Higher Technological Institute
Tenth of Ramadan City
Computer science department

STATISTICS and PROBABILITIES

Authors:
Prof: YOUSRIA ABO ELNAGA
Ass. Prof: TAREK NASSAR
Dr. : FARAHAT A. FARAHAT

http://www.hti.edu.eg

STATISTICS and PROBABILITIES

STATISTICS and PROBABILITIES

Prof: YOUSRIA ABO ELNAGA

Ass. Prof: TAREK NASSAR

Dr. : FARAHAT A. FARAHAT

Contents

Page

1. Chapter One

2. Chapter Two

Preface

It has become customary to many colleges and universities to teach undergraduate courses in statistics and probability. These courses usually emphasize the expectation of several events and treat some predicted problems.

In teaching such a course, the author has found two detrimental effects on students. Those students who are primarily interested in technical applications also get feeling that all probability distributions can be separated between discrete and continuous random variables.

This book is an attempt to present the material usually covered in such courses in the framework where the general properties of probability and a part of statistic. The first three chapters cover the probability and fourth chapter cover the correlation and regression analysis.

Chapter one present the opinions of voters concerning a new sales tax can also be considered as observations of an experiment. We are particularly interested in the observations obtained by repeating the experiment several times. In most cases, the outcomes will depend on chance and, therefore, cannot be predicted with certainty. If a chemist runs an analysis several times under the same conditions, he or she will obtain different measurements, indicating an element of chance in the experimental procedure.

Chapter two study of a random phenomenon is in the statements we can make concerning the events that can occur, and these statements are made based on probabilities assigned to simple outcomes. One of the immediate steps that can be taken in this unifying attempt is to require that each of the possible outcomes of a random experiment be represented by a real number. In this way, when the experiment is performed, each outcome is identified by its assigned real number rather than by its physical description.

Chapter three study involving testing the effectiveness of a new drug, the number of cured patients among all the patients who use the drug approximately follows binomial, hypergeometric, and Poisson distributions.

Chapter four use statistical terms we use correlation to denote association between two quantitative variables. We also assume that the association is linear, that one variable increases or decreases a fixed amount for a unit increase or decrease in the other.

Chapter 1

Probability - Sample Space

1. Introduction

Statisticians use the word experiment to describe any process that generates a set of data. A simple example of a statistical experiment is the tossing of a coin. In this experiment, there are only two possible outcomes, heads or tails.

Another experiment might be the launching of a missile and observing of its velocity at specified times. The opinions of voters concerning a new sales tax can also be considered as observations of an experiment. We are particularly interested in the observations obtained by repeating the experiment several times. In most cases, the outcomes will depend on chance and, therefore, cannot be predicted with certainty. If a chemist runs an analysis several times under the same conditions, he or she will obtain different measurements, indicating an element of chance in the experimental procedure. Even when a coin is tossed repeatedly, we cannot be certain that a given toss will result in a head. However, we know the entire set of possibilities for each toss.

2. Sample space

Consider an experiment whose outcome is not predictable with certainly in advance. However, although the outcome of the experiment will not be known in advance, let us suppose that the set of all possible outcomes of an experiment is known.

❖The set of all possible outcomes of a random experiment is called a sample space.

❖ We are going to denote the sample space by S.

Example 1:

Tossing a coin once.

S = {H, T}

Example 2:

Rolling a die once.

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

Example 3:

Tossing two coins, then the sample space consists of the following four points

S = {(H, H), (H, T), (T, H), (T, T)}

The outcome will be (H,H) if both coins are heads , (H,T) if the first coin is heads and the second tails , (T,H) if the first is tails and the second heads , and (T,T) if both coins are tails .

Example4: (Rolling two dice)

S = {(1,1), (1,2), ... , (1,6),

(2,1), (2,2), ... , (2,6),

……………………….

(6,1), (6,2), …… , (6,6)}

Or , we can write

S = { (i , j) : i , j = 1,2,3,4,5,6 }

where the outcome (i , j) is said to occur if i appears on the left most die and j on the other die.

Example 5:

If the experiment consists of measuring (in hours) the lifetime of a transistor, then the sample space consists of all nonnegative real numbers.

That is

$$S = \{x: 0 \le x < \infty\}$$

Example 6:

If the experiment consists selecting three items at random from a manufacturing process, such that each item is inspected and classified defective, D, or non defective, N. Then the sample space will be

S = {(D,D,D), (D,D,N), (D,N,D), (N,D,D),(D,N,N), (N,D,N), (N,N,D) , (N,N,N)}

3.Events

An **event**, A, is a subset of a sample space. If A is an event, we say that A has occurred if it contains the outcomes that occurred $A \subseteq S$. If an event A contains no outcomes, then *A* is an ***impossible*** event.

Example 7:

In example 3, if A = {(H, H), (H, T)}, then A is the event that a head appears on the first coin.

Example 8:

In example 4, if

A = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)},

then

A is the event that the sum of the dice equals 7.

Complementary Events

The **complement** of an event A with respect to S is the subset of all elements of S that are not in A. We denote the complement of A by A^c

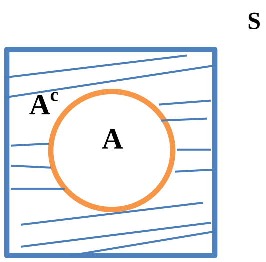
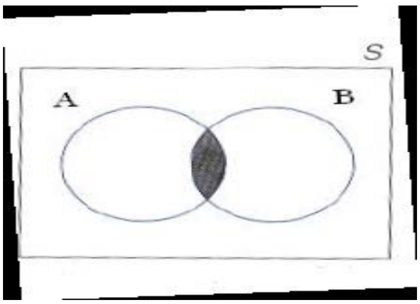


Figure 1: shaded region: A^c

Definition

The **intersection** of two events A and B, denoted by $A\cap B$, or AB is the event containing all elements that are common to A and B



Shaded region: $A\cap B$

Example 9:

In example 3, if

$$A = \{(H, H), (H, T), (T, H)\}$$

is the event that at least 1 head occurs, and

$$B = \{(H, T), (T, H), (T, T)\}$$

is the event that at least 1 tail occurs, then

$$A \cap B = \{(H, T), (T, H)\}$$

is the event that exactly 1 head and 1 tail appear.

Example 10:

In example 4, if $A = \{(1,6),(2,5),(3,4), (4,3), (5,2), (6,1)\}$,

is the event that the sum of the dice equals 7 and

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

is the event that the sum is 6, then the event

$$A \cap B = \emptyset.$$

That is A and B have no elements in common and therefore, cannot occur simultaneously.

Definition

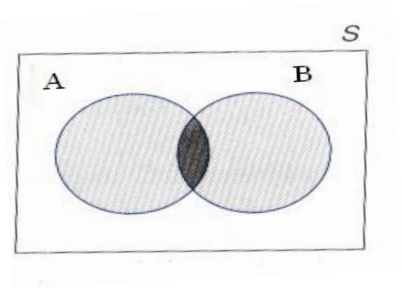
Two events A and B are **mutually exclusive**, or **disjoint** if

$A \cap B = \emptyset$, that is, if A and B have no elements in common.

It is clear that A and A^c are mutually exclusive.

Definition

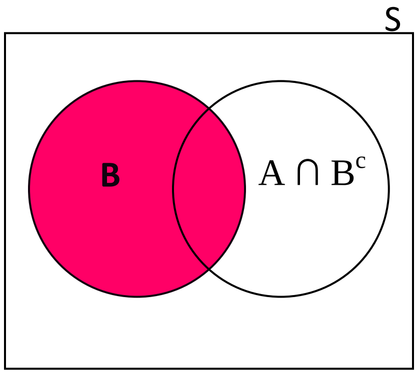
The **union** of the two events A and B, denoted by A U B, is the event containing all the elements that belong to A or B or both.



shaded region A U B

Definition

The **difference** of the two events A and B, denoted by A −B, is the event containing all the elements that belong to A and not belong to B, that is **A only** $(A \cap B^c)$



Example 11:

An electronic device is tested and its total time of service say t, is recorded. We shall assume the sample space to be $S = \{t| \; t \geq 0\}$. Let $A = \{t \mid t < 100\}$, $B = \{t \mid 50 \leq t \leq 200\}$, $C = \{t \mid t > 150\}$. Then

1- A U B = {t | t ≤ 200}

2- A ∩ B = {t| 50 ≤ t <100}

3- B U C = {t | t ≥ 50}

4- B ∩ C = {t| 150 < t ≤ 200}

5- A ∩ C = Ø

6- A U C = {t| t < 100 or t > 150}

7- A^c= {t| t ≥ 100}

8- C^c= {t| t ≤ 150}

Algebra's laws of the sets

1- $A \cup A = A, \quad A \cup S = S, \quad A \cup \varnothing = A$

$A \cap A = A, \quad A \cap S = A, \quad A \cap \varnothing = \varnothing$

2- Associative laws

$$(A \cup B) \cup C = A \cup (B \cup C), \qquad (A \cap B) \cap C = A \cap (B \cap C)$$

3- Commutative laws

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A$$

4-Distributive laws

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

5-Complement laws

$$A \cup A^c = S, \qquad A \cap A^c = \varnothing, \qquad (A^c)^c = A \quad S^c = \varnothing, \qquad \varnothing^c = S$$

6-De Morgen's law

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c$$

Remarks:-

1- *If* $A \subset B, \quad B \subset C \quad \Rightarrow \quad A \subset C$

2- *If* $A \subset B, \quad B \subset A \quad \Rightarrow \quad A = B$

3- *If* $A \subset B$, *then* $A \cup B = B$

4- $A \subset (A \cup B), \quad B \subset (A \cup B)$

5- $(A \cap B) \subset (A \cup B)$

6- $(A - B) \neq (B - A)$

7- $S - A = A^c, \quad \varnothing - A = \varnothing, \quad A - \varnothing = A, \quad A - A = \varnothing$

8- *If* $A \subset B, \quad \Rightarrow \quad A - B = \varnothing$

4. Axioms of probability

The probability of an event A is a number, P(A), such that

(1) $0 \leq P(A) \leq 1$

(2) $P(S) = 1$

(3) if $A_1, A_2, \ldots$ are mutually exclusive events

(i.e., $A_i \cap A_j = \varnothing,\ i \neq j$) then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Classical Probability

The (***classical***) probability, P(A), of an event A is given by

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

It is assumed here that the sample space is finite and all outcomes are equally likely to occur.

Example 12:

A coin is tossed twice. What is the probability that at least one head occurs?

Solution: The sample space for this experiment is

S = {(H, H), (H, T), (T, H), (T, T)} , n(S)=4

If the coin is balanced, each of these outcomes would be equally to occur.

If A represents the event of at least one head occurring, then

A = {(H, H), (H, T), (T, H)}, n(A)=3, and

$$P(A)=\frac{n(A)}{n(S)}=\frac{3}{4}$$

<u>Further Rules on probability</u>

1) Addition Rule

Let A and B are two not mutually exclusive events then:

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

2) Multiplication Rule

Let A and B are two independent events then:

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

3) Complement Rule

Let A an event subset from sample space then:

$$P(A^C) = P(S) - P(A) = 1 - P(A)$$

Properties:

1) $P(E \cup E^C) = P(E) + P(E^C) = 1$.

2) $P(E_1 \cap E_2 \cap E_3 ..... \cap E_n) = P(E_1) * P(E_2) * ..... * P(E_n)$

3) $P(E_1^c \cup E_2^c) = P(E_1 \cap E_2)^C = 1 - P(E_1 \cap E_2)$

4) $P(E_1^c \cap E_2^c) = P(E_1 \cup E_2)^C = 1 - P(E_1 \cup E_2)$.

