COURSE MANUAL

Distribution Theory I

STA 312

University of Ibadan Distance Learning Centre
Open and Distance Learning Course Series Development

ISBN 978-021-360-0

General Editor: Prof. Bayo Okunade

University of Ibadan Distance Learning Centre
University of Ibadan,
Nigeria

Telex: 31128NG

Tel: +234 (80775935727)
E-mail: ssu@dlc.ui.edu.ng
Website: www.dlc.ui.edu.ng

Vice-Chancellor's Message

The Distance Learning Centre is building on a solid tradition of over two decades of service in the provision of External Studies Programme and now Distance Learning Education in Nigeria and beyond. The Distance Learning mode to which we are committed is providing access to many deserving Nigerians in having access to higher education especially those who by the nature of their engagement do not have the luxury of full time education. Recently, it is contributing in no small measure to providing places for teeming Nigerian youths who for one reason or the other could not get admission into the conventional universities.

These course materials have been written by writers specially trained in ODL course delivery. The writers have made great efforts to provide up to date information, knowledge and skills in the different disciplines and ensure that the materials are user-friendly.

In addition to provision of course materials in print and e-format, a lot of Information Technology input has also gone into the deployment of course materials. Most of them can be downloaded from the DLC website and are available in audio format which you can also download into your mobile phones, IPod, MP3 among other devices to allow you listen to the audio study sessions. Some of the study session materials have been scripted and are being broadcast on the university's Diamond Radio FM 101.1, while others have been delivered and captured in audio-visual format in a classroom environment for use by our students. Detailed information on availability and access is available on the website. We will continue in our efforts to provide and review course materials for our courses.

However, for you to take advantage of these formats, you will need to improve on your I.T. skills and develop requisite distance learning Culture. It is well known that, for efficient and effective provision of Distance learning education, availability of appropriate and relevant course materials is a *sine qua non*. So also, is the availability of multiple plat form for the convenience of our students. It is in fulfilment of this, that series of course materials are being written to enable our students study at their own pace and convenience.

It is our hope that you will put these course materials to the best use.

Prof. Abel Idowu Olayinka
Vice-Chancellor

Foreword

As part of its vision of providing education for "Liberty and Development" for Nigerians and the International Community, the University of Ibadan, Distance Learning Centre has recently embarked on a vigorous repositioning agenda which aimed at embracing a holistic and all encompassing approach to the delivery of its Open Distance Learning (ODL) programmes. Thus we are committed to global best practices in distance learning provision. Apart from providing an efficient administrative and academic support for our students, we are committed to providing educational resource materials for the use of our students. We are convinced that, without an up-to-date, learner-friendly and distance learning compliant course materials, there cannot be any basis to lay claim to being a provider of distance learning education. Indeed, availability of appropriate course materials in multiple formats is the hub of any distance learning provision worldwide.

In view of the above, we are vigorously pursuing as a matter of priority, the provision of credible, learner-friendly and interactive course materials for all our courses. We commissioned the authoring of, and review of course materials to teams of experts and their outputs were subjected to rigorous peer review to ensure standard. The approach not only emphasizes cognitive knowledge, but also skills and humane values which are at the core of education, even in an ICT age.

The development of the materials which is on-going also had input from experienced editors and illustrators who have ensured that they are accurate, current and learner-friendly. They are specially written with distance learners in mind. This is very important because, distance learning involves non-residential students who can often feel isolated from the community of learners.

It is important to note that, for a distance learner to excel there is the need to source and read relevant materials apart from this course material. Therefore, adequate supplementary reading materials as well as other information sources are suggested in the course materials.

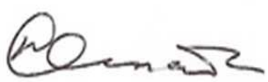
Apart from the responsibility for you to read this course material with others, you are also advised to seek assistance from your course facilitators especially academic advisors during your study even before the interactive session which is by design for revision. Your academic advisors will assist you using convenient technology including Google Hang Out, You Tube, Talk Fusion, etc. but you have to take advantage of these. It is also going to be of immense advantage if you complete assignments as at when due so as to have necessary feedbacks as a guide.

The implication of the above is that, a distance learner has a responsibility to develop requisite distance learning culture which includes diligent and disciplined self-study, seeking available administrative and academic support and acquisition of basic information technology skills. This is why you are encouraged to develop your computer skills by availing yourself the opportunity of training that the Centre's provide and put these into use.

In conclusion, it is envisaged that the course materials would also be useful for the regular students of tertiary institutions in Nigeria who are faced with a dearth of high quality textbooks. We are therefore, delighted to present these titles to both our distance learning students and the university's regular students. We are confident that the materials will be an invaluable resource to all.

We would like to thank all our authors, reviewers and production staff for the high quality of work.

Best wishes.



Professor Bayo Okunade

Director

Course Development Team

Content Authoring	Adedayo Adepoju
Content Editor	Prof. Remi Raji-Oyelade
Production Editor	Ogundele Olumuyiwa Caleb
Learning Design/Assessment Authoring	Folajimi Olambo Fakoya
Managing Editor	Ogunmefun Oladele Abiodun
General Editor	Prof. Bayo Okunade

Table of Contents

Study Session 1: Basic Concepts of Probability

Expected duration: 1 week or 2 contact hours

Introduction

It will prove useful to begin this course by defining some basic concepts of probability. One of the fundamental tools of statistics is probability, which had its formal origin with games of chance in the seventeenth century. Probability theory is a fascinating subject which can be studied at a variety of intellectual and mathematical levels. Probability lies at the foundation of statistical theory and application.

Learning Outcomes from Study Session

At the end of this study session, you should be able to:

1.1 Highlight of Basic Set Theory
1.2 Explain Basic concept of Probability
1.3 Enumerate on Counting Techniques
1.4 Discuss Combination

1.1 Basic Set Theory

The following are basic set theory:

(1) Random Experiment: A random experiment is an experiment in which

a. All the outcomes of the experiment know in advance;

b. Any performance of the experiment results in an outcome that is not known in advance; and

c. The experiment can be repeated in an identical condition.

Simply put, any experiment that can have more than one possible outcome or result is called a random experiment.

Example 1:

- A coin is tossed once; the result of the experiment is either head or tail.
- A dice is rolled once, the possible outcomes are 1, 2, 3, 4, 5, or 6.

2. Set: A set can be defined as any well-defined collection of objects. Individual objects that belong to a set are called members or elements of the set.

A set of vowels is {a, e, i, o, u}

A set of animals {dog, tiger, elephant}

A SET is denoted by upper case letters and an element, a lower case.

i. A unit set is a set composed of only one element.

ii. A set that contains no elements is called the empty set, or null set, and is designated by the symbol ϕ.

3. **Sample Space:** A set S which consists of all possible outcomes of a random experiment is called a sample space.

   $S = \{H, T\}$ $S = \{1, 2, \ldots, 6\}$

4. **Events:** An "event" is a subset of the sample space.

   i. An elementary event is a single possible outcome of an experimental trial. It is an event which cannot be subdivided into a combination of other events.

   e.g. $A = \{1\}$ $B = \{5\}$

   ii. Compound or composite event is an event that can be subdivided into small events.

   e.g. $A = \{1, 3, 5\}$ $B = \{2, 4, 6\}$

5. **Subset:** A collection made up of some of the objects in a set is called a subset.

1.2 Basic Concepts of Probability

Since probability originated from games of chance, actions such as the following are familiar in the theory of probability; tossing a coin, throwing a dice, spinning a roulette wheel, drawing a card etc.

Here, the outcome of a trial is uncertain, however, it is recognized that even though, the outcome of a trial is uncertain, there is a predictable long-term outcome (relative frequency). It is known, for example, that in many throws of an ideal (balanced, symmetrical) coin about half of the trials will result in heads. This concept of probability may be defined and interpreted in several different ways, the chief ones arising from the following:

1.2.1 Classical (A Priori) Approach to Probability

In this approach, the total number of all possible outcomes is fixed and known prior to the performance of any experiment. Similarly, the number of outcomes that have the particular characteristic associated with the event in question is fixed and known beforehand.

All the outcomes are mutually exclusive and equally likely.In this situation, the probability that an event occurs will be defined as the ratio of:

- ❖ The number of outcomes (results) in an experiment that have the characteristic associated with the event.
- ❖ The total number of possible outcomes of the experiment.

Therefore, the probability of a given event can be determined without necessarily performing the experiment.

That is, for event E in the sample space S.

$$P(E) = \frac{\text{No. of results in E}}{\text{No. of relsults in S}} = \frac{n(E)}{n(S)}$$

That is, if there are E possible outcomes favorable to the occurrence of an event A and F possible outcomes unfavorable to the occurrence of this event, then the probability that A will occur is

$$P(A) = \frac{E}{F+E} = \frac{\text{No. of outcomes favorable to A}}{\text{Total No. of possible outcomes}}$$

Example 2:

1. If a fair 6-sided dice is rolled, the probability that 1 will be observed is equal to 1/6, and is the same for other five sides.
2. If a card is picked at random from a well shuffled deck of ordinary playing cards, the probability of picking a heart is $^{13}/_{52}$.

