with respective probabilities of $0.7, 0.2, and\ 0.1$, we have

$$E[X^2] = 1(0.7) + 16(0.2) + 64(0.1) = 10.3$$

Therefore, $Var(X) = 10.3 - (0.9)^2 = 9.94$

Properties of Variance

1. For any random variance X and constant C, it can be shown that
   - $Var(CX) = C^2 Var(X)$
   - $Var(C + X) = Var(X)$
2. If X and Y are independent random variable, $Var(X + Y) = Var(X) + Var(Y)$
3. The square root of the $Var(X)$ is called the standard deviation of X, and we denote it by $D(X)$. That is, $SD(x) = \sqrt{Var(x)}$

6.3 Common Discrete Probability Distribution

6.3.1. Binomial Distribution

Many types of probability problems have only two outcomes, or they can be reduced to two outcomes. For example, when a coin is tossed, it can land heads or tails.

A probability experiment is a binomial probability experiment that satisfies the following four requirements:

1. Each trial can have only two outcomes or outcomes that can be reduced to two outcomes.
2. There must be a fixed number of trials
3. The outcomes of each trial must be independent
4. The probability of a success must remain the same for each trial

The outcomes of a binomial experiment and the corresponding probabilities of these outcomes are called a binomial distribution. The probability mass function of a binomial random variable having parameter (n, p) is given by

$$P(X = x) = \binom{n}{x} P^x (1 - P)^{n-x}, \quad i = 0, 1, \ldots, n$$

Example 4.11: Five fair coins are flipped. If the outcomes are assumed independent, find the probability of the number of heads obtained

Solution: If we let X equal the number of heads (successes) parameters$(n = 5, P = ½)$. Hence,

$$P\{X = 3\} = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 \equiv \frac{10}{32}$$

$$P\{X = 4\} = \binom{5}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right) \equiv \frac{5}{32}$$

$$P\{X = 5\} = \binom{5}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right) \equiv \frac{1}{32}$$

Example 4.12:

A. Determine $P\{X \le 12\}$ when X is a Binomial random variable with parameters $n = 20$ and $P = 0.4$

B. Determine $P\{Y \le 10\}$ when Y is a Binomial random variable with parameters $n = 16$and $P = 0.5$

Solution:

A. $P\{X \le 12\} = 1 - P\{X > 12\}$

$$= 1 - P\{X = 13\} - P\{X = 14\} - \cdots - p\{X = 20\}$$

$= 0.9790$

B. $P\{Y \geq 10\} = 1 - P\{Y < 10\} = 1 - P\{Y \leq 9\} = 0.2272$

If X is Binomial random variable with parameter n and P, then

$$E(X) = np$$

$$Var(X) = np(1-p) = npq, where\ q = 1 - p$$

Example6.17: Suppose that each screw produced is independently defective with probability 0.01. Find the expected value and variance of the number of defective screws in a shipment of size 1000.

Solution: The number of effective screws in the shipment of size 1000 is a Binomial random variable with parameters $n = 1000$, $P = 0.01$. Hence, the expected number of defective screws is $E[number\ of\ defectives] = 1000(0.01) = 10$ and the variance of the number of detective screws is $Var(number\ of\ defective) = 1000(0.01)(0.99) = 9.9$

643.2 The Poisson Distribution

A discrete probability distribution that is useful when n is large and p is small and when the independent variable occurs over a period of time is called the Poisson distribution, name for Simeon D. Poisson (1781-1840). In addition to being used for the stated conditions (i.e. n is large, p is small, and the variable occur over a period of time), the Poisson distribution can be used when a density of items is distributed over a given area or volume, such as the number of plants growing per acre of woods or the number of defects in a given length of videotape.

If X is Poisson random variable with parameter λ, then

$$P(x) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Example6.18: If X is a Poisson random variable with parameter $\lambda = 2$, find $P(X = 0)$

Solution: $P(X = 0) = \frac{e^{-2}2^{0}}{0!}$**,** Using the fact that $2^0 = 1$, $0! = 1$, we obtain $P(X = 0) = e^{-2} = 0.135$

Both the expected value and the variance of a Poisson random variable are equal to λ. That is, we have the following. If X is a Poisson random variable with parameter λ, $\lambda > 0$; then

$$E[X] = \lambda, \quad Var[X] = \lambda$$

Example6.19: Suppose the average number of accidents occurring weekly on a particular high way is equal to 1.2. Approximate the probability that there is at least one accident this week.