Example 13:

If A and B are two not mutually exclusive events, suppose $P(A) = 0.5$ $P(B) = 0.3$ and $P(A \cap B) = 0.2$. **Find** the following probabilities:

1) $P(A \cup B)$ 2) $P(A^c \cap B^c)$ 3) $P(A \cup B^c)$

Solution:

1- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.5 + 0.3 - 0.2 = 0.6$

2- $P(A^c \cap B^c) = P(A \cup B)^c = 1 - 0.6 = 0.4$

3- $P(A \cup B^c) = P(A) + P(B^c) - P(A \cap B^c) = 0.5 + 0.7 - (0.5 - 0.2) = 0.9$

Example 14:

If A and B are two mutually exclusive events, suppose $P(A) = 0.4$, and $P(B)= 0.2$. **Find** the following probabilities

1- Probability of A or B   2- $P(A^c \cup B^c)$ 3) $P(A \cup B^c)$

Solution:

1- $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) = 0.4 + 0.2 = 0.6$

2- $P(A^c \cup B^c) = P(A \cap B)^c = 1 - P(\varnothing) = 1$

3- $P(A \cup B^c) = P(A) + P(B^c) - P(A \cap B^c) = 0.4 + 0.8 - (0.4) = 0.8$

Example 15:

A student is taking two courses, Math & Phys, the probability the student will pass the math is 0.6, and the probability of passing Phys is 0.7 , the probability of passing both courses is 0.5, what is the probability of passing at least one course.

Solution:

Let the probability of student pass in Math $P(A) = 0.6$ and in Phys $P(B) = 0.7$, also pass in both them $P(A \cap B) = 0.5$

$P(\text{Math or Phys}) = P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.6 + 0.7 - 0.5 = 0.8$

5.Counting sample points:

1.Multiplication Rule

If an operation can be performed in n_1 ways, and if for each of these a second operation can be performed in n_2 ways, then the two operations can be performed together in

$$\mathbf{n_1\, n_2} \textbf{ ways}.$$

- The rule is sometimes called the ***Basic Principle of Counting***.

Example 16:

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 the Multiplication Rule , we have

26 . 26 . 26 . 10. 10 .10 .10 = 175760000 possible license plates.

Example 17:

In the previous example how many license plates would be possible if no letter or digit can be repeated?

Solution :

In this case there would be 26 . 25 . 24 . 10. 9 .8 .7 = 78624000 possible license plates.

2.Permutation:

The number of arrangements **of size *r* from a set of *n* distinct objects** is given by

$$^{n}P_{r}=\frac{n!}{(n-r)!}; \qquad 0\le r\le n$$

Note: if $r = n$,

$$^{n}P_{n}=\frac{n!}{(n-n)!}=\frac{n!}{0!}=\frac{n!}{1}=n!$$

If $r = 0$,

$$^{n}P_{0}=\frac{n!}{n!}=1$$

Example 17:

In a class of ten students, six are to be chosen and seated in a row for a picture. How many different pictures are possible?

Solution

Thus, there are

$10 \times 9 \times 8 \times 7 \times 6 \times 5 = 151200$ different pictures.

Note that:

$$10\times9\times8\times7\times6\times5$$

$$=10\times9\times8\times7\times6\times5\times\frac{4\times3\times2\times1}{4\times3\times2\times1}$$

$$=\frac{10!}{4!}={}^{10}P_{6}=151200$$

Example 19:

Mr. Jones has 10 books that he is going to put on his bookshelf. Of these, 4 are mathematics books, 3 are chemistry books, 2 are history books, and 1 is a language book. Jones wants to arrange his books so that all the books dealing with the same subject are together on the shelf. How many different arrangements are possible?

Solution:

There are **4! 3! 2! 1!** Arrangements such that the mathematics books are first in line, then the chemistry books, then the history books, and then the language book. Similarly, for each ordering of the subjects, there are 4! 3! 2! 1! possible arrangements. Hence, as there are **4! Possible orderings of the subjects**, the desired answer is:

$$4!\ 4!\ 3!\ 2!\ 1! = 6912.$$

Example 20:

A class in probability theory 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 ranking are possible?

(b) If the men are ranked just among themselves and the women among themselves, how many different rankings are possible?

(c) What is the probability that women receive the top 4 scores ?

Solution:

(a) As each ranking corresponding to a particular ordered arrangement of the 10 people, we see that the answer to this part is 10! = 3628800.

(b) As there are 6! Possible ranking of the men among themselves and 4! possible ranking of the women among themselves, it follows from there are

2 (6!)(4!) = 2 (720) (24) = 34560 possible ranking in this case.

(c) Let A be the event that women receive the top 4 scores, then, n(A) = (4!) (6!) = (24)(720) = 17280 .

n(S) = 10! = 3628800, so

$$P(A)=\frac{n(A)}{n(S)}=\frac{4!\ 6!}{10!}=\frac{1}{210}$$

Permutation with repetition

The number of distinct permutations of n objects of which n_1 are alike, n_2 are alike ,…, n_r are alike , is given by:

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

Example 21:

What is the number of permutations of the letters in the word "ball"?

Note that:

(i) The answer is not 4!since we do not have 4 distinct objects.

(ii) We have a set of only 3 distinct objects .But note that the answer is not 3!.The permutations here involve repetitions

Solution

- Suppose that there are no repetitions:

  ball

- In this case, there are 4! permutations.
- But for each of these permutations, there is exactly one permutation where l and l switch positions.
- These are, really, **the same permutations**.
- Thus, the number of permutations of the letters in "ball" is

  $4!/2! = 12$

Example 22:

How many distinct permutations can formed from all the letters of each word

(i) them **(ii)** that **(iii)** radar

(iv) unusual **(v)** sociological

Solution:

(i) $4! = 24$, since there are 4 letters and no repetition.

(ii) $\frac{4!}{2!} = 12$ since there are 4 letters of which 2 are t's.

(iii) $\frac{5!}{2!\ 2!} = 30$ since there are 5 letters of which 2 are r's and 2 are a's.

(iv) $\frac{7!}{3!} = 840$ since there are 7 letters of which 3 are u's.

(v) $\frac{12!}{3!\ 2!\ 2!\ 2!} =$ since there are 12 letters of which 3 are o's , 2 are i's , 2 are c's and 2 are l's.

3.Combinations:

The number of selections of size r from a set of n distinct objects is given by:

$$^{n}C_{r} = \binom{n}{r} = \frac{n!}{(n-r)!\ r!}$$

Example 23:

A committee of 3 is to be selected from a group of 20 people. How many different committees are possible?

Solution:

There are $^{20}C_{3} = \binom{20}{3} = \frac{20!}{3!\ 17!} = 1140$ possible committees.

Example 24:

A committee of 5 is to be selected from a group of 6 men and 9 women. If the selection is made randomly, what is the probability that the committee consists of 3 men and 2 women?

Solution:

If A represents the event of selecting the committee, then:

n(S) = $^{15}C_5 = \binom{15}{5}$, and n (A) = $^{6}C_3\,^{9}C_2 = \binom{6}{3}\binom{9}{2}$,

There for,

$$P(A) = \frac{n(A)}{n(S)} = \frac{\binom{6}{3}\binom{9}{2}}{\binom{15}{5}} = \frac{240}{1001}$$

Example 25:

Two balls are selected at random from a bag with four white balls and three black balls, where order is not important.

1- What would be an appropriate sample space S?

Solution: Denote the set of balls by

B = {w1, w2, w3, w4, b1, b2, b3}

The number of outcomes in S (which are sets of two balls) is then

$^{7}C_2 = 21$

2- What is the probability that both balls are white?

$$\frac{^{4}C_2}{^{7}C_2} = \frac{6}{21} = \frac{2}{7},$$

3- What is the probability that both balls are black?

$$\frac{{}^{3}C_{2}}{{}^{7}C_{2}} = \frac{3}{21} = \frac{1}{7},$$

4- What is the probability that one is white and one is black?

$$\frac{{}^{3}C_{1} * {}^{4}C_{1}}{{}^{7}C_{2}} = \frac{4*3}{21} = \frac{4}{7},$$

Example 26:

A young boy asks his mother to get 5 Game cartridges from his collection of 10 arcade and 5 sports games. How many ways are there that his mother can get 3 arcade and 2 sports games?

Solution:

The number of ways of selecting 3 cartridges from 10 is

$$\binom{10}{3} = \frac{10!}{3!(7)!} = 120$$

The number of ways of selecting 2 cartridges from

$$\binom{5}{2} = \frac{5!}{2!(2)!} = 10$$

Using the multiplication rule with $n_1 = 120$ and $n_2 = 10$, we have (120) (10) = 1200 ways.

6.Conditional Probability, Independence, and the Product Rule

One very important concept in probability theory is conditional probability. In some applications, the practitioner is interested in the probability structure under certain restrictions.

The probability of an event B occurring when it is known that some event A has occurred is called a **conditional probability** and is denoted by $P(B|A)$. The symbol $P(B|A)$ is usually read "the probability that B occurs given that A occurs" or simply "the probability of B, given A."

Definition:

Conditional Probability. Suppose that we learn that an event B has occurred and that we wish to compute the probability of another event A taking into account that we know that B has occurred. The new probability of A is called the conditional probability of the event A given that the event B has occurred and is denoted P (A|B).

If $P(B) > 0$,

we compute this probability as $P(A|B) = P(A \cap B)/P(B)$

Example 27:

The probability that a regularly scheduled flight departs on time is $P(D) = 0.83$; the probability that it arrives on time is $P(A) = 0.82$; and the probability that it departs and arrives on time is

$P(D \cap A) = 0.78$.

Find the probability that a plane

(a) arrives on time, given that it departed on time.

(b) departed on time, given that it has arrived on time.

Solution

Using Definition of *Conditional Probability*, we have the following.

(a) The probability that a plane arrives on time, given that it departed on time, is

$$P(A|D) = P(D \cap A)/P(D)$$
$$= 0.78/\ 0.83 = 0.94.$$

(b) The probability that a plane departed on time, given that it has arrived on time, is

$$P(D|A) = P(D \cap A)/P(A) = 0.78/\ 0.82 = 0.95.$$

The notion of conditional probability provides the capability of reevaluating the idea of probability of an event in light of additional information, that is, when it is known that another event has occurred. The probability $P(A|B)$ is an updating of $P(A)$ based on the knowledge that event B has occurred. The probability that it arrives on time, given that it did not depart on time i.e $P(A|D^c)$ is

$$P(A|D^c) = P(A \cap D^c)/P(D^c)$$
$$= 0.82 - 0.78\ /\ 0.17 = 0.24.$$

<u>1.Independent Events</u>

In the die-tossing experiment Example, we note that $P(B|A) = 2/5$ whereas $P(B) = 1/3$. That is, $P(B|A) \neq P(B)$, indicating that B depends on A. Now consider an experiment in which 2 cards are drawn in succession from an ordinary deck, with replacement. The events are defined as

A: the first card is an ace, B: the second card is a spade. Since the first card is replaced, our sample space for both the first and the second draw consists of 52 cards, containing 4 aces and 13 spades. Hence,

$$P(B|A) = 13/52 = 1/4 \text{ and } P(B) = 13/52 = 1/4$$

That is, $P(B|A) = P(B)$. When this is true, the events A and B are said to be **independent**.

Definition:

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

The condition $P(B|A) = P(B)$ implies that $P(A|B) = P(A)$, and conversely.