1.2.2 Relative Frequency (A Posteriori) Approach to Probability

This is when the probability ratio associated with a given event is not known before experiments are performed. Rather, the probability ratio is determined only after a relatively long period has passed in which many experiments have occurred under identical conditions. In this type of situation, the probability that as an event may occur is defined as the ratio of:

- ❖ The number of experiments in which the specified event occurred to
- ❖ The total number of experiments performed.

The two basic assumptions underlying this definition are that:

- ❖ A relatively large number of experiments are performed under identical conditions and
- ❖ As the total number of experiments is increased, the probability ratio approaches a given value.Consider an experiment in which there are independently repeated trials.

The number of outcomes "f" of an event A in which we are interested is recorded in n trials of experimental. Then, the relative frequency of occurrence of A is

$$P(A) = {}^{f}/_{n}$$

Example 3

An experiment has a box containing N balls, where N is a large but unknown number. Some of the balls may be red. The experiment defines the event E as "obtain a red ball" and wishes to assign a probability p to this event.

Note: Since the number of results is not known and the number of favorable results is unknown, the experiments cannot apply the classical definition of probability. He therefore decides to draw a sample of balls from the box, making sure that every ball has equal chance of being selected.

⇒He draws a total of 10 balls from the box, and 4 are red so

$$f/n = 4/10 \cong \rho = 0.4$$

⇒He draws and 100 balls. 38 are red.

$$f/n = 38/100 \approx \rho = 0.38$$

⇒He draws 1000 balls and 388 are red.

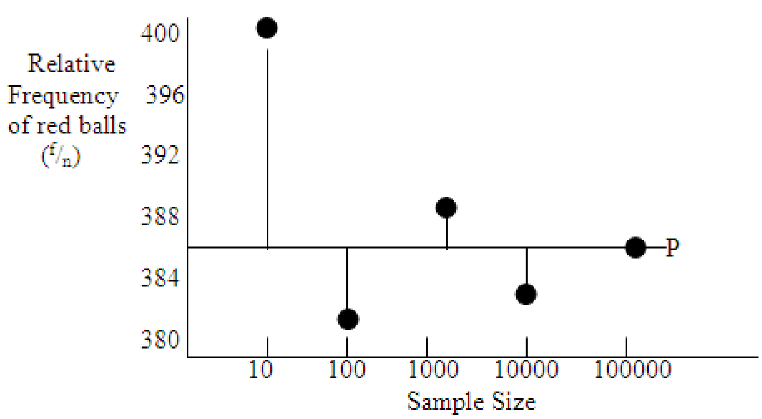
$$f/n = 388/1000 = 0.388$$

⇒10,000, 3840 are red

$$3840/10000 = 0.3840$$

⇒ Draws 100,000 and

$$38500/100000 = 0.385 = \rho$$



1.2.3 Mathematical (Axiomatic) Definition of Probability

The basis of this approach is embodied in three properties from which the whole system of probability theory is constructed through the use of mathematical logic.

The mathematical definition of probability is given as follows:

Let S be the sample space of an experiment. Then the probability is a function P that assigns real numbers to events in such a way that:

i. $0 \leq P(E) \leq 1$ for any event E

ii. $P(E_1 \cup E_2 \cup \ldots) = P(E_1) + P(E_2) + \ldots$ for any collection of mutually exclusive events $\{E_1, E_2, \ldots\}$

iii. $P(S) = 1$

1.3 Counting Techniques – Permutations and Combinations

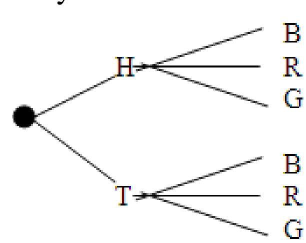
These techniques are helpful in computing the probability of an event when the total number of possible events are large.

1.3.1 Multiplication Principle

If one operation can be performed in n_1 ways and a second operation can be performed in n_2 ways, then there are $n_1 \cdot n_2$ ways in which both operations can be carried out.

Example 4

Suppose a coin is tossed and then a marble is selected at random from a box containing one black, one red, and one green marble. The possible outcome are HB, HR, HG, TB, TR, and TG for each of the two possible outcomes of the coin there are three marbles that may be selected for a total of 2 x 3 = 6 possible outcomes.



Note that the multiplication principle can be extended to more than two operations. In particular, if the i^{th} of r successive operations can be performed in n_i ways, the total number of ways to carry out all r operations is the product.

$$\prod_{i=1}^{r} n_i = n_1 n_2 \ldots n_r$$

1.3.1 Permutations

A permutation is an ordered arrangement of objects. The number of permutations of n distinct objects taken n at a time is

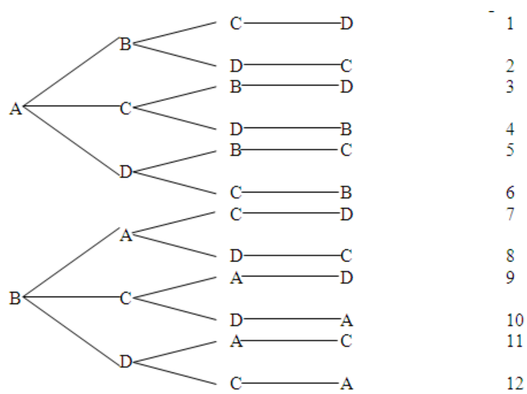
$$nP_r = \frac{n!}{(n-r)!}$$

Proof: In order to fill r positions of n objects, the first position may be filled in n ways using n objects, the second position may be filled with n – 1 ways, and so on until there are n – (r – 1) objects left to fill in the r^{th} position. Thus, the total number of ways of carrying out this operation is

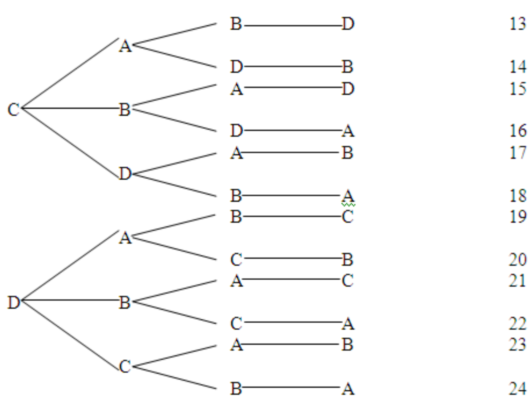
$$n.(n-1).(n-2)......(n-(r-1)) = \frac{n!}{(n-r)!}$$

The possible permutations of 4 objects taken 4 at a time is

Ways to fill	Ways to fill	Ways to fill	Ways to fill	counting the no. of
1^{st} Position	2^{nd} Position	3^{rd} Position	4^{th} Position	arrangements

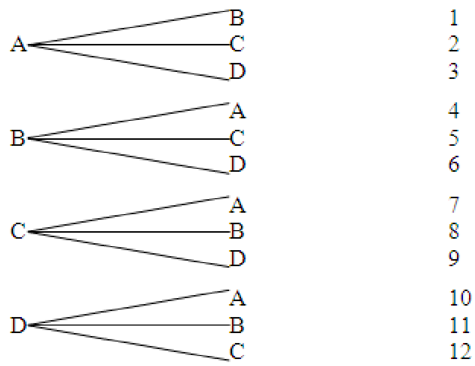


Ways to fill	Ways to fill	Ways to fill	Ways to fill	counting the no. of
1^{st} Position	2^{nd} Position	3^{rd} Position	4^{th} Position	arrangements



Suppose we have only two positions available on the shelf. In how many ways can we fill these two positions using four objects? 1st Determine the number of possible permutations of 4 things taken two at a time. Let the 4 objects be A, B, C, D.Four objects with which to fill the 1st position, once that has been filled, we have 3 objects, only using a tree diagram.

Ways to fill the 1st Position | Ways to fill the 2nd Positions | Counting the number of permutations



From above figure, there are 4 x 3 = 12 possible permutations of 4 takes 2 at a time.Designate the number of distinct objects by n from which the ordered arrangement is to be derived, and by r the number of objects in the arrangement. The number of possible such ordered arrangements is referred to as the number of permutations of n things taken r at a time, and written as nPr.

In general:

$$nPr = n(n-1)(n-2) \ldots (n-r+1)$$

nPr can be evaluated by means of a fraction involving factorials as follows:

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

Example 5

In a county health department, there are five adjacent offices to be occupied by five nurses, A, B, C, D and E. In how many ways can the five nurses be assigned to the offices?

$${}_5P_5 = \frac{5!}{(5-5)!} = 5 \text{ x } 4 \text{ x } 3 \text{ x } 2 \text{ x } 1 = 120.$$

Suppose that there were six nurses to whom adjacent offices are to be assigned, out of only four offices available.We need to determine the number of permutations of six thing taken four at a time which is

$$6P_4 = \frac{6!}{(6-4)!} = \frac{6 \text{ x } 5 \text{ x } 4 \text{ x } 3 \text{ x } 2}{2!} = 360$$

Permutation: Indistinguishable objects; objects that are not all different.

$$n! = \left(nP_{n_1}, n_2, \ldots\ldots, n_1\right) n_1! n_2! \ldots . n_k!$$

1.4 Combinations

A combination is an arrangement of objects without regard to order. The number of combinations of n distinct objects chosen r at a time is

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

Proof: nP_r may be interpreted as the number of ways of choosing r objects from n objects and then permuting the r objects r! ways.