Solution: Let x denote the number of accidents because it is reasonable to suppose that there are a large number of cars passing along the high way, each having a small probability of being involved in an accident, the number of such accidents should be approximately a Poisson random variable. That is, if x denotes the number of accidents that will occur this week, then x is approximately Poisson random variable with mean value $\lambda = 1.2$. The desired probability is now obtained as follows.

$$p\{x > 0\} = 1 - p\{x = 0\} = 1 - \frac{e^{-1.2}(1.2)^0}{0!} = 1 - e^{-1.2} = 0.6988$$

Therefore, there is approximately a 70% chance that there will be at least one accident this week.

We can approximate Binomial distribution to Poisson distribution if n is large and p is too small. Thus, the approximately Poisson distribution has a parameter.

$$\lambda = np$$

Example 6.20: Suppose that items produced by a certain machine are independently defective with probability 0.1.What is the Poisson approximation for this probability?

Solution: If we let x denote the number of defective items, then x is a Binomial random variable with parameters $n = 10$ and $P = 0.1$. Thus the desired probability is

$$p\{X = 0\} + P\{X = 1\} = \binom{10}{0}(0.1)^0 (0.9)^{10} + \binom{10}{1}(0.1)^1 (0.9)^9$$

$$= 0.7361$$

Since $nP = 10(0.1) = 1$, the Poisson approximation yields the value.

$$P\{x = 0\} + P\{x = 1\} = e^{-1} + e^{-1} = 0.7358$$

Thus, even in this case, where n is equal to 10 (which is not that large) and p is equal to 0.1 (which is not that small), the Poisson approximation to the Binomial probability is quite accurate.

6.4 Common Continuous Probability Distribution

Every continuous random variable X has a curve associated with it. This curve, formally known as a *probability density function*, can be used to obtain probabilities associated with the random variable. This is accomplished as follows, consider any two points a and b, where a is less than

b. The probability that x assumes a value that lies between a and b is equal to the area under the curve between a and b. That is,

$$P\{a \le x \le b\} = Area\ under\ curve\ between\ a\ and\ b$$

Since X must assume some value, it follows that the total area under the density curve must equal 1. Also, since the area under the graph of the probability density function between points a and b is the same regardless of whether the end points a and b are themselves included.

That is, $P\{a \le x \le b\} = P\{a < x < b\}$

6.4.1 Normal Random Variables

The most important type of random variable is the normal random variable. The probability density function of a normal random variable X is determined by two parameters: the *expected value* and the *standard deviation* of X. We designate these values as μ and σ, respectively.

$$\mu = E[X] \text{ And } \sigma = SD(X)$$

The normal probability density function is a bell-shaped density curve that is symmetric about the value μ; its variability is measured by σ. The larger σ is, the more variability there is in the curve.

Since the probability density function of a normal random variable X is symmetric about its expected value μ; it follows that X is equally likely to be on either side of μ. That is,

$$P\{X < \mu\} = P\{X \geq \mu\} = 0.5$$

Not all bell-shaped symmetric density curves are normal. The normal density curves are specified by a particular formula:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

A normal random variable having mean value 0 and standard deviation 1 is called a standard normal variable, and its density curve is called the standard normal curve. The letter Z represents a standard normal random variable.

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x)^2}{2}}$$

Probabilities Associated with a Standard Normal Random Variable

$$Z = \frac{value - mean}{standard\ deviation} = \frac{X - \mu}{\sigma}$$

Once the X values are transformed by using the above formula, they are called Z value is actually the number of standard deviations that a particular X value is a way from the mean.