For the card-drawing experiments, where we showed that

$$P(B|A) = P(B) = 1/4,$$

we also can see that $P(A|B) = P(A) = 1/13$.

The Product Rule, or the Multiplicative Rule (or **product rule**), which enables us to calculate the probability that two events will both occur.

Theorem:

If in an experiment the events A and B can both occur, then

$$P(A \cap B) = P(A)P(B|A), \text{ provided } P(A) > 0.$$

Thus, the probability that both *A* and *B* occur is equal to the probability that *A* occurs multiplied by the conditional probability that *B* occurs, given that *A* occurs. Since the events *A B* and *B A* are equivalent, it follows that we can also write

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

In other words, it does not matter which event is referred to as *A* and which event is referred to as *B*.

Theorem:

Two events *A* and *B* are *independent* if and only if

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

Therefore, to obtain the probability that two independent events will both occur, we simply find the product of their individual probabilities.

Example 29:

You are given the following information on Events A, B, C, and D. **Where** $P(A)=0.1, P(C)=0.4$, $P(B)=0.3$, $P(C|B)=0.6$, $P(A\cap C)=0.04$, $P(B|A)=0.9$

(i) Compute $P\left(A^c \cap C\right)$.

(ii) Compute $P\left(C^c \cup B\right)$.

(iii) Are A and B mutually exclusive? Explain.

(iv) Are A and C independent? Explain.

Solution:

(i) $P\left(A^c \cap C\right) = P(C) - P(A \cap C) = 0.4 - 0.04 = 0.36$.

(ii)

$$P\left(C^c \cup B\right) = 1 - P\left(B^C \cap C\right) = 1 - \left[P(C) - P(B \cap C)\right] = 1 - (0.4 - 0.18)$$
$$= 1 - 0.22 = 0.78$$

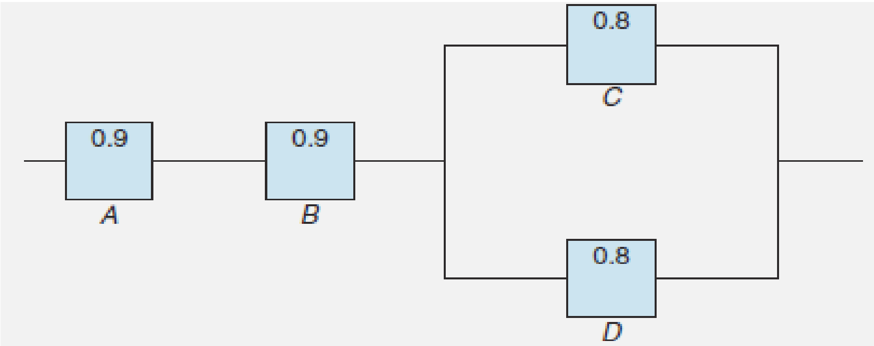
since $P(B \cap C) = P(B)P(C \mid B) = 0.3 \cdot 0.6 = 0.18$

(iii) Since $P(A \cap B) = P(A)P(B \mid A) = 0.1 \cdot 0.9 = 0.09 \neq 0$, A and B are not mutually exclusive.

(iv) Since $P(A \cap C) = 0.04 = P(A)P(C) = 0.1 \cdot 0.4$, A and B are independent.

Example 30:

An electrical system consists of four components as illustrated in the next Figure . The system works if components *A* and *B* work and either of the components *C* or *D* works. The reliability (probability of working) of each component is also shown In the Figure . Find the probability that (a) the entire system works and (b) the component *C* does not work, given that the entire system works. Assume that the four components work independently.



Solution:

In this configuration of the system, *A*, *B*, and the subsystem *C* and *D* constitute a serial circuit system, whereas the subsystem *C* and *D* itself is a parallel circuit system.

(a) Clearly the probability that the entire system works can be calculated as follows:

$$P[A \cap B \cap (C \cup D)] = P(A)P(B)P(C \cup D)$$
$$= P(A)P(B)[1 - P(C^c \cap D^c)]$$
$$= P(A)P(B)[1 - P(C^c)P(D^c)]$$
$$= (0.9)(0.9)[1 - (1 - 0.8)(1 - 0.8)]$$
$$= 0.7776.$$

The equalities above hold because of the independence among the four components.

(b) To calculate the conditional probability in this case, notice that

P = *P*(the system works but *C* does not work)/*P*(the system works) = $P(A \cap B \cap C^c \cap D)$/P(the system works)

$= (0.9)(0.9)(1 - 0.8)(0.8)/0.7776 = 0.1667.$

The multiplicative rule can be extended to more than two-event situations.

Theorem: If, in an experiment, the events $A_1, A_2, \ldots, A_k$ can occur, then

$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A3|A_1 \cap A_2) \cdots P(A_k|A_1 A_2 \cdots A_{k-1})$.

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

$P(A_1 \cap A_2 \cap \cdots \cap A_k) = P(A_1)P(A_2) \cdots P(A_k)$.

<u>7.Bayes' Theorem</u>:

Let $A_1, A_2, \ldots, A_k$ be a collection of k mutually exclusive and exhaustive events with *prior* probabilities $p(A_i)$, I = 1, 2,…k, . Then for any other event B for which $p(B) > 0$, the *posterior* probability of A_j given that B has occurred is

$$P(A_j / B) = \frac{P(A_j \cap B)}{p(\mathrm{B})} = \frac{P(\mathrm{B}/A_j)\mathrm{p}(\mathrm{A}_j)}{\sum_{i=1}^{k} P(\mathrm{B}/A_i)\mathrm{p}(\mathrm{A}_i)}$$

The transition from the second to the third expression above rests on using the multiplication rule in the numerator and the law of total probability in the denominator.

Example 31:

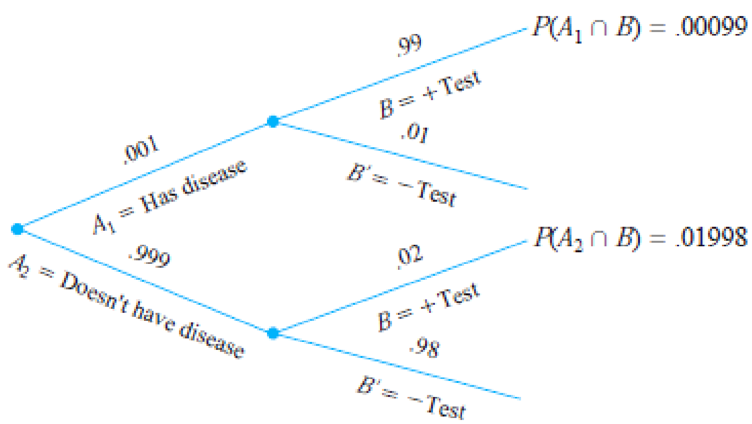
Incidence of a rare disease. Only 1 in 1000 adults is afflicted with a rare disease for which a diagnostic test has been developed. The test is such that when an individual actually has the disease, a positive result will occur 99% of the time, whereas an individual without the disease will show a positive test result only 2% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease?

Solution

To use Bayes' theorem, let A_1 = individual has the disease, A_2= individual does not have the disease, and B = positive test result.

Then $p(A_1)$ =0.001, $p(A_2)$ = 0.99, , $p(B \backslash A_1)$ =0.99, and $p(B \backslash A_2)$ = 0.02 .

The tree diagram for this problem is the following Figure .



Next to each branch corresponding to a positive test result, the multiplication rule yields the recorded probabilities. Therefore, $P(B) = .00099 + .01998 = 0.02097$,

from which we have

$$P(A_1/B) = \frac{P(A_1 \cap B)}{p(\mathrm{B})} = \frac{0.00099}{0.02097} = 0.047$$

This result seems counterintuitive; the diagnostic test appears so accurate that we expect someone with a positive test result to be highly likely to have the disease, whereas the computed conditional probability is only .047. However, the rarity of the disease implies that most positive test results arise from errors rather than from diseased individuals. The probability of having the disease has increased by a multiplicative factor of 47 (from prior .001 to posterior .047); but to get a further increase in the posterior probability, a diagnostic test with much smaller error rates is needed.

8. Exercises

1. If a multiple-choice test consists of 5 questions, each with 4 possible answers of which only 1 is correct,

(a) in how many different ways can a student check off one answer to each question?

(b) in how many ways can a student check off one answer to each question and get all the answers wrong?

2. A contractor wishes to build 9 houses, each different in design. In how many ways can he place these houses on a street if 6 lots are on one side of the street and 3 lots are on the opposite side?

3. A box contains 500 envelopes, of which 75 contain $100 in cash, 150 contain $25, and 275 contain $10. An envelope may be purchased for $25. What is the sample space for the different amounts of money?

Assign probabilities to the sample points and then find the probability that the first envelope purchased contains less than $100.

4. If 3 books are picked at random from a shelf containing 5 novels, 3 books of poems, and a dictionary, what is the probability that

(a) the dictionary is selected?

(b) 2 novels and 1 book of poems are selected?

5. A class in advanced physics is composed of 10 juniors, 30 seniors, and 10 graduate students. The final grades show that 3 of the juniors, 10 of the seniors, and 5 of the graduate students received an *A* for the course. If a student is chosen at random from this class and is found to have earned an *A*, what is the probability that he or she is a senior?

6. In the senior year of a high school graduating class of 100 students, 42 studied mathematics, 68 studied psychology, 54 studied history, 22 studied both mathematics and history, 25 studied both mathematics and psychology, 7 studied history but neither mathematics nor psychology, 10 studied all three subjects, and 8 did not take any of the three. Randomly select a student from the class and find the probabilities of the following events.

(a) A person enrolled in psychology takes all three subjects.

(b) A person not taking psychology is taking both history and mathematics.

7. The probability that a married man watches a certain television show is 0.4, and the probability that a married woman watches the show is 0.5. The probability that a man watches the show, given that his wife does, is 0.7. Find the probability that

(a) a married couple watches the show;

(b) a wife watches the show, given that her husband does;

(c) at least one member of a married couple will watch the show.

8. In a certain region of the country it is known from past experience that the probability of selecting an adult over 40 years of age with cancer is 0.05. If the probability of a doctor correctly diagnosing a person with cancer as having the disease is 0.78 and the probability of incorrectly diagnosing a person without cancer as having the disease is 0.06, what is the probability that an adult over 40 years of age is diagnosed as having cancer?

Chapter 2

Random Variables and Expectation

1. Introduction

We have mentioned that our interest in the study of a random phenomenon is in the statements we can make concerning the events that can occur, and these statements are made based on probabilities assigned to simple outcomes. One of the immediate steps that can be taken in this unifying attempt is to require that each of the possible outcomes of a random experiment be represented by a real number. In this way, when the experiment is performed, each outcome is identified by its assigned real number rather than by its physical description. For example, when the possible outcomes of a random experiment consist of success and failure, we arbitrarily assign the number one to the event 'success' and the number zero to the event 'failure'. The associated sample space has now 1, 0 as its sample points instead of success and failure, and the statement 'the outcome is 1' means 'the outcome is success'.

Consequently, sample spaces associated with many random experiments of interest are already themselves sets of real numbers. The real-number assignment procedure is thus a natural unifying agent. On this basis, we may introduce a variable, which is used to represent real numbers, the values of which are determined by the outcomes of a random experiment.

2. Random variable

Consider a random experiment to which the outcomes are elements of sample space in the underlying probability space. In order to construct a model for a random variable, we assume that it is possible to assign a real number $X(s)$ for each outcome s following a certain set of rules. We see that the 'number' $X(s)$ is really a real-valued point function defined over the domain of the basic probability space

Definition1. The point function $X(s)$ is called a random variable if (a) it is a finite real-valued function defined on the sample space S of a random experiment for which the probability function is defined, and (b) for every real number X, the set $\{s : X(s) < X\}$ is an event.