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

Permutations of 4 objects taken two at a time consisted of

AB	AC	AD	BC	BD	CD
BA	CA	DA	CB	DB	DC

Whereas, there are only six combinations, i.e. there a 2 permutations of each combination.In certain cases we may not want to make a distinction between arrangements AB and BA for example, we may want to consider them as the same subset, we say that order does not count, and refer to the arrangements as combinations.In general, we have r! Permutations for each combination of n objects taken r at a time.

$$nP_r = r! \binom{n}{r}$$

$$\binom{n}{r} = \frac{nP_r}{r!}$$

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

e.g

$$\binom{4}{2} = \frac{4!}{2!\ 2!}$$

$$= \frac{4 \text{ x } 3 \text{ x } 2!}{2!\ 2!} = 6$$

Example 6

Suppose that a group therapy leader in a mental health clinic has 10 patients from which to form a group of six. How many combinations of patients are possible?

$$\binom{10}{6} = \frac{10!}{6!\ 4!} = 210$$

Example 7: Combination

A steering committee of 7 is to be chosen at random from a club with 40 members. How many committees can be formed?

$$40C_7 = \frac{40!}{33!\ 7!} = 18{,}643{,}560$$

1.4.1 Properties of Probability

1. If A is an event and A′ is its complement, then

   $$P(A) = 1 - P(A')$$

   Proof: $S = A \cup A'$ since $A \cap A' = \phi$ A and A′ are mutually exclusive

   So

   $1 = P(S) = P(A \cup A') = P(A) + P(A')$

2. For any event A, $P(A) \leq 1$.

   Proof: $P(A) = 1 - P(A')$, we know that $P(A') \geq 0$

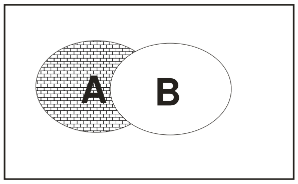
   $\therefore P(A) \leq 1$.

3. For any two events A and B

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

   Proof: $A \cup B = (A \cap B') \cup B$

   $A = (A \cap B) \cup (A \cap B')$



ie $(A \cup B) \cap (B' \cup B)$

$(A \cup B) \cap S = A \cup B$

$(A \cap B) \cup (A \cap B') = A \cap (B \cup B') = A \cap S$

$A = (A \cap B) \cup (A \cap B')$

Then it follows that $(A \cap B')$ and B are mutually exclusive

Since $(A \cap B') \cap B = (A \cap B) \cap (B' \cap B) = \phi$

Since $A \cap B = \phi$

$B' \cap B = \phi$

Then $P(A \cup B) = P(A \cap B') - P(B)$

Similarly, $A \cap B$ and $A \cap B'$ are mutually exclusive, so

$P(A) = P(A \cap B) + P(A \cap B')$

$\therefore P(A \cap B') = P(A) - P(A \cap B)$

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

For any three events, A, B, and C

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

$+ P(A \cap B \cap C)$.

1.4.2 Conditional Probability

A major objective of probability modeling is to determine how likely it is that an event A will occur when a certain experiment is performed. However, there are numerous cases in which the probability assigned to A will be affected by knowledge of the occurrence or non-occurrence of another event B. In such a case, we use "Conditional Probability of A given B" and write as P(A/B).

Definition: The Conditional Probability of an event A, given the event B, is defined by

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

If $P(B) \neq 0$

Example 8

Two cards are drawn without replacement from a deck of cards.

Let A_1 denote the event of getting "an ace on the first draw" and

A_2 denote the event of getting "an ace on the second draw"

The number of ways in which different outcomes can occur can be enumerated as follows. (Multiplicative principle is used).

	A_1	A_1'	
A_2	4 x 3	48 x 4	**4 x 51**
A_2'	4 x 48	48 x 47	**48 x 51**
	4 x 51	**48 x 51**	52 x 51

i. The probability of getting "an ace on the first draw and an ace on the second draw" is given by

$$P(A_1 \cap A_2) = \frac{4 \text{ x } 3}{52 \text{ x } 51}$$

ii. What is the probability that an ace is drawn on the second drawn given that an ace was obtained on the first draw?

$$P(A_2/A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{(4 \text{ x } 3)/(52 \text{ x } 51)}{(4 \text{ x } 51)/(52 \text{ x } 51)} = \frac{3}{51}$$

1.4.3 Independent Events

In some situations, knowledge that an event A has occurred will not affect the probability that an event B will occur is P(B/A) = P(B) That is $P(A \cap B) = P(A)P(B/A) = P(A)\,P(B)$

In general, when this happens, the two events are said to be independent or stochastically independent.

Definition: Two events A and B are called independent events if

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

Otherwise, A and B are called dependent events. Dependent events occur in connection with repeated sampling without replacement from a finite collection.

Example 9

If two cards are drawn in succession from a deck, what is the probability of ace on the 1st draw and ace on the 2nd draw.

$P(A_2) = 4/52$ $\quad P(A_2/A_1) = 3/51$ the events are dependent.

But if the sampling is with replacement; then the draws are independent trials.

$$P(A_1 \cap A_2) = P(A_1)\,P(A_2)$$

i.e. $P(A_2) = 4/52 \quad P(A_2/A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)} = 4/52$

Example 10

In a certain high school class, consisting of 60 girls and 40 boys, it is observed that 24 girls and 16 boys wear eyeglasses. If a student is picked at random from the class, what is the probability that the student wears eye glasses?

$$P(E) = 40/100 = 0.4$$

What is the probability that a student picked at random wears eye glasses, given that the student is a boy?

$$P(E/B) = \frac{P(E \cap B)}{P(B)} = \frac{16/100}{40/100}$$

i.e $\frac{16/40 \text{ x } 40/100}{40/100} = 0.4$

Thus, the additional information that the student is a boy does not alter the probability that the student wears eyeglasses.

: . P(E) = P(E/B)

The event being a boy and wearing glasses for this example are independent. Show that the event of wearing eye-glasses, E_1 and not being a boy $\bar{B}$ are also independent.

$$P(E/\bar{B}) = \frac{P(E \cap \bar{B})}{P(B)} = \frac{24/100}{60/100} = \frac{24}{60} = 0.4$$

Summary

In this study session you have learnt about:

(1) Basic Set Theory

The following are basic set theory:

- **Random Experiment:** A random experiment is an experiment in which
  a. All the outcomes of the experiment know in advance;
  b. Any performance of the experiment results in an outcome that is not known in advance; and the experiment can be repeated in an identical condition.
  c. Simply put, any experiment that can have more than one possible outcome or result is called a random experiment.

(2) Basic Concepts of Probability

Since probability originated from games of chance, actions such as the following are familiar in the theory of probability; tossing a coin, throwing a dice, spinning a roulette wheel, drawing a card etc.

Here, the outcome of a trial is uncertain, however, it is recognized that even though, the outcome of a trial is uncertain, there is a predictable long-term outcome (relative frequency). It is known, for example, that in many throws of an ideal (balanced, symmetrical) coin about half of the trials will result in heads

(3) Counting Techniques – Permutations and Combinations

These techniques are helpful in computing the probability of an event when the total number of possible events are large.

❖ **Multiplication Principle**

If one operation can be performed in n_1 ways and a second operation can be performed in n_2 ways, then there are $n_1 \cdot n_2$ ways in which both operations can be carried out.

(4) Combinations

A combination is an arrangement of objects without regard to order. The number of combinations of n distinct objects chosen r at a time is

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

Proof: nP_r may be interpreted as the number of ways of choosing r objects from n objects and then permuting the r objects r! ways.

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

Self-Assessment Questions (SAQs) for study session 1

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 1.1 (Testing Learning Outcomes 1.1)

A box contains 4 white balls, two red, and two green balls. Use the classical definition of probability to find the probability of

i. Drawing a red ball on one draw from the box

ii. Drawing a black ball on one draw from the box.

SAQ 1.2 (Testing Learning Outcomes 1.2)

If two fair dice are rolled once, what is the probability that the probability that the total number of spots shown is :

i. Equal to 5?

ii. Divisible by 3?

SAQ 1.3 (Testing Learning Outcomes 1.3)

Twenty balls numbered from 1 to 20 are mixed in an urn and two balls are drawn successively and without replacement. If x_1 and x_2 are the numbers written on the first and second ball drawn, respectively, what is the probability that:

i. $x_1 + x_2 = 8$?

ii. $x_1 + x_2 \leq 5$?

SAQ 1.4 (Testing Learning Outcomes 1.4)

If $P(A/B) > P(A)$, then show that $P(B/A) > P(B)$.

5. Show that:

i. $P(A'/B) = 1 - P(A/B)$

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

SAQ 1.5 (Testing Learning Outcomes 1.5)

Define the following terms with relevant examples:

i. An experiment

ii. A sample space

iii. An event

References

Bain L. J and Engelhardt Max (1989). Introduction to Probability and Mathematical Statistics. Boston: PWS-KENT Publishing Company.

Hogg,R. V and Tanis E. A (1993). Probability and Statistical Inference. New York: Macmillan Publishing Company.