Steps to find areas under the normal distribution curve

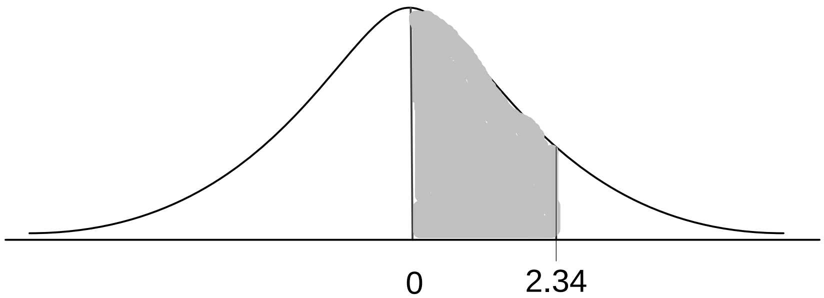
1. Between 0 and any Z value: Look up the Z value in the table to get the area
2. In any tail
   a. Look up the Z value to get the area
   b. Subtract the area from 0.5
3. Between Z values on the same side of the mean
   a. Look up both Z values to get the area
   b. Subtract the smaller area from the larger area
4. Between two Z values on opposite sides of the mean
   a. Look up both Z values to get the area
   b. Add the areas
5. Less than any Z value to get the right of the mean
   a. Look up the Z value to get the area
   b. Add 0.5 to the area
6. Greater than any Z value to the left of the mean
   a. Look up the Z value in the table to get the area
   b. Add 0.5 to the area
7. In any two tailed
   a. Look up Z values in the table to get the areas
   b. Subtract both areas from 0.5
   c. Add the answer

General procedure is

- Draw the picture
- Shade the area desired
- Find the correct figure
- Follow the direction

Example 6.15: Find the area under the normal distribution curve between $Z = 0$ and $Z = 2.34$

Solution: Draw the area as follows:



Since Z table gives the area between 0 and any Z value to the right of 0, one need look up the Z value in the table. Find 2.3 in the left column and 0.04 in the top row. The value where the column and row meet in the table is the answer, 0.4904.

Z	0.00	0.01	0.02	0.03	0.04	…
0.0						
0.1						
0.2						
⋮						
2.2						
2.3					0.4904	
⋮						

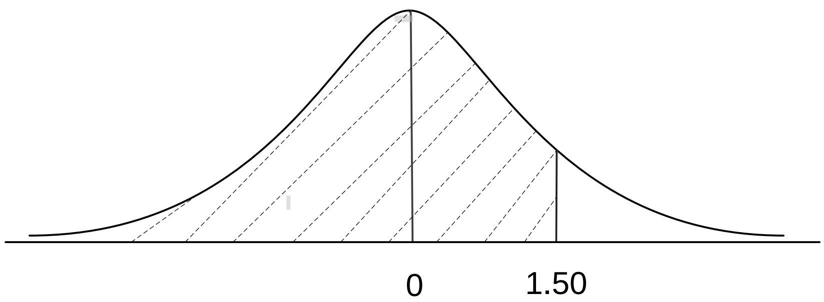
Example 6.16: Find

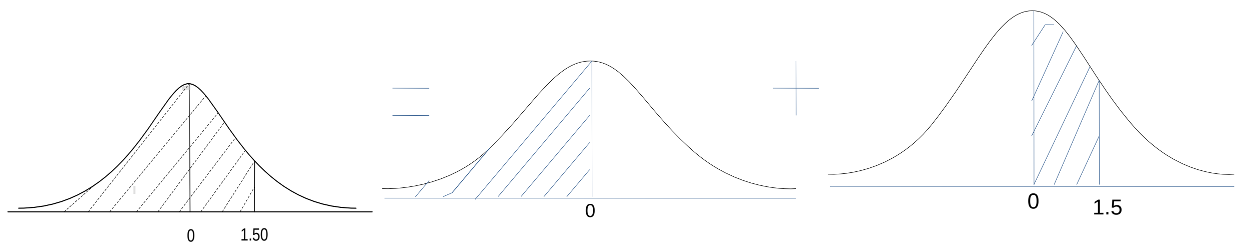
A. $P\{Z < 1.5\}$

B. $P\{Z \geq 0.8\}$

Solution:

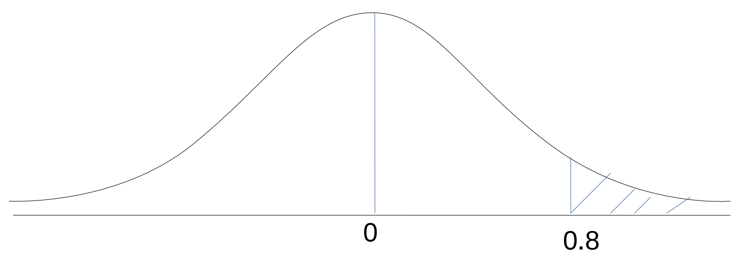
A. Draw the area as follows:

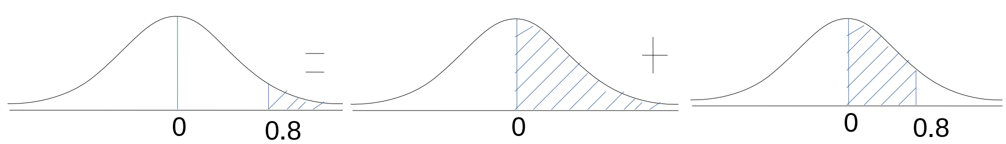




$$P\{Z < 1.5\} = 0.5 + P\{0 < Z < 1.5\} = 0.5 + 0.4332 = 0.9332$$

B. Draw the required area as follows:





$$P\{Z \geq 0.8\} = 0.5 - P\{0 < Z < 0.8\} = 0.5 - 0.2881 = 0.2119$$

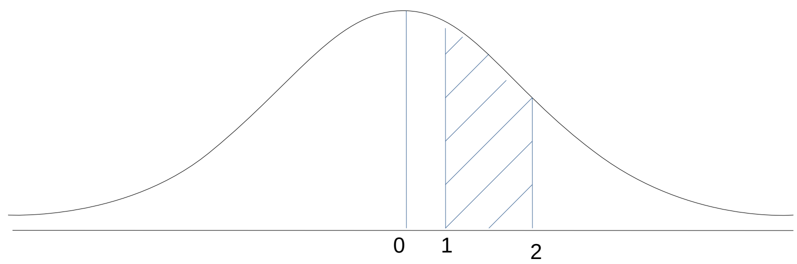
Example 6.17: Find

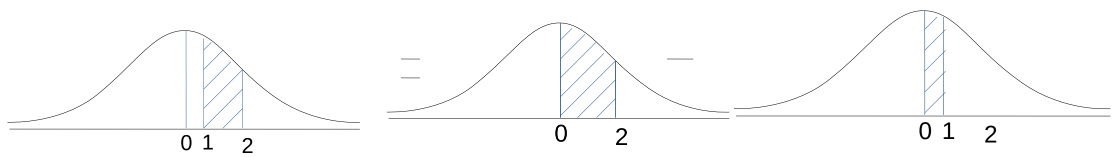
A. $P\{1 < Z < 2\}$

B. $P\{-1.5 < Z < 2.5\}$

Solution:

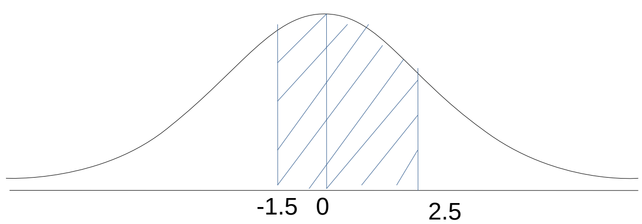
A. Draw the graph as follows:

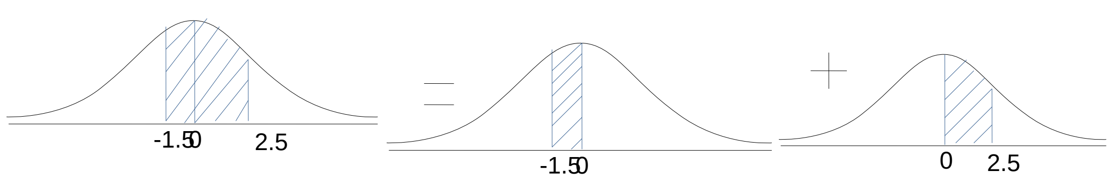




$$P\{1 < Z < 2\} = P\{0 < Z < 2\} - P\{0 < Z < 1\} = 0.4772 - 0.3159 = 0.1359$$

B. Draw the graph as follows:





$$P\{-1.5 < Z < 2.5\} = P\{-1.5 < Z < 0\} + P\{0 < Z < 2.5\}$$

Since $P\{-1.5 < Z < 0\} = P\{0 < Z < 1.5\}$, due to symmetric property of normal distribution

$$P\{-1.5 < Z < 2.5\} = P\{0 < Z < 1.5\} + P\{0 < Z < 2.5\} = 0.4332 + 0.4938 = 0.9270$$

Finding Normal Probabilities: Conversion to the Standard Normal

Let Xbe a normal random variable with mean μ and standard deviation σ. We can determine probabilities concerning X by using the fact that the variable Z defined by

$$Z = \frac{X - \mu}{\sigma}$$

has a standard normal distribution.