To see more clearly the role a random variable plays in the study of a random phenomenon; consider again the simple example where the possible outcomes of a random experiment are success and failure. Let us again assign number one to the event success and zero to failure. If X is the random variable associated with this experiment, then X takes on two possible values: 1 and 0. Moreover, the following statements are equivalent:

1- The outcome is success.

2- The outcome is 1.

3- X=1

The random variable is called a random variable if it is defined over a sample space having a finite or a countably infinite number of sample points. In this case, random variable takes on discrete values, and it is possible to enumerate all the values it may assume. In the case of a sample space having an uncountable infinite number of sample points, the associated random variable is called a random variable, with its values distributed over one or more continuous intervals on the real line. We make this distinction because they require different probability assignment considerations. Both types of random variables are important in science and engineering.

3. Discrete Random variables

To each point in the sample space we will assign a real number denoting the value of the variable X. The value assigned to X will vary from one sample point to another, but some points may be assigned the same numerical value. Thus, we have defined a variable that is a function of the sample points in S, and {all sample points where X = a} is the numerical event assigned the number a. Indeed, the sample space S can be partitioned into subsets so that points within a subset are all assigned the same value of X. These subsets are mutually exclusive since no point is assigned two different numerical values. The partitioning of S is symbolically indicated in for a random variable that can assume values 0, 1, 2, 3, and 4.

Definition 2: A random variable *X is* said to be discrete if it can assume only a finite or countably infinite number of distinct values.

Example.1:

Let X is the random variable defined by the Head appear, in times, for toss a coin 3 times respectively. Find the value of x.

Solution: The sample space

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

Let Head is the random variable X, then the conjugate space

$$X(s)=\{3,2,1,0\} \text{ then } x_i=0,1,2,3$$

Example.2:

Suppose a sampling plan involves sampling items from a process until a defective is observed. The evaluation of the process will depend on how many consecutive items are observed. In that regard, let X be a random variable defined by the number of items observed before a defective is found. With N a non-defective and

D a defective, sample spaces are S = {D} given X = 1, S = {ND} given X = 2, S = {NND} given X = 3, and so on.

3.1. Discrete probability distribution

A discrete random variable assumes each of its values with a certain probability. For a discrete variable X can be represented by a formula, a table, or a graph that provides (i.e: example 1)

1- $P(x) = P(X = x)\ \forall\ x \in X(s)$

2- $0 \le P(x) \le 1$

3- $\sum P(x_i) = 1$

x	**0 1 2 3**
***P*(*x*)**	**1/8 3/8 3/8 1/8**

Example.3:

A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Solution: Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can only take the numbers 0, 1, and2.

Now we calculate, $P(0) = P(X = 0) = \frac{136}{190}$,

$P(1) = P(X = 1) = \frac{51}{190}$,

$P(3) = P(X = 3) = \frac{3}{190}$

x	**0**	**1**	**2**
***P*(*x*)**	**136/190**	**51/190**	**3/190**

Thus, the probability distribution of X is

Notice that the probabilities associated with all distinct values of a discrete random variable must sum to 1.

Example.4:

Suppose that a pair of fair dice is tossed and let the discrete random variable X denote the sum of the points. Obtain the range of the discrete random variable X.

Solution:

Random Variable	**Events**
$x_1=2$	$A_1=\{(1,1)\}$
$x_2=3$	$A_2=\{(1,2),(2,1)\}$
$x_3=4$	$A_3=\{(2,2),(3,1),(1,3)\}$
$x_4=5$	$A_4=\{(1,4),(4,1),(3,2),(2,3)\}$
$x_5=6$	$A_5=\{(3,3),(2,4),(4,2),(5,1),(1,5)\}$
$x_6=7$	$A_7=\{(3,4),(4,3),(5,2),(2,5),(1,6),(6,1)\}$
$x_7=8$	$A_7=\{(4,4),(5,3),(3,5),(6,2),(2,6)\}$
$x_8=9$	$A_8=\{(4,5),(5,4),(3,6),(6,3)\}$
$x_9=10$	$A_9=\{(5,5),(4,6),(6,4)\}$
$x_{10}=11$	$A_{10}=\{(5,6),(6,5)\}$
$x_{11}=12$	$A_{11}=\{(6,6)\}$

Then $x=\{2,3,4,5,6,7,8,9,10,11,12\}$

Example.5:

Find the probability distribution corresponding to the random variable X of a coin is tossed twice. And Construct a probability graph.

Solution:

Assuming that the coin is fair we have

$$S=\{(H,H),(H,T),(T,H),(T,T)\}$$

Let $A_1=\{(T,T)\}$, $A_2=\{(H,T),(T,H)\}$, and $A_3=\{(H,H)\}$

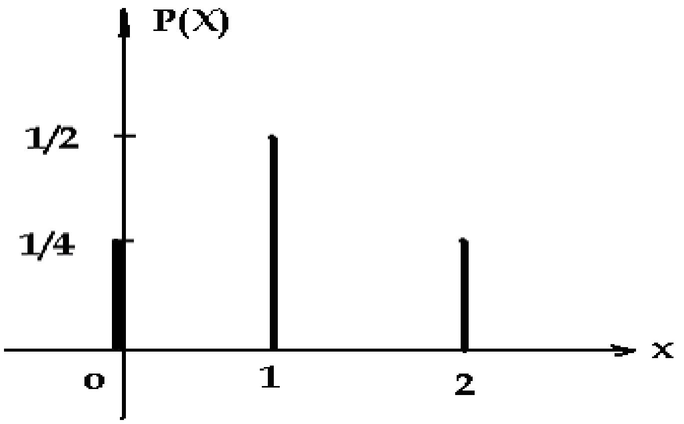
$$P(A_1)=\frac{1}{4},\quad P(A_2)=\frac{2}{4}=\frac{1}{2},\quad P(A_3)=\frac{1}{4}$$

Then $P(X=0)=P(A_1)=\frac{1}{4}$,

$P(X=1)=P(A_2)=\frac{1}{2}$ and $P(X=2)=P(A_3)=\frac{1}{4}$

X	***0***	***1***	***2***
P(X)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

A probability graph can be obtained by use of a bar chart.



3.2. The cumulative distribution function, or distribution function for a random variable X is defined by

$$P(X \le x) = F(x).$$

Where x is any real number, i.e. $-\infty < x < \infty$. The distribution function can be obtained from the probability function by

$$F(x) = P(X \le x) = \sum_{0 \le x} P(x)$$

There are many problems where we may wish to compute the probability that the observed value of a random variable X will be less than or equal to some real number x. Writing $F(x) = P(X \le x)$ for every real number x, we define $F(x)$ to be the **cumulative distribution function** of the random variable X.

Example.6:

Consider the simple condition in which components are arriving from the production line and they are stipulated to be defective or not defective.

Solution:

Define the random variable X by

X = 1, if the component is defective,

0, if the component is not defective.

Clearly the assignment of 1 or 0 is arbitrary though quite convenient. The random variable for which 0 and 1 are chosen to describe the two possible values is called a Bernoulli random variable.

Example.6:

A store carries flash drives with either1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The accompanying table gives the distribution of X = the amount of memory in a purchased drive:

X	1	2	4	8	16
p(X)	0.05	0.10	0.35	0.40	0.10

Let's first determine $F(x)$ for each of the five possible values of X:

$F(16) = P(X \leq 16) = 1$

$F(8) = P(X \leq 8) = p(1) + p(2) + p(4) + p(8) = .90$

$F(4) = P(X\leq 4) = P(Y = 1 \text{ or } 2 \text{ or } 4) = p(1) + p(2) + p(4) = .50$

$F(2) = P(X \leq 2) = P(Y = 1 \text{ or } 2) = p(1) + p(2) = .15$

$F(1) = P(X\leq 1) = P(Y = 1) = p(1) = .05$

Now for any other number x, $F(x)$ will equal the value of F at the closest possible value of X to the left of x.

4. Continuous Random variables

A continuous random variable has a probability of 0 of assuming exactly any of its values. Consequently, its probability distribution cannot be given in tabular form. We shall concern ourselves with computing probabilities for various intervals of continuous random variables such as:

$$P(a < X < b), \qquad P(W \geq c), \text{ and so forth.}$$

Note that when X is continuous, for a continuous random variable, its Probability Density Function, is a continuous function of and the derivative

$$f(x) = \frac{dF(x)}{dx}.$$

The **cumulative distribution function** $F(x)$ of a continuous random variable

X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx, \quad -\infty < x < \infty.$$

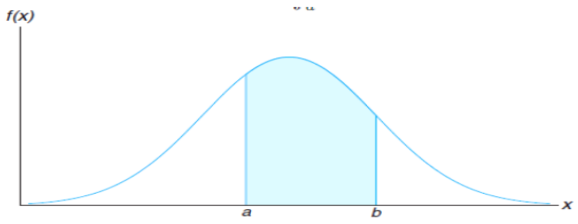
A function f(x) is called probability density function (p.d.f) if the following conditions satisfied

1- $f(x) \geq 0$

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

3- $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) =$

$$= P(a < X < b) = \int_{a}^{b} f(x)\, dx$$



Should this range of *X* be a finite interval, it is always possible to extend the interval to include the entire set of real numbers by defining *f*(*x*) to be zero at all points in the extended portions of the interval. In the Figure, the probability that *X* assumes a value between *a* and *b* is equal to the shaded area under the density function between the ordinates at *x* = *a* and *x* = *b*, and from integral calculus is given by

$$P(a < x < b) = \int_a^b f(x)dx$$

Example.7:

A continuous random variable X has a density function given by

$$f(x) = \begin{cases} \frac{1}{4} & 0 < x \le 4 \\ 0 & \text{elsewhere} \end{cases}$$

Show that $f(x)$ is a valid probability density function.

Solution

It is clear that $f(x) \ge 0$

$$\int_{a}^{b} f(x)dx = \int_{0}^{4} \frac{1}{4}\,dx + 0 = 1$$

Then $f(x)$ is a density function.

Example.8:

A continuous random variable X has a density function given by

$$f(x) = \begin{cases} k\,x^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the constant k

(b) Compute $P(1<X<2)$

Solution

a) Since $f(x)$ satisfies property (1) if $k \geq 0$, it must satisfy property (2) in order to be a density function. Now:

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx$$

$$= 0 + \int_{0}^{3} kx^2\,dx + 0 = \frac{k}{3}\left[x^3\right]_0^3 = 1$$

$$9k = 1 \quad \Rightarrow \quad k = \frac{1}{9}$$

b) $$P(1<X<2) = \int_{1}^{2} \frac{1}{9}x^2\,dx = \left.\frac{x^3}{27}\right|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

Example.9:

Suppose that the error in the reaction temperature, in ◦C, for a controlled laboratory experiment is a continuous random variable *X* having the probability density function

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 \le x \le 2 \\ 0 & elsewhere \end{cases}$$

(a) **Verify** that $f(x)$ is a density function.

(b) **Find** $P(0 < x \le 1)$.

Solution

We use Definition.

(a) Obviously, $f(x) \ge 0$. To verify condition 2, we have

$$\int_a^b f(x)dx = \int_{-1}^{2} \frac{x^2}{3} dx = 1$$

(b) $p(0 < x \le 1) = \int_0^1 \frac{x^2}{3} dx = \frac{1}{9}$

5. Expectations and Moments

While a probability distribution [$F_X(x)$, $P_X(x)$, or $f_X(x)$] contains a complete description of a random variable *X*, it is often of interest to seek a set of simple numbers that gives the random variable some of its dominant features. These numbers include moments of various orders associated with *X*.