Larson Harold J (1982). Introduction to Probability Theory and Statistical Inference. Third Edition. New York: John Wiley & Sons.

Spiegel, Murray R (1972). Theory and Problems of Statistics. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 2: Functions and Random Variables

Expected duration: 1 week or 2 contact hours

Introduction

If as we observe a characteristic, we find that it takes on different values in different persons, places, or things, we label the characteristic a **VARIABLE** e.g., heights of adults, weights of children and ages of patients seen in a dental clinic.

Whenever we determine the height, weight or age of an individual, the result is frequently referred to as a value of the respective variable. When the values obtained arise as a result of chance factors, the variable is called a random variable.

Learning Outcomes from Study Session 2

At the end of this study session, you should be able to:

2.1 Highlight on Functions and Random Variables
2.2 Discrete Random Variable
2.3 Continuous Random Variable

2.1 Functions and Random Variables

A function is simply a rule by which every member of one set is assigned to or paired with one member of another set.

Let X and Y be sets. If f is a rule that assigns to every element x in the set X a unique element y in the set Y, then f is said to be a function that maps X into Y.

$$X \quad Y$$

$$x_1 \rightarrow y_1$$

$$x_2 \rightarrow y_2$$

$$\left.\begin{array}{l} x_3 \\ x_4 \\ x_5 \end{array}\right\} \rightarrow y_3$$

For example: suppose that X is the set of 5 students in a seminar and that Y is the set of term paper topics. Let a rule be defined as "choose a topic for a term paper". The rule "Choose a topic ….." is called a function. A random variable is therefore a function which assigns numerical values to the different outcomes defined by the sample space.

2.1.1 Definition Random

Given a random experiment with a sample space S. A function X that assigns to each element c ∈ S one and only one real number $X(c) = x$ is called a *Random Variable*. The space of X is the set of real numbers

$$A = \{x : x = X(c), c \in S\}$$

In a group of 10 people, a, b, c, d… j, a person is selected at random from this group (random selection can have 10 possible results). Suppose that the height of each person has been measured to the nearest inch, the weight measured to the nearest 10 pounds

S	X	S	Y
a →	64	a →	130
b →	67	b →	150
c →	63	c →	110
d →	68	d →	160
e →	72	e →	170
f →	62	f →	180
g →	68	g →	210
h →	66	h →	160
i →	68	i →	130
j →	65	j →	150

The arrow that connects a value in X to every individual in S represents the operation "measure the height of each person to the nearest inch". A rule that assigns a numerical value to every result in the sample space of a random experiment is called a random variable, and the numbers that are assigned by this rule are called the value set of the random variable.Draw an individual at random; what is the chances that the height of the individual drawn is 68 inches? Ans. $^{3}/_{10}$

Example 1: Suppose a coin is tossed twice so that the sample space is S. Let X represents the number of heads which can come up.

Elements of the sample space **S**	No. of Heads (Random Variable) **X**	Probability **f(x)**
HH	2	¼
HT	1	¼
TH	1	¼
TT	0	¼

The probability distribution (function) is a function which assigns probabilities to these numerical values. If X stands for the random variable "number of heads" then x is the values that the random variable can assume. The probability that a random variable X takes on the value x is written as:

P(X = x) or f(x) and is called the probability density function, p.d.f.

P(X = 0) = f(0) = ¼

P(X = 1) = f(1) = ¼

P(X = 2) = f(2) = ¼

Random variables can be classical as discrete or continuous.

2.2 Discrete Random Variable

Let X denote a random variable with space R, suppose we can compute P(X $\in$ A); A $\subset$ R. Let X denote a random variable with one-dimensional space R, a subset of the real numbers. Suppose that the space R contains a countable number of points; that is, R contains either a finite number of points or the points of R can be put into a one-to-one correspondence with the positive integers.

Such a set R is called a set of discrete points or simply a discrete sample space. The random variable X is therefore called a random variable of the discrete type and X is said to have a distribution of the discrete type.

For a random variable X of the discrete type, the probability P(X = x) is frequently denoted by f(x) and this function f(x) is called the probability density function. Some authors refer to f(x) as the probability function, the frequency function, or the probability mass function.

If the set of all possible values of a random variable X, is a countable set, $x_1, x_2, \ldots$ then X is called a discrete random variable. The function

$$f(x) = P[X = x],\ x = x_1, x_2, \ldots$$

That assigns the probability to each possible value x will be called the discrete probability density function (or probability mass function, p.m.f).

Property: A function f(x) is a discrete p.d.f if and only if is satisfies the following properties for at most a countably infinite set of real numbers x_1, x_2 ….

i. $f(x_i) \geq 0 \quad \forall x_i$ (Non-negative)

ii. $\sum_{all\ x_i} = 1$

iii. $P(X \in A) = \sum_{x \in A} f(x)$ where A$\subset$R

Example 2 : Let random variable X of pdf $f(x) = x/6$, x = 1, 2, 3

= 0 otherwise

find i. P[X = 1 or 2]

ii. P[X ≥ 2]

Solution: i. P[X = 1 or 2] = P[X = 1] or P[X = 2]

$$\sum_{x=1}^{2} f(x) = \sum_{1}^{2} x/6$$

$$= 1/6 + 2/6$$

$$= 3/6 \qquad \text{if } x = 1,\ f(1) = 1/6$$

$$= 1/2 \qquad x = 2,\ f(2) = 2/6$$

$$P[X \geq 2] = \sum_{x=2}^{3} f(x) = \sum_{2}^{3} x/6$$

$$= 2/6 + 3/6$$

$$= 5/6$$

or

$$P[X \geq 2] = 1 - P[X < 2]$$

$$= 1 - f(1)$$

$$= 1 - 1/6 = 5/6$$

2. For each of the following determine the constant c so that f(x) satisfies the conditions of being a p.d.f for a random variable X.

   a. $f(x) = x/c$, x = 1, 2, 3, 4

   b. f(x) = cx, x = 1, 2, 3, ….., 10

   c. $f(x) = c(x + 1)^2$, x = 0, 1, 2, 3

   d. $f(x) = x/c$, x = 1, 2, 3, …., n

Solution:

i. $f(x) = x/c, \qquad x = 1,2,3,4$

$$\sum f(x) = 1 \Rightarrow \sum x/c = 1$$

$$\frac{1}{c}\sum x = 1$$

$$\frac{1}{c}[1+2+3+4] = 1$$

$$\frac{10}{c} = 1 \qquad \therefore c = 10$$

ii. f(x) = cx, x = 1, 2, 3, ….. 10

$\sum cx = 1$

$c\sum x = 1$

$c[1 + 2 + 3 + ….. + 10] = 1$

$55c = 1$

$c = \frac{1}{55}$

iii.f(x) = $c(x + 1)^2$, x = 0, 1, 2, 3

$\sum c(x + 1)^2 = 1$

$c\sum(x + 1)^2 = 1$

$c[1^2 + 2^2 + 3^2 + 4^2] = 1$

$30c = 1$

$c = \frac{1}{30}$

iv. $f(x) = \frac{x}{c}$, x = 1, 2, 3, …., n

$$\sum \frac{x}{c} = 1$$

$$\frac{1}{c}\sum_{x=1}^{n} x = 1$$

$$\frac{1}{c}\left[\frac{(n+1)}{2}\right] = 1$$

$$\therefore c = \frac{n(v+1)}{2}$$

Note the following:

1. $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

2 $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$

3. $\sum_{i=1}^{n} i^3 = \left[\frac{n(n+1)}{2}\right]^2$

4. $\sum_{i=1}^{n} i^4 = \frac{n(n+1)(2n+1)(3n^2+3n+1)}{30}$

5. $(1+t)^n = \sum_{k=0}^{n} \binom{n}{k} t^k$ i.e $\sum \binom{n}{k} = 1$

6. $\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = (a + b)^n$

7. $(1-t)^n = \sum \binom{n}{k} (-1)^k 1^k$

8. $2^n = \sum_{k=0}^{n} \binom{n}{k}$

9. $\sum_{k=0}^{n} \binom{a}{k} \binom{b}{n-k} = \binom{a+b}{n}$

2.3 Continuous Random Variable

Random variable whose spaces are not composed of a countable number of points, but are intervals or a union of intervals are said to be of the continuous type.The probability density function, pdf, of a random variable X of the continuous type, with space R that is an interval or union of intervals, is an integrable function f(x) satisfying the following conditions.

i. $f(x) > 0, \; x \in R$

ii. $\int_R f(x)dx = 1$

iii. The probability of the event $X \in A$ is

$$P(X \in A) = \int_A f(x)dx$$

Example 3: A machine produced copper wire, and occasionally there is a flaw at some point along the wire. The length of wire (in meters) produced between successive flaws is a continuous random variable X with p.d.f of the form

$$f(x) = \begin{cases} c(1+x)^{-3} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

where c is a constant

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

$$\int_0^{\infty} c(1+x)^{-3}\,dx = 1$$

Let U = (1 + x) and du = dx apply power rule for integral

$$c\int_0^{\infty} U^{-3}du = 1$$

$$c\left[\frac{U^{-2}}{-2}\right]_0^{-\infty} = 1$$

$$c\left[\frac{1}{2}\right] = 1$$

c = 2

Example 4: For each of the following functions, find the constant c so that f(x) is a p.d.f of a random variable X.