We can compute any probability statements in terms of Z. For example,

$$P\{X < a\} = P\left\{\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right\} = P\left\{Z < \frac{X-\mu}{\sigma}\right\}$$

where Z is a standard normal random variable

Example 6.18: IQ examination scores for sixth-graders are normally distributed with mean value 100 and standard deviation 14.2.

A. What is the probability a randomly chosen sixth-grader has a score greater than 130?

B. What is the probability a randomly chosen sixth-grader has score between 90 and 115?

Solution: Let X denote the score of a randomly chosen student. We compute probabilities concerning X by making use of the fact that the standardized variable

$$Z = \frac{X-100}{14.2}$$

has a standard normal distribution

A. $$P\{X > 130\} = P\left\{\frac{X-100}{14.2} > \frac{130-100}{14.2}\right\} = P\{Z > 2.1127\} = 0.0170$$

B. The inequality $90 < X < 115$ is equivalent to

$$\frac{90-100}{14.2} < \frac{X-100}{14.2} < \frac{115-100}{14.2}$$

Or equivalently,

$$-0.7042 < Z < 1.0560$$

Therefore,

$$P\{90 < X < 115\} = P\{-0.7042 < Z < 1.0560\}$$

$$= P\{0 < Z < 0.7042\} + P\{0 < Z < 1.0560\} = 0.6120$$

Properties of the Normal distribution

1. The normal distribution curve is bell-shaped
2. The mean, median and mode are equal and located at the center of the distribution
3. The normal distribution curve is unimodal
4. The curve is symmetrical about the mean, which is equivalent to saying that is shape the same on both sides of vertical line passing through the center

5. The curve is continuous. That is, no gaps or holes
6. The curve never touches the x axis
7. The total area under the normal distribution curve is equal to 1

CHAPPTER FIVE

5 Sampling and sampling Technique

Introduction:

Sampling in statistics is as common and important as sugar in preparing any sweet dish in all the spheres of life such as economical, social, business, agricultural, health science and Biological science the need for statistical investigation and data analysis is rising day by day. Most of us use sampling in our daily life for example, when we go to buy provisions from a grocery, we might sample a few grains of rice or wheat to inner the quality of the population. This could be the average life of a florescent tube, the percentage of customers of Talcum powder who prefer the new improved Talcum powder to the old one or the percentage of time a machine is being used.

In this unit we begin our study of sampling. A sampling is a statistical tool that allows as to take a sample from a population to infer something about population characteristics. Next, we construct a distribution of the sample means to understand how the sample means to tend to cluster around the population mean and using the concept of central limit theorem, we will see that the shape of this distribution tends to follow the normal distribution for a large sample size.

Objectives

After completing this unit you are expected to be able to

- Define what sample is
- Define the basic terms of sampling
- Distinguish the difference between sample survey and census
- Define the concept of sampling distribution of sample means and it's properties
- Define the concept of central limit theorem

7.1 The concept of sampling

Definition of some basic terms:

Sampling: The process or method of sample selection from the population.

Sampling unit: The ultimate unit to be sampled or elements of the population to be sampled.

Examples:

- ❖ If somebody studies Scio-economic status of the households, households is the sampling unit.
- ❖ If one studies performance of freshman students in some college, the student is the sampling unit.

Sampling frame: is the list of all elements in a population.

Examples:

- ❖ List of households.
- ❖ List of students in the registrar office.

Reasons for Sampling

- ❖ Reduced cost
- ❖ Greater speed
- ❖ Greater accuracy
- ❖ Greater scope
- ❖ Avoids destructive test
- ❖ The only option when the population is infinite

Because of the above consideration, in practice we take sample and make conclusion about the population values such as population mean and population variance, known as parameters of the population.