Definition.

Let g(X) be a real-valued function of a random variable X. The mathematical expectation, or simply expectation, of g(X), denoted by

$$E(g(X)) = \sum_i g(x_i) P_X(x_i)$$

If X is discrete, where x_1, x_2, . . . are possible values assumed by X. When the range of i extends from 1 to infinity, the sum $\sum_i g(x_i) P_X(x_i)$ exists if it converges absolutely.

If random variable X is continuous, the expectation is defined by

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

Let us note some basic properties associated with the expectation operator.

For any constant c and any functions g(X) and h(X) for which expectations

exist, we have

1- $E(c) = c$

2- $E(cg(X)) = cE(g(X))$

3- $E(g(X) + h(X)) = E(g(X)) + E(h(X))$

5.1. Moments

Let $g(X) = X^n$, n=1, 2, …; the expectation $E(X^n)$, when it exist, is called the n^th moment of X. It is denoted by n and is given by

$$\mu_n = E(X^n) = \sum_i x_i^n P(x_i) \qquad \text{for X is discrete.}$$

$$\mu_n = E(X^n) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{for X is continuous.}$$

Example.10:

Let X is the waiting time (in minutes) of a customer waiting to be served at a ticket counter has the density function.

$$f(x)\begin{cases} 2e^{-2x} & \text{for } x \geq 0 \\ 0, & elsewhere. \end{cases}$$

Determine the average waiting time.

Solution:

$$E(X) = \int_0^{\infty} xf(x)dx$$

$$= \int_0^{\infty} 2xe^{-2x} dx = \frac{1}{2}$$

Proof that $\quad Var(X) = E((X - \mu)^2) = E(X^2) - \mu^2$

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

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

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

$$= \sum x^2 p(x) - 2\mu \sum xp(x) + \mu^2 \sum p(x)$$

$$= E(X^2) - 2\mu E(X) + \mu^2 \times 1$$

$$= E(X^2) - 2\mu^2 + \mu^2, \text{ as } E(X) = \mu$$

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

6. Exercises.2

Ex.1: In the experiment of tossing a fair die once, the range of the random variable X is {1, 3} where "1" is specified for the appearance of an odd number on the upper face and "3" is specified for the appearance of an even number. Find the probability distribution of this variable.

Ex.2: In the experiment of tossing a coin three consecutive times to observe the type of the appearing faces, if he random variable "X" is defined by "Twice the number of the appearing heads." Write the probability distribution of X.

Ex.3: Let X be a continuous random variable with probability density function given by $f(x) = \begin{cases} 3x^2, & 0 \le x \le 1 \\ 0 & elsewhere \end{cases}$

Find $F(x)$. Graph both $f(x)$ and $F(x)$.

Ex.4: Let X be a discrete random variable taking the values x = 0, 1, 2with probabilities 1/4, 1/2, 1/4 respectively. Plot F(x).

Ex.5: Let the distribution function of random variable X given as follows

x	***1***	***2***	***3***	***4***
f(x)	***1/8***	***3/8***	***3/8***	***1/8***

Determine:

(a) Probability function
(b) $P(1<X\leq 3)$
(c) $P(X\geq 2)$
(d) $P(X<3)$

Exe.6: Let $f(x)=\begin{cases} K(2-x) & \text{for} \quad 0\leq x\leq 2 \\ 0 & \text{elsewhere} \end{cases}$

(a) Find K such that f(x).density function
(b) Plot f(x).
(c) $E(X)$

Ex.7: A continuous random variable X that can assume values between $x = 1$ and $x = 3$ has a density function given by $f(x) = 1/2$.

(a) **Show** that the area under the curve is equal to 1.

(b) **Find** $P(2 < X < 2.5)$. (c) **Find** $P(X \leq 1.6)$.

(d) **Find** $F(x)$. Use it to evaluate $P(2 < X < 2.5)$.

(e) $E(X)$

Ex.8: Consider the density function

$$f(x)=\begin{cases} k\sqrt{x}, & 0\leq x\leq 1 \\ 0 & elsewhere \end{cases}.$$

(a) **Evaluate** k.

(b) **Find** $F(x)$ and use it to evaluate $P(0.3 < X < 0.6)$.

(e) $E(X)$

Chapter 3

Probability Distributions

1. Discrete probability distributions

No matter whether a discrete probability distribution is represented graphically by a histogram, the behavior of a random variable is described. Often, the observations generated by different statistical experiments have the same general type of behavior. Consequently, discrete random variables associated with these experiments can be described by essentially the same probability distribution and therefore can be represented by a single formula. In fact, one needs only a handful of important probability distributions to describe many of the discrete random variables encountered in practice. Such a handful of distributions describe several real-life random phenomena. For instance, in a study involving testing the effectiveness of a new drug, the number of cured patients among all the patients who use the drug approximately follows binomial, hypergeometric, and Poisson distributions.

1.1. The Binomial Distribution

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled success or failure. The most obvious application deals with the testing of items as they come off an assembly line, where each trial may indicate a defective or a non-defective item. We may choose to define either outcome as a success. The process is referred to as a Bernoulli process. Each trial is called a Bernoulli trial. Observe, for example, if one were drawing cards from a deck, the probabilities for repeated trials change if the cards are not replaced. That is, the probability of selecting a heart on the first draw is 1/4, but on the second draw it is a conditional probability having a value of 13/51 or 12/51, depending on whether a heart appeared on the first draw: this, then, would no longer be considered a set of Bernoulli trials.

If we define "success" as a defective component, then p is the proportion of defective components in the population. Many experiments consist of performing a sequence of Bernoulli trials. For example, we might sample several components from a very large lot and count the number of defectives among them. This amounts to conducting several independent Bernoulli trials and counting the number of successes. The number of successes is then a random variable, which is said to have a **binomial distribution**.

The number X of successes in n Bernoulli trials is called a binomial random variable. The probability distribution of this random variable is called the binomial distribution, and its values will be denoted by B (n, p) since they depend on the number of trials and the probability of a success on a given trial.

Probability density function of a Binomial distribution

We can now generalize this result to produce a formula for the probability of *x* successes in *n* independent Bernoulli trial can result in a success with probability p and a failure with probability q = 1-p. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials. In other words, we can compute P(X = x) where $X \sim Bin\ (n, p)$.

We can now define the probability mass function for a binomial random variable.

$$P(x) = P(X = x) = {}^{n}C_{x}\, p^{x} q^{n-x}, \quad x = 0, 1, 2, \ldots, n$$

Properties

The discrete random variable of Binomial distribution for $X \sim$ Bin (n, p),

Then

1) Mean: E(X) = μ = *np*.
2) Variance: Var (X) = σ = npq

Example.1:

The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Solution:

The three characters of Binomial are n=4, p=3/4, and

q=1-3/4=1/4

Then $P(X=x)={}^{n}C_{x}p^{x}q^{n-x} \quad x=0,1,2,3,4$

$$P(X=2)={}^{4}C_{2}(\frac{3}{4})^{2}(\frac{1}{4})^{2}$$

$$=\frac{27}{128}$$

Example.2:

Of all the new vehicles of a certain model that are sold, 20% require repairs to be done under warranty during the first year of service. A particular dealership sells 14 such vehicles.

1) What is the probability that fewer than five of them require warranty repairs?

2) What is the probability that more than 2 of the 14 vehicles require warranty repairs?

Solution:

Let X represents the number of vehicles that require warranty repairs. Then

$$X \sim \text{Bin}(14, 0.2).$$

1) The probability that fewer than five vehicles require warranty repairs is

$$P(X \leq 4) = [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)] = 0.870.$$

2) What is the probability that more than 2 of the 14 vehicles require warranty repairs?

Let X represents the number of invoices in the sample that receive discounts.

$$P(X > 2) = 1 - P(X \leq 2).$$

We find that $P(X \leq 2) = 0.448$.

Therefore $P(X > 2) = 1 - 0.448 = 0.552$.

Example.3:

A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%. The inspector randomly picks 20 items from a shipment. What is the probability that there will be at least one defective item among these 20?

Solution

n=20, p=0.03, q=0.97

$$P(X \geq 1) = 1 - P(X = 0)$$

$$P(X \geq 1) = 1 - {}^{20}C_0 (0.03)^0 (0.97)^{20} = 0.4562$$

Example.4:

Refer to the above example. What is the probability that more than or equal 2 of the 10 items are defective?

n=10, p=0.03, q=0.97

$$P(X > 2) = 1 - (P(X = 0) + P(X = 1) + P(X = 2))$$

$$P(X > 2) = 1 - (0.77374 + 0.2281 + 0.0317) = 0.0028$$

1.2. The Poisson distribution

Experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region, are called Poisson experiments. The time interval may be of any length, such as a minute, a day, a week, a month, or even a year. The specified region could be a line segment, an area, a volume, or perhaps a piece of material. In such instances, X might represent the number of field mice per acre, the number of bacteria in a given culture, or the number of typing errors per page. A Poisson experiment is derived from the Poisson process and possesses the following properties:

1. The number of outcomes occurring in one time interval or specified region of space is independent of the number that occur in any other disjoint time interval or region.

2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the length of the time interval.

3. The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.

Probability density function of a Poisson distribution

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

If X is a random variable whose probability mass function is given by the above equation, then X is said to have the **Poisson distribution** with parameter λ. The notation is $X \sim$ Poisson (λ).

Properties

The discrete random variable of Poisson distribution for $X \sim$ Poisson (λ),

Then

1) Mean: $E(X) = \mu = \lambda$.
2) Variance: $Var(X) = \lambda$.

Example.5:

During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution:

Using the Poisson distribution with x = 6 and λ = 4, we have

$$P(X = 6) = \frac{4^6 e^{-4}}{6!} = 0.1042$$

Example.6:

If $X \sim$ Poisson (4), Determine:

(1) $P(X \leq 2)$ **(2)** $P(X > 1)$.

Solution

(a) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$

$$= e^{-4}\, 4^0 / 0! + e^{-4}\, 4^1 / 1! + e^{-4}\, 4^2 / 2!$$

$$= 0.0183 + 0.0733 + 0.1465 = 0.2381$$

(b) $P(X > 1)$, we might try to start by writing

$$P(X > 1) = P(X = 2) + P(X = 3) + \cdot\cdot\cdot$$

$$P(X > 1) = 1 - P(X \leq 1)$$

$$= 1 - [P(X = 0) + P(X = 1)]$$

$$= 1 - (0.0183 + 0.0733) = 0.908$$

Example.7:

The number of email messages received by a computer server follows a Poisson distribution with a mean of 6 per minute. Find the probability that exactly 20 messages will be received in the next 3 minutes.

Solution

Let X is the number of messages received in 3 minutes.

The mean number of messages received in 3 minutes is (6) (3) = 18,

So $X \sim$ Poisson (18).

Using the Poisson (18) probability mass function, we find that

$$P(X = 20) = e^{-18}\, 18^{20} / 20! = 0.0798$$

Remark 1: If X ~ Poisson (λ), then

1. X is a discrete random variable whose possible values are the non-negative integers.
2. The parameter λ is a positive constant.
3. The Poisson probability mass function is very close to the binomial probability mass function when n is large, p is small, and λ = np.

Remark 2: For a discrete random variable X, whose probability density function that is used to defined the mean and variance. Then for two real numbers **a** and **b** that is satisfied:

1) $E(aX \pm b) = a\mu \pm b$

2) $Var.(aX \pm b) = a^2 \sigma^2$

Example.8:

In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles. What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

Solution:

Let X represents the number of bubbles in manufacturing process.