i. $f(x) = 4x^c, \quad 0 \le x \le 1$

ii. $f(x) = c\sqrt{x}, \quad 0 \le x \le 4$

iii. $f(x) = c/x^{3/4}, \quad 0 < x < 1$

Solution

i. $f(x) = 4x^c$

$$\int_0^1 f(x)dx = 1$$

$$\int_0^1 4x^c dx = 1$$

$$4\int_0^1 x^c dx = 1$$

$$4\left[\frac{x^{c+1}}{c+1}\right]_0^1 = 1$$

$$4\left[\frac{1}{c+1}\right] = 1$$

c + 1 = 4

c = 3

ii. $f(x) = c\sqrt{x}$

$$c\int_0^4 x^{1/2}dx = 1$$

$$c\left[\frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1}\right]_0^4 = 1$$

$$c\left[\frac{4^{\frac{3}{2}}}{\frac{3}{2}}\right] = 1$$

$$c = \frac{3}{16}$$

iii. $f(x) = \frac{c}{x^{\frac{3}{4}}}$

$$c\int_0^1 x^{-\frac{3}{4}} = 1$$

$$c\left[\frac{x^{-\frac{3}{4}+1}}{-\frac{3}{4}+1}\right]_0^1 = 1$$

$$\frac{c}{\frac{1}{4}} = 1$$

$$c = \frac{1}{4}$$

Summary

In this study session you have learnt about:

(1) Functions and Random Variables

A function is simply a rule by which every member of one set is assigned to or paired with one member of another set.

(2) Discrete Random Variable

Let X denote a random variable with space R, suppose we can compute P(X $\in$ A); A $\subset$ R. Let X denote a random variable with one-dimensional space R, a subset of the real numbers. Suppose that the space R contains a countable number of points; that is, R contains either a finite number of points or the points of R can be put into a one-to-one correspondence with the positive integers.

(3) Continuous Random Variable

Random variable whose spaces are not composed of a countable number of points but are intervals or a union of intervals are said to be of the continuous type.

Self-Assessment Questions (SAQs) for study session 2

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study

Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 2.1 (Testing Learning Outcomes 2.1)

1. Define the following terms;

   i. discrete random variable

   ii. continuous random variable

SAQ 2.2 (Testing Learning Outcomes 2.2)

2. Let $f(x) = \frac{x}{15}$, $x = 1,2,3,4,5$, zero elsewhere be the p.d.f. of X.

   Find;

   i. Pr [1 or 2]

   ii. Pr [$1 \leq X \leq 3$]

   iii. Pr [$\frac{1}{2} \leq X \leq \frac{5}{2}$]

SAQ 2.3 (Testing Learning Outcomes 2.3)

3. For each of the following functions, find the constant c so that $f(x)$, satisfies the conditions of being a p.d.f. of a random variable X.

   i. $f(x) = cx^3$, $o < x < 1$

   ii. $f(x) = \frac{c}{x^4}$, $0 < x < \infty$

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference.* Third Edition. New York: John Wiley & Sons.

Ross S (1994)**.***A First Course in Probability*. 4th Ed., New Jersey: Prentice-Hall, Inc. A Simon & Schuster Company, Englewood Cliffs.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods Nigeria*: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 3: Distribution Functions of Random Variable

Expected duration: 1 week or 2 contact hours

Introduction

Frequently, we are interested in the probability that a random variable is equal to or less than some specified value or greater than a given value. The cumulative distribution function is particularly useful in this connection. The probability function for a discrete random variable X gives the probability of occurrence of the elements in the range of X.

It can then be used to compute the probability of occurrence of any event defined by the observed value of X. In this study session, we shall consider the distribution function (also frequently called the cumulative distribution function or cdf) for a random variable X. It is simply an alternative function that can be used to evaluate probabilities of events defined by the observed value of random variables.

Learning outcomes from Study Session 3

At the end of this study session, you should be able to:

3.1 Explain Distribution of Random Variables

3.1Distribution Functions of Random Variable

Given a random variable X, the value of the cumulative distribution function at x, denoted by $F(x)$, is the probability that X takes on values less than or equal to x. Hence,

$$F(x) = P(X \leq x)$$

In the case of a discrete random variable, it is clear that

$$F(c) = \sum_{x \leq c} f(x)$$

The symbol, $\sum_{x \leq c} f(x)$ means "sum the values of $f(x)$ for all values of x less than or equal to c".

The relationship between $F(x)$ and $f(x)$ for a discrete distribution is given by the following theorem;

1. Theorem: Let X be a discrete random variable with p.d.f. $f(x)$ and cdf $F(x)$. If the possible values of X are indexed in increasing order, $x_1 < x_2 < x_3 < ...$, then $f(x_1) = F(x_1)$, for any $i > 1$,

$$f(x_i) = F(x_i) - F(x_{i-1})$$

Additionally, if $x < x_1$ then $F(x) = 0$, and for any other real x

$$F(x) = \sum_{x_i \le x} (fx_i)$$

where the summation is taken over all indices i such that $x_i < x$.

The distribution function of a random variable X of the continuous type defined in terms of the p.d.f. of X is given by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \text{ Note that } F'(x) = f(x)$$

The following are the properties of a distribution function $F(x)$ as a consequence of the fact that probability must be a value between 0 and 1, inclusive.

i. $0 \le F(x) \le 1$ since $F(x)$ is a probability.

ii. $F(x)$ is a non decreasing function of x.

iii. $F(w) = 1$, where w is any value greater than or equal to the largest value in R; and $F(z) = 0$, where z is any value less than the smallest value in R.

iv. If X is a random variable of the discrete type, then $F(x)$ is a step function, and the height of a step at $x, x \in R$, equal the probability $P(X = x)$.

Example 1: Let the random variable X of the discrete type have the p.d.f $f(x) = x/7$, $x = 1,2,4$. Find the distribution function of X..

Note that $P(X \le 0) = f(0) = 0$

But $P(X \le 1) = f(1) = \frac{1}{7}$,

$$P(X \le 2) = f(1) + f(2) = \frac{1}{7} + \frac{2}{7} = \frac{3}{7}$$

$$P(X \le 4) = f(1) + f(2) + f(4) = \frac{1}{7} + \frac{2}{7} + \frac{4}{7} = 1$$

So let $F(x) = P(X \le x)$ be defined for each real number x. Then

$$F(x) = \begin{cases} 0, -\infty < x < 1 \\ \frac{1}{7}, 1 \le x < 2 \\ \frac{3}{7}, 2 \le x < 4 \\ 1, 4 \le x < \infty \end{cases}$$

Note that $F(x)$ cumulates all the probability from points that are less than or equal to x.

Example 2: Let the random variable X be the distance in feet between bad records on a used computer tape. Suppose that a reasonable probability model of X is given by the p.d.f.

$$f(x) = \begin{cases} 0 & -\infty < x < 0 \\ \dfrac{1}{40} e^{-x/40} & , 0 \le x < \infty \end{cases}$$

the distribution function of X is

F(x) = 0 for x $\le$ 0 and for x > 0

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \frac{1}{40} e^{-t/40} dt$$

$$= \quad -e^{-t/40}\Big|_0^x = 1 - e^{-x/40}$$

Note that,

$$F'(x) = \begin{cases} 0 & -\infty < x < 0 \\ \dfrac{1}{40} e^{-x/40} & , 0 \le x < \infty \end{cases}$$

Example 3: Let the random variable X have the p.d.f

$f(x) = 2(1-x), \quad 0 \le x \le 1$, zero elsewhere.

Determine the distribution function of X.

Solution:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} 2(1-t)dt$$

$$= \quad 2t\Big|_0^x - t^2\Big|_0^x$$

$$= \quad 2x - x^2 = x(2-x)$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x(2-x) & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

Example 4: For each of the following functions.

i. Find the constant c so that *f*(*x*) is a p.d.f of a random variable X.

ii. Find the distribution function, F(*x*) = P(X ≤*x*)

a. $f(x) = 4x^c, \quad 0 \le x \le 1 \qquad c = 3$

b. $f(x) = c\sqrt{x}, \quad 0 \le x \le 4 \qquad c = 3/16$

c. $f(x) = c/x^{3/4}, \quad 0 < x < 1 \qquad c = 1/4$

Solution: $f(x) = 4x^3$

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} 4t^3 dt = t^4 \Big|_0^x = x^4$$

$$F(x) = \begin{cases} 0 & x < 0 \\ x^4 & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

Summary

In this study session you have learnt about:

1. Distribution Functions of Random Variable

Given a random variable X, the value of the cumulative distribution function at x, denoted by $F(x)$, is the probability that X takes on values less than or equal to x. Hence,

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

2. Distribution Function of Random Variable

The distribution function of a random variable X of the continuous type defined in terms of the p.d.f. of X is given by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \text{ Note that } F'(x) = f(x)$$

The description of cumulative distribution function.The properties of distribution functions. The relationship between $F(x)$ and $f(x)$ for a discrete distribution.

How to determine the distribution function of a random variable X, of the discrete and continuous type.