Sometimes taking a census makes more sense than using a sample. Some of the reasons include:

- ❖ Universality
- ❖ Qualitativeness
- ❖ Detailedness
- ❖ Non-representativeness

Sampling Techniques

Sampling technique can be grouped in to two categories. This are

1. Random (probability) sampling method and
2. Non random sampling (non probability) sampling methods.

A. **Random Sampling or probability sampling.**

- Is a method of sampling in which all elements in the population have a pre-assigned none zero probability to be included in to the sample. The basis of all random sampling is the simple random sample. Other important methods include Stratified sampling, cluster sampling and systematic sampling.

Examples:

• Simple random sampling

• Stratified random sampling

• Cluster sampling

• Systematic sampling

1. Simple Random Sampling*:*

- Is a method of selecting items from a population such that every possible sample of specific size has an equal chance of being selected? In this case, sampling may be with or without replacement. Or

- All elements in the population have the same pre-assigned non zero probability to be included in to the sample.

- Simple random sampling can be done either using the lottery method or table of random numbers.

Table of Random Numbers

Table of random numbers are tables of the digits 0, 1, 2,…,, 9, each digit having an equal chance of selection at any draw. For convenience, the numbers are put in blocks of five. In using these tables to select a simple random sample, the steps are:

i Number the units in the population from 1 to N (prepare frame of the population).

ii. Then proceed in the following way

If the first digit of N is a number between 5 and 9 inclusively, the following method of selection is adequate. Suppose N=528 and we want n=10.

Select three columns from the table of random numbers, say columns 25 to 27. Go down the three columns selecting the first 10 distinct numbers between 001 & 528. These are 36, 509, 364, 417, 348, 127, 149, 186, 439, and 329. Then the units with these roll numbers are our samples.

Note: If sampling is without replacement, reject all the numbers that comes more than once.

2.* Stratified Random Sampling*:

- The population will be divided in to none overlapping but exhaustive groups called strata.
- Simple random samples will be chosen from each stratum.
- Elements in the same strata should be more or less homogeneous while different in different strata.
- It is applied if the population is heterogeneous.
- Some of the criteria for dividing a population into strata are: Sex (male, female); Age (under 18, 18 to 28, and 29 to 39); Occupation (blue-collar, professional, and other).

3. Cluster Sampling*:*

- The population is divided in to non overlapping groups called clusters.
- A simple random sample of groups or cluster of elements is chosen and all the sampling units in the selected clusters will be surveyed.
- Clusters are formed in a way that elements within a cluster are heterogeneous, i.e. observations in each cluster should be more or less dissimilar.
- Cluster sampling is useful when it is difficult or costly to generate a simple random sample. For example, to estimate the average annual household income in a large city we use cluster sampling, because to use simple random sampling we need a complete list of households in the city from which to sample. To use stratified random sampling, we would again need the list of households. A less expensive way is to let each block within the city represent a cluster. A sample of clusters could then be randomly selected, and every household within these clusters could be interviewed to find the average annual household income.

4. Systematic Sampling*:*

- A complete list of all elements within the population (sampling frame) is required.
- The procedure starts in determining the first element to be included in the sample.
- Then the technique is to take the k^{th} item from the sampling frame.
- let N = population size , n = sample size, K= $\frac{N}{n}$ =sampling interval
- Chose any number between 1 and *k*. Suppose it is $j(1 \leq j \leq K)$
- The j^{th} unit is selected at first and then the $(j + k)^{th}$,$(j + 2k)^{th}$ …. etc until the required sample size is reached.

B. **Non Random Sampling or non probability sampling.**

- It is a sampling technique in which the choice of individuals for a sample depends on the basis of convenience, personal choice or interest.

We usually want to avoid a non random sampling but, sometimes random sampling is too costly or inconvenient to perform. We might then use methods such as Judgment Sampling, quota sampling and convenience sampling methods.

Examples:

• Judgment sampling.

• Convenience sampling

• Quota Sampling.

1. Judgment Sampling

- In this case, the person taking the sample has direct or indirect control over which items are selected for the sample.

2. Convenience Sampling

- In this method, the decision maker selects a sample from the population in a manner that is relatively easy and convenient.

3. Quota Sampling

- In this method, the decision maker requires the sample to contain a certain number of items with a given characteristic. Many political polls are, in part, quota sampling.

Note:

Let N =population size, n =sample size

1. Suppose simple random sampling is used

• We have N^npossible samples if sampling is with replacement.

• We have (Ncn)possible samples if sampling is without replacement.