The mean number of bubbles is 1 per 1000,

If we take n=8000, and p=0.001 is binomial distribution

Then $P(X<7)=P(X=0)+P(X=1)+...+P(X=6)=0.3134$

Or

If we the average of bubbles $\mu = \lambda$ =8000*0.001=8 is Poisson distribution

Then $P(X<7)=P(X=0)+P(X=1)+...+P(X=6)=0.3134$

1.3. The Hypergeometric Probability Distribution

The simplest way to view the distinction between the binomial distribution of this chapter and the hypergeometric distribution is to note the way the sampling is done. The types of applications for the hypergeometric are very similar to those for the binomial distribution. We are interested in computing probabilities for the number of observations that fall into a particular category. But in the case of the binomial distribution, independence among trials is required. As a result, if that distribution is applied to, say, sampling from a lot of items (deck of cards, batch of production items), the sampling must be done with replacement of each item after it is observed.

On the other hand, the hypergeometric distribution does not require independence and is based on sampling done without replacement. Applications for the hypergeometric distribution are found in many areas, with heavy use in acceptance sampling, electronic testing, and quality assurance. Obviously, in many of these fields, testing is done at the expense of the item being tested.

That is, the item is destroyed and hence cannot be replaced in the sample. Thus, sampling without replacement is necessary from a finite population of size N elements, r elements in group A and the other N-r elements in group B. Suppose we select n elements from the two groups and the random variable X represent the number of elements selected from group A. The following characterizes the hypergeometric distribution:

1) The result of each draw can be classified into one of two mutually exclusive events (e.g. Pass/Fail or Employed/Unemployed).
2) The probability of a success changes on each draw, as each draw decreases the population.

Probability density function of a Poisson distribution

For a random variable X, the hypergeometric distribution density function is given as:

$$P(X = x_i) = f(x_i) == \frac{\binom{r}{x_i}\binom{N-r}{n-x_i}}{\binom{N}{n}}$$

Where $\binom{r}{x_i}$ is the number of combinations as selecting x_i elements from group A while $\binom{N-r}{n-x_i}$ is the number of combinations as selecting n-x_i elements from group B. $\binom{N}{n}$ is the total number of combinations as selecting n elements from the two groups while $\binom{r}{x_i}\binom{N-r}{n-x_i}$ is the total number of combinations as selecting x_i and n-x_i elements from groups A and B, respectively.

Properties

The discrete random variable of hypergeometric distribution for

$$X \sim \mathrm{H}(x_i; N, n, r)$$

Then

1) Mean: E(X) = $\mu = \frac{nr}{N}$

2) Variance: Var (X) = $\sigma^2 = (\frac{N-n}{N-1})n(\frac{r}{N})(1-\frac{r}{N})$

Example.9:

A particular part that is used as an injection device is sold in lots of 10. The producer feels that the lot is deemed acceptable if no more than one defective is in the lot. Some lots are sampled and the sampling plan involves random sampling and testing 3 of the parts out of 10. If none of the 3 is defective, the lot is accepted. Comment on the utility of this plan.

Solution:

Let us assume that the lot is truly unacceptable (i.e., that 2 out of 10 are defective). The probability that our sampling plan finds the lot acceptable is

$$P(X=0)=\frac{\binom{2}{0}\binom{8}{3}}{\binom{10}{3}}=0.467$$

Thus, if the lot is truly unacceptable with 2 defective parts, this sampling plan will allow acceptance roughly 47% of the time. As a result, this plan should be considered faulty.

Example.10:

Lots of 40 components each are called unacceptable if they contain as many as 3 defectives or more. The procedure for sampling the lot is to select 5 components at random and to reject the lot if a defective is found.

1) What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

2) calculate the mean and variance of the variable X.

Solution:

1) Using the hypergeometric distribution with n = 5, N = 40, r= 3, and x = 1, we find the probability of obtaining one defective to be

$$P(X=1)=\frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}}=0.3011$$

Once again this plan is likely not desirable since it detects a bad lot (3 defectives) only about 30% of the time.

2) Mean: E(X) = $\mu=\frac{nr}{N}=\frac{5*3}{40}=\frac{3}{8}$

Variance: Var (X) = $\sigma^2=(\frac{35}{39})*5*(\frac{3}{40})*(1-\frac{3}{40})=0.311$

Example.11:

Suppose there are 50 officers, 10 female officers and 40 male officers. Suppose 20 of them will be promoted. Let X represents the number of female promotions. Then,

Therefore, the probability distribution function for X is

$$P(X=i)=\frac{\binom{10}{i}\binom{40}{20-i}}{\binom{50}{20}},\ i=0,1,\ldots,10.$$

2. Continuous probability distribution

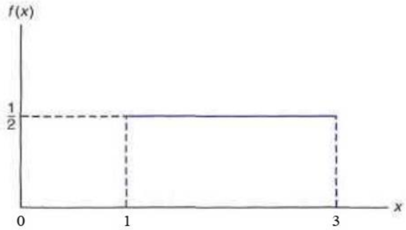
Continuous Uniform Distribution, one of the simplest continuous distributions in all of statistics is the continuous uniform distribution. This distribution is characterized by a density function that is “flat,” and thus the probability is uniform in a closed interval, say [a, b]. Although applications of the continuous uniform distribution are not as abundant as those for other distributions discussed in this chapter, it is appropriate for the novice to begin this introduction to continuous distributions with the uniform distribution.

2.1. The Uniform Distribution

The density function of the continuous uniform random variable X on the Distribution interval [a, b] is

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0 & elsewhere \end{cases}$$

It should be emphasized to the reader that the density function forms a rectangle with **base** *(b— a)* and **constant height** $\frac{1}{b-a}$. As a result, the uniform distribution is often called the **rectangular distribution.** The density function for a uniform random variable on the interval [1, 3] is shown in the Figure



Probabilities are simple to calculate for the uniform distribution due to the simple nature of the density function. However, note that the application of this distribution is based on the assumption that the probability of falling in an interval of fixed length within [a, b] is constant.

Properties

The continuous random variable of uniform distribution for X, with

1) Mean: E(X) = $\mu = \dfrac{a+b}{2}$

2) Variance: Var (X) = $\sigma^2 = \dfrac{(b-a)^2}{12}$

Example.12:

Suppose that a large conference room for a certain company can be reserved for no more than 4 hours. However, the use of the conference room is such that both long and short conferences occur quite often. In fact, it can be assumed that length X of a conference has a uniform distribution on the interval [0, 4].

1) What is the probability density function?

2) What is the probability that any given conference lasts at least 3 hours?

3) Find the mean and variance.

Solution:

1) The appropriate density function for the uniformly distributed random variable X in this situation is

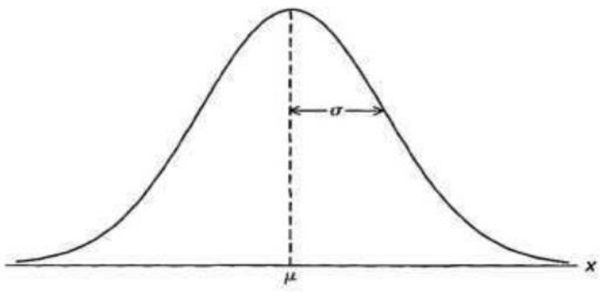
$$f(x) = \begin{cases} \frac{1}{4}, & 0 \le x \le 4 \\ 0 & elsewhere \end{cases}$$

2) $P(X \ge 3) = \int_3^4 \frac{1}{4} dx = 0.25$

3) $\mu = \frac{4}{2} = 2 \quad , \quad \sigma^2 = \frac{(4)^2}{12} = \frac{4}{3}$

2.2. The Normal Distribution

The most important, continuous probability distribution in the entire field of statistics is the normal distribution. Its graph, called the normal curve, is the bell shaped curve: of the present figure, which describes approximately many phenomena that occur in nature, industry, and research. Physical measurements in areas such as meteorological experiments, rainfall studies, and measurements of manufactured parts are often more than adequately explained with a normal distribution. In addition, errors in scientific measurements are extremely well approximated by a normal distribution.



The normal distribution is continuous rather than discrete. The mean of a normal random variable may have any value, and the variance may have any positive value. The probability density function of a normal random variable with mean μ and variance σ^2 is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad , -\infty < x < \infty$$

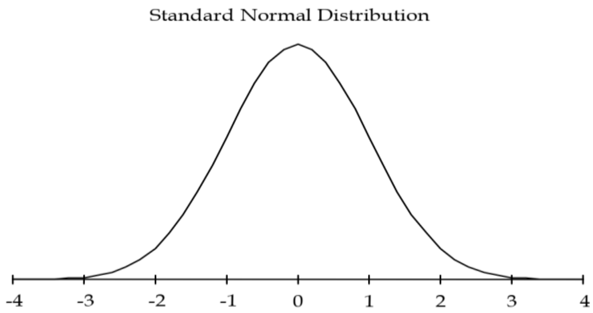
If X is a random variable whose probability density function is normal with mean μ and variance σ^2, we denote as $X \sim N(\mu, \sigma^2)$.

The next figure presents a plot of the normal probability density function with mean μ and standard deviation σ. The normal probability density function is sometimes called the **normal curve**. Note that the normal curve is symmetric around μ.

2.3. The Standard Normal Distribution

A normal random variable with location parameter 0 and scale parameter 1 is called a standard normal random variable. Because of the form of the normal density, it is possible to determine probabilities for any normal random variable from the distribution function of the standard normal variable. Consequently, the standard normal random variable has been given the special symbolic destination, Z, from which the z-score derives. The standard normal distribution function is given the special symbol, $f(z)$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad , -\infty < z < \infty$$



Occasionally, we are required to find the value of z corresponding to a specified probability that falls between values listed in Table STANDARD NORMAL (see convenience, we shall always choose: the z value corresponding to the tabular probability that comes closest, to the specified probability. The preceding two examples were solved by going first; from a value of x to a z value and then computing the desired area. In normal distribution we reverse the process and begin with a known area or probability, find the z value, and then determine x by rearranging the formula: $z = (x - \mu) / \sigma$.

Then Z is a random variable whose probability density function is standard normal with mean *0* and variance *1*, we denote as $Z \sim N(0, 1)$.

Example.13:

Resistances in a population of wires are normally distributed with mean 20m*Ω* and standard deviation 3m*Ω*. The resistance of two randomly chosen wires is 23m*Ω* and 16m*Ω*. Convert these amounts to standard units.

Solution

A resistance of 23m*Ω* is 3m*Ω* more than the mean of 20, and 3m*Ω* is equal to one standard deviation. So 23m*Ω* is one standard deviation above the mean and is thus equivalent to one standard unit. A resistance of 16m*Ω* is 1.33 standard deviations below the mean, so 16m*Ω* is equivalent to −1.33 standard units.

In general, we convert to standard units by subtracting the mean and dividing by the standard deviation. Thus, if *x* is an item sampled from a normal population with mean μ and variance $\sigma 2$, the standard unit equivalent of *x* is the number *z*, where

$$z = (x - \mu) / \sigma$$

The number *z* is sometimes called the "z-score" of *x*. The *z*-score is an item sampled from a normal population with mean 0 and standard deviation 1. This normal population is called the **standard normal population**.

Example.14:

The yield, in grams, of a dye manufacturing process is normally distributed with mean1500 and standard deviation 50. The yield of a particular run is 1568 grams. **Find** 1) The z-Value.