Self-Assessment Questions (SAQs) for study session 3

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study

Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 3.1 (Testing Learning Outcomes 3.1)

1. Let $f(x)$ be the p.d.f. of a random variable X. Find the distribution function $F(x)$ of X.

    i. $f(x)=1,\ x=3$

    ii. $f(x)=\frac{1}{3},\ x=1,2,3.$

    iii. $f(x)=\frac{x}{15},\ x=1,2,3,4,5.$

    iv. $f(x)=\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^{x},\ x=0,1,2,....$

SAQ 3.1 (Testing Learning Outcomes 3.1)

For each of the followings, find the distribution function F(x) = P(X $\leq x$).

i. $f(x)=\left(\frac{3}{16}\right)x^{\frac{1}{2}}$

ii. $f(x)=\frac{\left(\frac{1}{4}\right)}{x^{\frac{3}{4}}}$

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hamburg Morris (1970).*Statistics Analysis for Decision Making*. New York: Harcourt, Brace & World, Inc.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods Nigeria.* Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 4: Mathematical Expectation

Expected duration: 1 week or 2 contact hours

Introduction

The probability distribution for a random variable can de defined by either its distribution function, $F(x)$, or its density function, $f(x)$, for continuous or discrete type. Once the probability distribution of X is known, the probabilities of occurrence for any event of interest can thus be computed.

However, in many applications, we may be interested in describing various aspects of different probability distributions, ways of describing certain properties of probability distributions. For instance, what is a “typical” value the random variable can assume? “Typical” here may be defined in various ways.

How much variability is exhibited by the probability distribution or how spread out are the possible observed values for a random variable can be our concern. In this lecture, we will discuss some common measures of certain aspects of probability distributions such as a value that describes the “middle” or the “spread” of the probability distribution.

Learning Outcome from Study Session 4

At the end of this study session, you should be able to:

4.1 Discuss Mathematical Expetation

4.1 Mathematical Expectation

An extremely important concept in summarizing important characteristics of distributions of probability is that of mathematical expectation, which is introduced by an example.

Example: A young man who needs a little extra money devises a game of chance in which some of his friends might wish to participate. The game that he proposes is to let the participant cast an unbiased dice and then receive a payment according to the following schedule.

If the event A = {1, 2, 3} occurs, he receives ₦1; if B = {4, 5} occurs, he receives ₦5; and if c = {6} occurs, he receives ₦35.

The probabilities of the respective events are assumed to be $^3/_6$, $^2/_6$, $^1/_6$.

The problem that now faces the young man is the determination of the amount that should be charged for the opportunity of playing the game. He reasons, that if the game is played a large number of times about $^3/_6$ of the trials will require a payment of ₦1; about $^2/_6$ of them will require one of ₦5 and about $^1/_6$ of them will require one of ₦35. Thus the approximate average payment is

$$(1)\left(\frac{3}{6}\right)+(5)\left(\frac{2}{6}\right)+(35)\left(\frac{1}{6}\right)=8$$

He expects to pay ₦8 "on the average". He never pays ₦8, the payment is either ₦1, ₦5 or ₦35. The weighted average of 1, 5 and 35 in which the weights are the respective probabilities $^3/_6$, $^2/_6$, and $^1/_6$, equals eight.

Such a weighted average is called the Mathematical Expectation of payment. If he charges ₦10 per play, he would make on the average ₦2 per play. The most that a player would lose at the charge of ₦10 per play is ₦9; the most he would gain is ₦25.

Definition: If f(x) is the pdf of the random variable X of the discrete type with space R and if the summation.

$\sum_R U(x)f(x)=\sum_{x\in R}U(x)f(x)$ exists, then the sum is called the Mathematical expectation or the expected value of the function is U(X) and it is denoted by E[U(X)]. That is

$$E[U(X)]=\sum_R U(x)f(x)$$

The expected value E[U(X)] is thought of as a weighted mean of U(x), $x\in R$ where the weights are the probabilities f(x) = P(X = x), $x\in R$.

Example 1: Let the random variable X have the pdf

$$f(x)=1/3, \quad x\in R$$

Where R = {- 1, 0, 1}. Let U(X) = X2. then,

$$E[U(X)]=E(X^2)=\sum U(x)f(x)=\sum x^2 f(x)$$

$$=(-1)^2\left(\frac{1}{3}\right)+(0)^2\left(\frac{1}{3}\right)+(1)^2\left(\frac{1}{3}\right)$$

$$=\frac{2}{3}$$

For continuous type random variable, the definitions associated with mathematical expectation are the same as those in the discrete case except that integrals replace summation.If X is a continuous random variable with p.d.f f(x), then the expected value of X is defined by

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

Properties: When it exists, mathematical expectation E satisfies the following properties.

a. If c is a constant,E(c) = c

b. If c is a constant, and U is a function E[cU(X)] = cE[U(X)]

c. If c_1 and c_2 are constants and U_1 and U_2 are functions, then

$$E[c_1U_1(X) + c_2U_2(X)] = c_1E[U_1(X)] + c_2E[U_2(X)]$$

Proof:

a. $E(c) = \sum_R cf(x) = c\sum_R f(x) = c.$ Since $\sum_R f(x) = 1$

In the continuous case,

$E(c) = \int_{-\infty}^{\infty} cf(x)dx = c\int_{-\infty}^{\infty} f(x)dx = c.$ Since $\int_{-\infty}^{\infty} f(x)dx$

b. $E[cU(X)] = \sum_R cU(x)f(x)$

$= c\sum_R U(x)f(x)$

$= cE[U(X)]$ Since $E[U(X)] = \sum U(x)f(x)$

For the continuous case,

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

$$= c\int_{-\infty}^{\infty} U(x)f(x)dx$$

$= cE[U(X)]$ Since $E[U(X)] = \int_{-\infty}^{\infty} U(x)f(x)dx$

c. $E[c_1U_1(X) + c_2U_2(X)] = \sum_R [c_1U_1(X) + c_2U_2(X)]f(x)$

$= \sum_R c_1U_1(X)f(x) + \sum_R c_2U_2(X)f(x)$

$= c_1\sum_R U_1(X)f(x) + c_2\sum_R U_2(X)f(x)$

$= c_1E[U_1(X)] + c_2E[U_2(X)]$

The Mean, Variance and Standard Deviation

If X is a random variable with p.d.f. f(x) of the discrete type and space $R = \{b_1, b_2, b_3, ...\}$ then

$$E(X) = \sum_R xf(x)$$

$= b_1f(b_1) + b_2f(b_2) + b_3f(b_3) + \ldots\ldots$ is the weighted average of the numbers belonging to R, where the weights are given by the p.d.f. f(x). If X is a continuous random variable having p.d.f f(x), then $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

E(X) is called the mathematical expectation, or mean value or just mean of X (or the mean of the distribution) and denoted by μ. That is, μ = E(X).

Example 2: Let X have the p.d.f

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

The mean of X is

$$\mu = E(X) = 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right) = \frac{3}{2}$$

Note that the mean μ = E(X) is the centroid of a system of weights or a measure of the central location of the probability distribution of X. A measure of the dispersion or spread of a distribution is defined as follows.

If $U(x) = (x - \mu)^2$ and $E[(X - \mu)^2]$ exists, the variance denoted by σ^2 or V(X) of a random variable X of the discrete type is defined by

$$\sigma^2 = E[(X-\mu)^2] = \sum_R (x-\mu)^2 f(x)$$

The positive square-root of the variance is called the standard deviation of X and denoted by

$$\sigma = \sqrt{Var(x)}$$

Example 2

A. Let the pdf of X be given by f(0) = $^3/_{10}$, f(1) = $^3/_{10}$, f(2) = $^1/_{10}$ and f(3) = $^3/_{10}$, compute the mean, variance and standard deviation of X.

B . Find the mean and variance for the following discrete distributions.