2) The yield of a certain run has a z-score of −1.4. Find the yield in the original units of grams.

Solution

The quantity 1568 is an observation from a normal population with mean $\mu = 1500$ and standard deviation $\sigma = 50$.

1) $z = (1568 - 1500)/50 = 1.36$

2) We use the Equation $z = (x - \mu)/\sigma$ substituting −1.4 for z and solving for x. We obtain

$$-1.4 = (x - 1500)/50$$

Solving for x yields $x = 1430$. The yield is 1430 grams.

Example.15:

Find the area under the normal curve to the right of $z = 1.38$.

Solution

From the z table, the area to the *left* of

$z = 1.38$ is 0.9162.

Therefore the area to the right is

$1 - 0.9162 = 0.0838$

Example.16:

Lifetimes of batteries in a certain application are normally distributed with mean 50 hours and standard deviation 5 hours. **Find** the probability that a randomly chosen battery lasts between 42 and 52 hours.

Solution

Let *X* represent the lifetime of a randomly chosen battery.

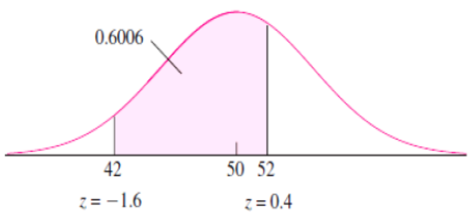
Then *X* is *N (*50, 52*)*. The next Figure presents the probability density function of the *N (*50, 52*)* population.

The shaded area represents *P (*42 < *X* < 52*)*, the probability that a randomly chosen battery has a lifetime between 42 and 52 hours.

To compute the area, we will use the *z* table. First we need to convert the quantities 42 and 52 to standard units.

We have $z = (42 – 50)/5 = -1.60$

$z = (52 – 50)/5 = 0.40$



3. Exercises.3

Ex.1: Let $X \sim$ Binomail $(8, 0.25)$. **Find**

a. $P(X < 2)$ b. $P(X \geq 1)$ c. $P(1 \leq X \leq 2)$

d. $P(2 < X < 6)$ e. $P(X = 0)$ f. $P(X = 8)$

Ex.2: Ten percent of the items in a large lot are defective. A sample of four items is drawn from this lot.

a. **Find** the probability that all of the sampled items are defective.

b. **Find** the probability that one or less of the sampled items is defective.

c. **Find** the probability that exactly two of the sampled items is defective.

d. **Find** the probability that bigger than two of the sampled items are defective.

Ex.3: A warehouse contains nine printing machines, four of which are defective. A company selects four of the machines at random, thinking all are in working condition.

1) What is the probability that less than one defective?

2) What is the probability that all four of the machines are non-defective?

Ex.4: Let $X \sim$ Poisson (5).

Find a. $P(X = 1)$ b. $P(X = 0)$ c. $P(X < 2)$

d. $P(X > 3)$ e. $\mu = E(X)$ f. $Var(2x\text{-}1)$

Ex.5. It is know that 2% of the circuit boards from a production line are defective. If a random sample of 900 circuit boards is taken from this production lines. **Estimate**

a) The probability exactly 2 defective boards.

b) The probability at least 2 defective boards.

Ex.6: In a Poisson distribution, the probability of particular value P(X=0) = 0.7408. **Find** the probability of X bigger than 2.

Ex.7: In a manufacturing process where glass products are made defects occur occasionally rending the piece undesirable for marketing. It is known that the average 0.002 produced. If we take a random sample of 600, what is the probability of 3 items possessing defects at most?

Ex.8: A company is interested in evaluating its current inspection procedure on shipments of 50 identical items. The procedure is to take a sample of 5 and pass the shipment if no more than 2 are found to be defective. What proportion of 20% defective shipments will be accepted?

Ex.9: A manufacturer of automobile tires reports that among a shipment of 5000 sent to a local distributor, 1000 are slightly blemished. If one purchases 10 of these tires at random from the distributor, what, is the probability that exactly 3 are blemished?

Ex.10: A manufacturing company uses an acceptance scheme on production items before they are shipped. The plan is a two-stage one. Boxes of 25 are readied for shipment and a sample of 3 is tested for defectives. If any defectives are found, the entire box is sent back for 100% screening. If no defectives are found, the boxes shipped.

(a) What is the probability that a box containing 3 defectives will be shipped?

(b) What is the probability that a box containing only 1 defective will be sent back for screening?

Ex.11: Scores on a standardized test are approximately normally distributed with a mean of 460 and a standard deviation of 80.

a. **What** proportion of the scores are above 550?

b. **What** is the 35th percentile of the scores?

c. If someone's score is 600**, what** percentile is she on?

d. **What** proportion of the scores is between 420 and 520?

Ex.12: Weights of female cats of a certain breed are normally distributed with mean 4.1 kg and standard deviation 0.6 kg.

a. **Wha**t proportion of female cats has weights between 3.7 and 4.4 kg?

b. A certain female cat has a weight that is 0.5 standard deviations above the mean. **What** proportion of female cats is heavier than this one?

c. **How** heavy is a female cat whose weight is on the 80th percentile?

d. A female cat is chosen at random. **What** is the probability that she weighs more than 4.5 kg?

e. Six female cats are chosen at random. **What** is the probability that exactly one of them weighs more than 4.5 kg?

Ex.13: The life time of a light bulb in a certain application is normally distributed with mean μ = 1400 hours and standard deviation σ = 200 hours.

a. **What** is the probability that a light bulb will last more than 1800 hours?

b. **Find** the 10th percentile of the lifetimes.

c. A particular battery lasts 1645 hours. **What** percentile is its lifetime on?

d. **What** is the probability that the lifetime of a battery is between 1350 and 1550 hours?

Ex.14: At a certain university, math SAT scores for the entering freshman class averaged 650 and had a standard deviation of 100. The maximum possible score is 800. Is it possible that the scores of these freshmen are normally distributed? Explain.

Ex.15: The strength of an aluminum alloy is normally distributed with mean 10 gig pascals (GPa) and standard deviation 1.4 GPa.

a. **What** is the probability that a specimen of this alloy will have strength greater than 12 GPa?

b. **Find** the first quartile of the strengths of this alloy.

c. **Find** the 95th percentile of the strengths of this alloy.

Ex.16: The temperature recorded by a certain thermometer when placed in boiling water (true temperature 100◦C) is normally distributed with mean μ = 99.8◦C and standard deviation 0.1◦C.

a. **What** is the probability that the thermometer reading is greater than 100◦C?

b. **What** is the probability that the thermometer reading is within ±0.05◦C of the true temperature?

Ex.17: The quality assurance program for a certain adhesive formulation process involves measuring how well the adhesive sticks a piece of plastic to a glass surface. When the process is functioning correctly, the adhesive strength X is normally distributed with a mean of 200 N and a standard deviation of 10 N. Each hour, you make one measurement of the adhesive strength. You are supposed to inform your supervisor if your measurement indicates that the process has strayed from its target distribution.

a. **Find** $P(X \leq 160)$, under the assumption that the process is functioning correctly.

b. Based on your answer to part (a), if the process is functioning correctly, would a strength of 160 N be unusually small? Explain.

c. If you observed an adhesive strength of 160 N, would this be convincing evidence that the process was no longer functioning correctly? Explain.

d. **Find** $P(X \geq 203)$, under the assumption that the process is functioning correctly.

e. Based on your answer to part (d), if the process is functioning correctly, would strength of 203 N be unusually large? Explain.

f. If you observed an adhesive strength of 203 N, would this be convincing evidence that the process was no longer functioning correctly? Explain.

g. **Find** $P(X \leq 195)$, under the assumption that the process is functioning correctly.

h. Based on your answer to part (g), if the process is functioning correctly, would strength of 195 N be unusually small? Explain.

i. If you observed an adhesive strength of 195 N, would this be convincing evidence that the process was no longer functioning correctly? Explain.

Chapter 4

Linear Regression and Correlation

1. Introduction

The word correlation is used in everyday life to denote some form of association. We might say that we have noticed a correlation between stress and strain of material. However, in statistical terms we use correlation to denote association between two quantitative variables. We also assume that the association is linear, that one variable increases or decreases a fixed amount for a unit increase or decrease in the other. The other technique that is often used in these circumstances is regression, which involves estimating the best straight line to summarize the association. When you have completed this chapter, you will be able to:

1-: Draw a scatter diagram.

2-: Understand and interpret the terms *dependent* variable and *independent* variable.

3-: Calculate and interpret the coefficient of correlation, the coefficient of determination, and the standard error of estimate.

4-: Calculate the least squares regression line and interpret the slope and intercept values.

In this chapter we will first discuss correlation analysis, which is used to quantify the association between two continuous variables (e.g., between an independent and a dependent variable or between two independent variables). Regression analysis is a related technique to assess the relationship between an outcome variable and one or more risk factors or confounding variables. The outcome variable is also called the **response** or **dependent variable** and the risk factors and confounders are called the **predictors**, or **explanatory** or **independent variables**. In regression analysis, the dependent variable is denoted "*y*" and the independent variables are denoted by "*x*".

2. <u>Correlation Analysis</u>

Correlation Analysis is a group of statistical techniques used to measure the strength of the association between two variables.

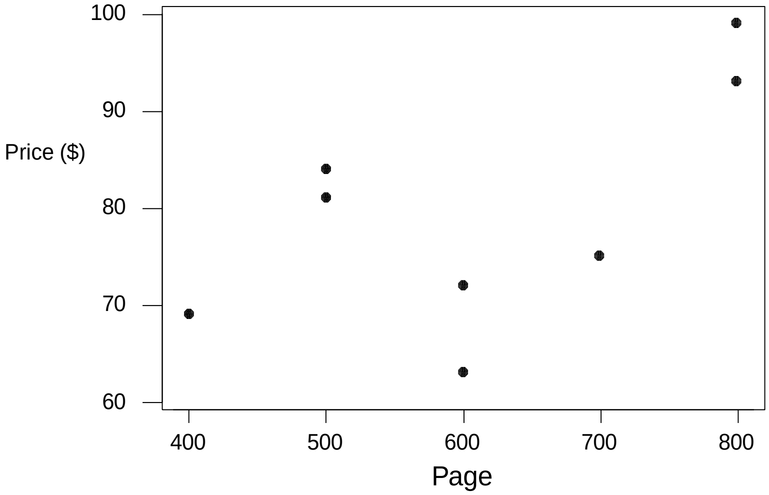
A Scatter Diagram is a chart that portrays the relationship between the two variables.

The Dependent Variable is the variable being predicted or estimated.