i. $f(x) = \frac{1}{5}, \quad x = 5,10,15,20,25$

ii. $f(x) = 1, \quad x = 5$

iii. $f(x) = \frac{4-x}{6}, \quad x = 1, 2, 3$

Solution: 2a. $E(x) = \sum xf(x) = 0\left(\frac{3}{10}\right) + 1\left(\frac{3}{10}\right) + 2\left(\frac{1}{10}\right) + 3\left(\frac{3}{10}\right)$

$$= \frac{3}{10} + \frac{2}{10} + \frac{9}{10}$$

$$= \frac{14}{10} = 1.4$$

$$E\left[(X-\mu)^2\right] = V(X) = \sum (x-\mu)^2 f(x)$$

$$= (0-1.4)^2\left(\frac{3}{10}\right) + (1-1.4)^2\left(\frac{3}{10}\right) + (2-1.4)^2\left(\frac{1}{10}\right) + (3-1.4)^2\left(\frac{3}{10}\right)$$

$$= \frac{5.88}{10} + \frac{0.48}{10} + \frac{0.36}{10} + \frac{7.68}{10}$$

$$= \frac{14.4}{10} = 1.44$$

$$\sigma = \sqrt{V(X)} = 1.2$$

or $V(X) = E(X^2) - E^2(X)$

$$E(X^2) = \sum x^2 f(x)$$

$$= 0^2\left(\frac{3}{10}\right) + 1^2\left(\frac{3}{10}\right) + 2^2\left(\frac{1}{10}\right) + 3^2\left(\frac{3}{10}\right)$$

$$= \frac{3}{10} + \frac{4}{10} + \frac{27}{10}$$

$$= \frac{34}{10}$$

$$V(X) = \frac{34}{10} - \left(\frac{14}{10}\right)^2$$

$$= \frac{34}{10} - \frac{196}{100}$$

$= 3.4 - 1.96 = 1.44$

b i. $E(X) = \sum xf(x) = \sum x\left(\frac{1}{5}\right)$

$$= \frac{5}{5} + \frac{10}{5} + \frac{15}{5} + \frac{20}{5} + \frac{25}{5}$$

$= 1 + 2 + 3 + 4 + 5$

$= 15$

Example 3 Given $E(X+4) = 10$ and $E\left[(X+4)^2\right] = 116$. Determine

i. Var (X + 4)

ii. μ

iii. σ^2

Solution

i. $Var(X + 4) = E[(X + 4)^2] - E^2(X + 4)$

$= 116 - 10^2$

$= 116 - 100$

$= 16$

ii. $\mu = E(X + 4) = 10$

$E(X) + 4 = 10$

$E(X) = \mu = 6$

iii. $\sigma^2 = Var(X) = E(X^2) - E^2(X)$

But $E[(X + 4)^2] = E[X^2 + 8X + 10] = 116$

$= E(X^2) + 8E(X) + 16 = 116$

$= E(X^2) + 48 = 100$

$= E(X^2) = 52$

$Var(X) = E(X^2) - E^2(X)$

$= 52 - 6^2$

$= 52 - 36$

$= \quad 16$

As a measure of variability or spread in a continuous distribution, we will again consider the variance, Var (X) = E[(X - μ)²], and the standard deviation is denoted by $\sigma = \sigma_X = \sqrt{Var(X)}$

The relationship Var(X) = $E(X^2) - \mu^2$ holds, where $E\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f(x)dx$

Example 4: Let Y be a continuous random variable with p.d.f.

$$f(y) = 2y, \ 0 < y < 1$$

Obtain E(Y) and Var(Y)

Solution:

$$\mu = E(Y) = \int_0^1 yf(y)dy$$

$$= \int_0^1 y(2y)dy$$

$$= \int_0^1 2y^2 dy$$

$$= \left[\frac{2y^3}{3}\right]_0^1 = \frac{2}{3}$$

$$\sigma^2 = Var(Y) = E(Y^2) - \mu^2$$

$$E(Y^2) = \int_0^1 y^2 f(y)dy$$

$$= \int_0^1 2y^3 dy$$

$$= \left[\frac{2y^4}{4}\right]_0^1$$

$$= \frac{2}{4}$$

$$= \frac{2}{4} - \frac{4}{9} = \frac{1}{2} - \frac{4}{9}$$

$$= \frac{1}{18}$$

Summary

In this study session you have learnt about:

1. Mathematical Expectation

An extremely important concept in summarizing important characteristics of distributions of probability is that of mathematical expectation.

- The expected value of a random variable X of the discrete type, if it exists, is given by $E(x) = \sum_x xf(x)$.
- The expected value of a random variable X of the discrete type, if it exists, is given by $E(x) = \sum_x xf(x)$.

2. The Mean, Variance and Standard Deviation

If X is a random variable with p.d.f. f(x) of the discrete type and space $R = \{b_1, b_2, b_3, ...\}$ then

$$E(X) = \sum_R xf(x)$$

= $b_1f(b_1)$ + $b_2f(b_2)$ + $b_3f(b_3)$ + …… is the weighted average of the numbers belonging to R, where the weights are given by the p.d.f. f(x). If X is a continuous random variable having p.d.f f(x), then $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

E(X) is called the mathematical expectation, or mean value or just mean of X (or the mean of the distribution) and denoted by μ. That is, μ = E(X).

Self-Assessment Questions (SAQs) for study session 4

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 4.1 (Testing Learning Outcomes 4.1)

1. Solve question 1b (ii) and (iii).

SAQ 4.2 (Testing Learning Outcomes 4.2)

Show that $E[cU(X)+d]=cE[U(X)]+d$, and, in particular, $E(cX+d)=cE(X)+d$ where c and d are constants.

SAQ 4.3 (Testing Learning Outcomes 4.3)

Suppose $f(x)=\frac{1}{5}$, $x=1,2,3,4,5$, zero elsewhere, is the p.d.f. of the discrete type of random variable X. Compute $E(X)$ and $E(X^2)$. Hence or otherwise find $Var(X)$ and $E[(X+2)^2]$.

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1995). *Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory*. Third Edition. New York: Macmillan Publishing Co., I

Study Session 5: Probability Distributions of Discrete Random Variables

Expected duration: 1 week or 2 contact hours

Introduction

In many situations, it is useful to present the probability distribution of a random variable by a general algebraic expression. Probability calculations can then be made conveniently by substitutingappropriatevalues into the algebraic model.

The mathematical expression is a compact form of summarizing the nature of the process that has generated the probability distribution. In this lecture, three probability distributions of the discrete type will be discussed: the Bernoulli, Binomial and Poisson distributions.

Learning Outcome from Study Session 5

At the end of this study session, you should be able to:

5.1 Explain Probability Distribution
5.2 Discuss Poisson Distribution

5.1 Probability Distribution

A probability distribution is a mathematical idealization, or model of the relative frequency distribution of outcomes of a random experiment. If a random variable can assign only a countable number of values to the result of a random experiment, it is said to be a discrete random variable. However, a random variable that can assume any real value is called a continuous random variable.

5.1.1 Bernoulli Distribution

The Trial: An experiment with exactly two possible results is called a trial e.g. guilty and not guilty, success and failure, defective and non defective, male or female, etc. The distribution is derived from a process known as a Bernoulli trial, when a single trail of an experimental result in only one of two mutually exclusive outcomes e.g. dead or alive, sick or well, the trial is called a Bernoulli trial. A random variable, X, that assumes only the values 0 or 1 is known as a Bernoulli variable, and a performance of an experiment resulting in only two types of outcomes is called a Bernoulli trial.

The p.d.f. of a Bernoulli distribution is given as

$f(x) = p^x q^{1-x}, \; x = 0,1$

Where p is the probability of success and it remains constant from trial to trial, the corresponding probability of failure is denoted by q which is equal to 1-p. The trials are independent. That is, the outcome of any given trial or sequence of trials does not affect the outcomes on subsequent trials. The outcome of any specific trial is determined by chance. Such processes are referred to as “random process” or “stochastic process.” Bernoulli trials are one example of such processes.

Properties

Mean $\mu = p$

Variance $\sigma^2 = pq$

Standard Deviation $\sigma = \sqrt{pq}$

An important distribution arising from counting the number of successes in a fixed number of independent Bernoulli trials is the Binomial distribution.

Example 1: An urn contains 5 red and 15 green balls. Draw one ball at random from the urn. Let X=1 if the ball drawn is red, and X=0 if a green ball is drawn. Obtain:

i. The p.d.f. of X,
ii. Mean of X and
iii. Variance of X.

Solution:

The p.d.f. of a Bernoulli distribution is $f(x) = p^x q^{1-x}, \; x = 0,1$

where $p = 5/20$ and $q = 15/20$

$$f(x) = \left(\frac{5}{20}\right)^x \left(\frac{15}{20}\right)^{1-x}, x = 0,1$$

Mean of $X = E(X) = \sum_{x=0}^{1} x\left(\frac{5}{20}\right)^x \left(\frac{15}{20}\right)^{1-x} = (0)\left(\frac{5}{20}\right)^0 \left(\frac{15}{20}\right)^1 + (1)\left(\frac{5}{20}\right)^1 \left(\frac{15}{20}\right)^0$

$$= \left(\frac{5}{20}\right)$$

Variance of X$= V(X) = \sum_{0}^{1} x^2 f(x) - [E(X)]^2 = (1)\left(\frac{5}{20}\right)^1 \left(\frac{15}{20}\right)^0 - \left(\frac{5}{20}\right)^2$

$$= \left(\frac{5}{20}\right) - \left(\frac{5}{20}\right)^2 = \left(\frac{3}{16}\right)$$

5.1.2 Binomial Distribution

A binomial experiment is an experiment that can be regarded as a sequence of n trials and meeting the following conditions.

i. The underlying experiment consists of n repeated trials (n is defined before the experiment begins).
ii. The result of every trial can be classified into one of two mutually exclusive categories.
iii. The probability of success p does not change from trial to trial.
iv. The result of any trial is independent of the results of all other trials.

The shape of the distribution depends on the two parameters p and n:

i. When $p < 0.5$ and n is small, the distribution will be skewed to the right.
ii. When $p > 0.5$ and n is small, the distribution will be skewed to the left
iii. When $p = 0.5$ the distribution will be symmetric.
iv. In all cases, as n gets larger the distribution gets closer to being a symmetric, bell-shaped distribution.