The Independent Variable provides the basis for estimation. It is the predictor variable.

i.e The relationship between the number of pages and selling price of text

Scatter Diagram of Number of Pages and Selling Price of Text

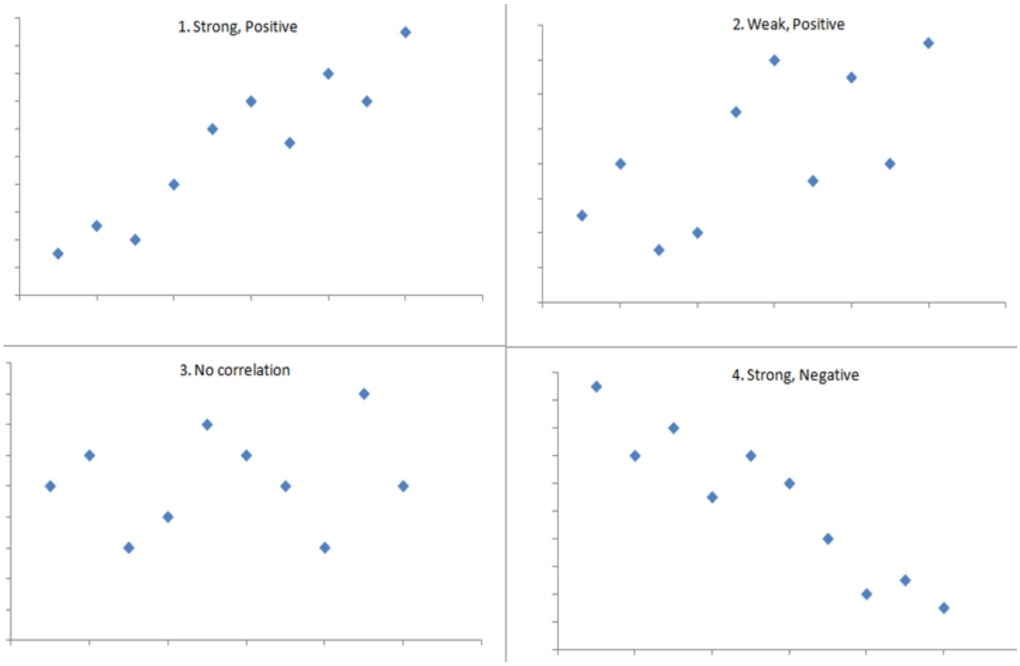


<u>The Coefficient of Correlation</u> *<u>(r)</u>*

is a measure of the strength of the relationship between two variables. The characteristics of the coefficient of correlation are:

1) It requires ratio-scaled data
2) It can range from -1 to 1
3) Values of -1.00 or 1.00 indicate perfect and strong correlation.
4) Values close to 0.0 indicate weak correlation.

5) Negative values indicate an inverse relationship and positive values indicate a direct relationship.



The coefficient of correlation will be Calculated from the following formulas:

i. Spearman's Coefficient Correlation

Spearman's r is a statistic for measuring the relationship between two qualitive or quantative variables. It is a nonparametric measure that avoids assumptions that the variables have a straight line relationship and can be used when one or both measures is measured on an ordinal scale. The difference is that

$$r = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}.$$

Spearman's r refers to the *ranked values* rather than the original measurements.

Where the n is number of order pairs and $\mathbf{d}^2$ is the difference between rank of x and rank of y

Example 1:

The following observations are selected from 8 students in course of statistics the observation measure the course degree and total GPA after the final exam

Degree	A^+	*D*	A^{--}	*B*	A^+	B^{--}	B^+	*F*
G.P. A	*2.8*	*2.8*	*3.2*	*3.0*	*3.2*	*2.9*	*3.2*	*1.99*

Compute the spearman's coefficient correlation.

Solution:

X	Y	R(X)	R(Y)	d^2
A^+	2.8	**7.5**	**2.5**	**2.5**
D	2.8	**2**	**2.5**	**0.25**
A^-	3.2	**6**	**7**	**1**
B	3.0	**4**	**5**	**1**
A^+	3.2	**7.5**	**7**	**0.25**
B^-	2.9	**3**	**4**	**1**
B^+	3.2	**5**	**7**	**4**
F	1.99	**1**	**1**	**0**
				32.5

$$r = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}.$$

$$\therefore r = 1 - \frac{6(32.5)}{8(63)} = 0.61$$

Strong positive correlation coefficient

Example 2:

The data below gives the marks of 10 students in physics and math

Math	***20***	***23***	***8***	***29***	***14***	***11***	***11***	***20***	***17***	***17***
Physics	***30***	***35***	***21***	***33***	***33***	***16***	***22***	***31***	***33***	***36***

Calculate the spearman's correlation coefficient.

Solution:

X	Y	Order(x)	Order(y)	d	(d^2)
20	30	7.5	4	3.5	12.25
23	35	9	9	0	0
8	21	1	1	0	0
29	33	10	7	3	9
14	33	4	7	-3	9
11	16	2.5	3	-0.5	0.25
11	22	2.5	2	0.5	0.25
20	31	7.5	5	2.5	6.25
17	33	5.5	7	-1.5	2.25
17	36	5.5	10	-4.5	20.25
					59.5

$$r = 1 - \frac{6\Sigma d^2}{n(n^2 - 1)}.$$

$$\therefore r = 1 - \frac{6(59.5)}{10(99)} = 0.639$$

Strong positive correlation coefficient.

ii. *Pearson's Coefficient of Correlation*

There is a simple and straightforward way to measure correlation between two variables. It is called the Pearson correlation coefficient (r). It's longer name, the *Pearson product-moment correlation*, is sometimes used.

The formula for computing the Pearson r is as follows:

$$r = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{\sqrt{\left[n(\Sigma X^2) - (\Sigma X)^2\right]\left[n\left(\Sigma Y^2\right) - \left(\Sigma Y\right)^2\right]}}$$

b. Coefficient of Determination

The coefficient of determination (r^2) is the proportion of the total variation in the dependent variable (Y) that is explained or accounted for by the variation in the independent variable (X).

3. Regression Analysis

In regression analysis, we use the independent variable (*X*) to estimate or predict the dependent variable (*Y*).

- The relationship between the variables is linear.
- Both variables must be at least interval scale.
- The least squares criterion is used to determine the equation.
  That is the term $(Y - Y^*)^2$ is minimized. (Y is observed variable and Y* is estimated variable).

The regression equation:

At its simplest level, linear regression is a method for fitting a straight line through an x-y scatter plot. Recall from other math courses that a straight line is described by the following formula:

$$Y^* = a + bX,$$

where:

- Y^* is the average predicted value of *Y* for any *X*.
- *a* is the *Y*-intercept. It is the estimated *Y* value when $X=0$
- *b* is the slope of the line, or the average change in *Y'* for each change of one unit in *X*
- The least squares principle is used to obtain *a* and *b*.

The least squares principle is used to obtain *a* and *b*. The equations to determine *a* and *b* are:

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2}$$

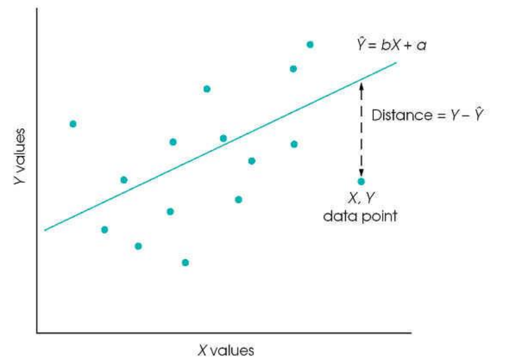
$$a = \bar{Y} - b\bar{X} = \frac{\sum Y}{n} - b\frac{\sum X}{n}$$

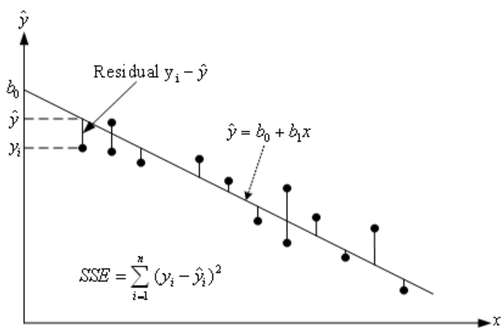
<u>The standard error of estimate measures the scatter, or dispersion, of the observed values around the line of regression</u>

The formulas that are used to compute the standard error

$$S_{y.x} = \sqrt{\frac{\sum (Y - Y^*)^2}{n-2}}$$

The difference $(y - y^*)$ is called a residual, and their sum is called the residual sum of squares or sum of squared errors (SSE).





<u>Example 3</u>

A road transport company might want to asses the relationship between the age of a vehicle and its yearly maintenance cost and it might take a random sample of observation such as:

Age of vehicle(years)	***5***	***3***	***8***	***2***	***11***	***7***
maintenance cost (1000 LE)	***3.5***	***2***	***7.5***	***3***	***15***	***9***

1) Determine the coefficient correlation (using pearson method)
2) Estimate the maintenance cost for vehicle with age 4 years.
3) Compute the standard error estimation.

Solution:

X	**Y**	**XY**	$\mathbf{X^2}$	$\mathbf{Y^2}$
5	3.5	17.5	25	12.25
3	2	6	9	4
8	7.5	60	64	56.25
2	3	6	4	9
11	15	165	121	225
7	9	63	49	81
36	40	317.5	272	387.5

1) Pearson's coefficient correlation:

$$r=\frac{n(\Sigma XY)-(\Sigma X)(\Sigma Y)}{\sqrt{\left[n(\Sigma X^2)-(\Sigma X)^2\right]\left[n\left(\Sigma Y^2\right)-(\Sigma Y)^2\right]}}$$

$$r=\frac{6(317.5)-(36)(40)}{\sqrt{\left[6(272)-(36)^2\right]\left[6(387.5)-(40)^2\right]}}=0.942$$

Strong positive correlation coefficient

2) Regression equation predicted the maintenance cost on vehicle age:

Let x=vehicle age and y= maintenance cost

$Y^{*} = a + bX,$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n(\sum X^2) - (\sum X)^2} = 1.384$$

$$a = \bar{Y} - b\bar{X} = \frac{\sum Y}{n} - b\frac{\sum X}{n} = \frac{40}{6} - 1.384(\frac{36}{6}) = -1.637$$

$Y^{*} = -1.637 + 1.384X$

At X=4 then Y=3.89~4 (thous LE)

3) Standard error estimation:

$$S_{y.x} = \sqrt{\frac{\sum (Y - Y^{*})^2}{n-2}} = \sqrt{\frac{11.045}{6-2}} = 1.662$$

Y	$(Y - Y^{*})^2$
5.283	3.179
2.515	0.265
79.435	3.744
1.131	0.755
13.587	1.997
8.051	1.105
Sum	11.045

4. Exercises 4

1) A recent article in engineering week listed the best small companies. We are interested in the current results of the company's sales and earning with million dollars. A random sample of 7 companies was selected as in the following table.

Sales(m.$)	***15***	***13***	***21***	***17***	***18***	***61***	***24***
Earning (m.$)	***25***	***15***	***29***	***31***	***26***	***81***	***31***

1- Compute the Pearson coefficient correlation.

2- Estimated the Sales for company with Earning 50 m.$.

3- Determine the standard error of estimate.

4- **Draw** the scatter diagram

2) The production department of National Electronics wants to explore the relationship between the number of employees who assemble and the number produced. As an experiment the complete set of paired observation follows (Lets the dependent variable is production).

Number. of employee	***2***	***4***	***1***	***6***	***5***	***3***
Production (unit/ 1 hr)	***15***	***25***	***10***	***38***	***40***	***30***

1- **Draw** the scatter diagram.

2- **Determine** the Regression equation.

3- **Find** the standard error of estimate as a measure of fitting the regression line.

4- Find the coefficient determination.

3) The production department of National Electronics wants to explore the relationship between the numbers of employees (Y) who work in the produced lines (X). As an experiment, the complete set of paired observation follows

X	*1st*	*3rd*	*4th*	*2nd*	*5th*
Y	*12*	*10*	*23*	*17*	*16*

a) Find the correlation coefficient, showing the type of the correlation.
b) Draw the scatter diagram between numbers of employees and the produced lines.

REFERENCES

1) Probability & Statistics for Engineers and Scientists, By R. E. Walpole, R. H. Myers, S. L. Myers and K. Ye., 9^{TH} Edition.
2) Statistics: concepts and applications, By Harry Frank and Steven C Althoen, Cambridge University Press, 2004.
3) Introduction to Probability, By Charles M. Grinstead and J. Laurie Snell, Swarthmore College.