Properties

Mean $\mu = np$

Variance $\sigma^2 = npq$

Standard Deviation $\sigma = \sqrt{npq}$

If X is a random variable with probability of a success p, then the probability of obtaining x success in n trials is

$f(x) = \binom{n}{x} p^x q^{n-x}$ $x = 0, 1, ---- n$ i.e the probability of x number of successes in n number of trials. This is the p.d.f of a Binomial distribution.

Example: 2 If 20% of the bolts produced by a machine are bad. Determine the probability that out of 4 bolts chosen at random.

i. one is defective
ii. none is defective
iii. at most 2 bolts will be defective.

Solution: $n = 4$, $p = 0.2$, $q = 0.8$

$$P[X=1]=f(1)=\binom{4}{1}0.2^{1}0.8^{3}=0.4096$$

$$P[X=0]=f(0)=\binom{4}{0}0.2^{0}0.8^{4}=0.4096$$

$P[X \le 2] = P[X = 0, 1, 2] = P[X = 0] + P[X = 1] + P[X = 2]$

$0.4096 + 0.4096 + 0.1536 = 0.9728$

or $\quad = 1 - P[X > 2] = 1 - P(X = 3) - P(X = 4)$

or $\quad 1 - P[X \ge 3] = 1-\binom{4}{3}0.2^{3}0.8^{1}-\binom{4}{4}0.2^{4}0.8^{0}$

$= 1 - 0.0256 - 0.0016$

$= 0.9728$

Example 3:

a. Suppose that it is known that 30% of a certain population is immune to some disease. If a random sample of 10 is selected from this population. What is the probability that it will contain exactly 4 immune persons?

$n = 10, \quad p = 0.3, \quad x = 4$

$$f(4)=\binom{10}{4}(0.3)^{4}(0.7)^{6}$$

$= 0.2$

b. In a certain population 10% of the population is color-blind. If a random sample of 25 people is drawn from this population (use table). Find the probability that

i. $P(X \ge 5) = 1 - P(X < 5) = 0.0980$

ii. $P(X \le 4) = 0.902$ or $1 - P(X \ge 5) = 1 - 0.0980 = 0.902$

iii. $P(6 \le X \le 10) = p(6) + p(7) + p(8)$ ----- $+ p(10) = 0.0333$

$= 0.0334$ i.e $P(X \ge 6)$

Example 4: From the experiment "toss four coins and count the number of tails" what is the variance of X?

$n = 4, \quad p = ½, \quad q = ½$

$V(X) = npq = 4 \times ½ \times ½ = 1$

Example 5: Roll a fair 6 – sided dice 20 times and count the number of times that 6 shows up what is the standard development of your random variable?

$$n = 20, \quad p = \frac{1}{6} \quad q = \frac{5}{6}$$

$$V(X) = npq$$

$$= 20 \times \frac{1}{6} \times \frac{5}{6} = \frac{100}{36}$$

$$\sigma = \sqrt{V(X)} = \sqrt{\frac{100}{36}} = \frac{10}{6}$$

5.2 The Poisson Distribution

This is concerned with occurrences that can be described by a discrete random variable. The random variable can take on values, x = 0, 1, 2 ---- (i.e. non-negative integers)

Countably infinite distribution, e.g

- ✓ A number of telephone calls per minute at a switchboard;
- ✓ A number of mistakes per page in a large document; and
- ✓ A numbers of traffic arrivals such as trucks at terminals, air planes at airports, ships at docks etc.

All these have something in common, the given occurrences can be described in terms of a discrete random variable which takes values 0, 1, 2, ----.

The Poisson distribution can be used to find the probability that a certain number of events will occur in a given period of time provided that the following criteria are satisfied:

1. The time interval used can be divided into many sub-intervals so small that the probability of the event occurring in any one sub-interval is almost zero.
2. The probability of more than one occurrence in any subinterval is negligible.
3. The occurrences of the events are independent. The occurrence of an event in an interval of space or time has no effect on the probability of a second occurrence of the event in the same or any other interval.
4. The probability of the single occurrence of the event in a given interval is proportional to the length of the interval.
5. The probability of an occurrence in any of the subintervals (or the mean rate of occurrence) remains constant throughout the entire time under consideration.

When the above mentioned criteria are satisfied, the probability of X occurrences per unit of time is given by

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, ----\infty, \quad \lambda > 0$$

λ is the average number of occurrences of the random event in the interval and e is the constant 2.7183.

The Poisson distribution gives a very good approximation to the binomial distribution. The Poisson distribution can be used when the sample size is very large, and the probability of an event occurring is very small.

Properties

Mean $E(X) = \mu = \lambda$

Variance $V(X) = \sigma^2 = \lambda$

Standard deviation $\sigma = \sqrt{\lambda}$

Example 6: Suppose that an urn contains 100,000 marbles and 120 are red. If a random sample of 1000 is drawn what are the probabilities that 0, 1, 2, 3, and 4 respectively will be red.

$$n = 1000, \qquad p = \frac{120}{100000} = 0.0012, \; q = 0.9988$$

Solution: Binomial $= \binom{1000}{x} 0.0012^2 . 0.9988^{1000-x}$

For x=3, $\binom{1000}{3} 0.0012^3 \text{ x } 0.9988^{997} = 0.0867$

i.e $166167000 \text{ x } 1.728^{-09} \text{ x } 0.30206$

Using the Poisson method,

$\lambda = np = 1000 \text{ x } 0.0012 = 1.2$

$e^{-1.2} = 0.3012$

$$f(3) = \frac{e^{-1.2} 1.2^3}{3!} = 0.0867$$

$P(X > 5) = 1 - p(X \le 4)$

$= 1 - 0.9985 = 0.0015$

Example 7: Let X have a Poisson distribution with a mean of $\lambda = 5$. Find

i. $P(X \le 6)$
ii. $P(X > 5)$
iii. $P(X = 6)$
iv. $P(X \ge 4)$

Solution:

i. $P(X \le 6) = \sum_{x=0}^{6} \frac{5^x e^{-5}}{x!} = 0.762$

ii. $P(X > 5) = 1 - P(X \le 5) = 1 - 0.616 = 0.384$

iii. $P(X = 6) = P(X \le 6) - P(X \le 5) = 0.762 - 0.616 = 0.146$

iv. $P(X \ge 4) = 1 - P(X < 4)$

Example 8: A hospital administrator, who has been studying daily emergency admissions over a period of several years, has come to the conclusion that they are distributed according to the Poisson law. Hospital records reveal that emergency admissions have averaged three per day during this period. If the administrator is correct in assuming a Poisson distribution. Find the probability that

1. exactly two emergency admissions will occur on a given day;
2. no emergency admissions will occur on a particular day; and
3. either 3 or 4 emergency cases will be admitted on a particular day.

Solution:

1. $p(X = 2) = \frac{e^{-\lambda} \lambda^x}{x!} \qquad \lambda = 3$

$$= \frac{e^{-3} 3^2}{2!} = \frac{0.05(9)}{2} = 0.225$$

2. $p(X = 0) = \frac{e^{-3} 3^0}{0!} = 0.05$

3. $p(X = 3) + p(X = 4) = \frac{e^{-3} 3^3}{3!} + \frac{e^{-3} 3^4}{4!}$

$$= e^{-3}\left[\frac{27}{6} + \frac{81}{24}\right]$$

$$= \quad 0.05\left[\frac{9}{2}+\frac{27}{8}\right]$$

$$= \quad 0.05\ (7.875)$$

$$= \quad 0.394$$

Summary

In this study session you have learnt about:

1. **Probability Distribution**

   A probability distribution is a mathematical idealization, or model of the relative frequency distribution of outcomes of a random experiment. If a random variable can assign only a countable number of values to the result of a random experiment, it is said to be a discrete random variable. However, a random variable that can assume any real value is called a continuous random variable.

2. **The Poisson Distribution**

This is concerned with occurrences that can be described by a discrete random variable. The random variable can take on values, x = 0, 1, 2 ---- (i.e. non-negative integers)

Countably infinite distribution, e.g

- ✓ A number of telephone calls per minute at a switchboard;
- ✓ A number of mistakes per page in a large document; and
- ✓ A numbers of traffic arrivals such as trucks at terminals, airplanes at airports, ships at docks etc.

Self-Assessment Questions (SAQs) for study session 5

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 5.1 (Testing Learning Outcomes 5.1)

Suppose that 24% of a certain population have blood group B, for a sample of size 20 drawn from this population, find the probability that

a. Exactly 3 persons with blood group B will be found.

b. Three or more persons ≡ the characteristics of interest will be found

c. Fewer than three will be found.

d. Exactly five will be found.

SAQ 5.2 (Testing Learning Outcomes 5.2)

In a large population, 16% of the members are left-handed. In a random sample of size 10, find

a. The probability that exactly 2 will be left-handed $p(X = 2)$

b. $P(X \geq 2)$

c. $P(X < 2)$

d. $P(1 \leq X \leq 4)$

SAQ 5.3 (Testing Learning Outcomes 5.3)

Suppose mortality rate of a certain disease is 0.1, suppose 10 people in a community contract the disease, what is the probability that

a. None will survive

b. 50% will be

c. At least 3 will die

d. Exactly 3 will die

SAQ 5.4 (Testing Learning Outcomes 5.4)

Suppose it is known that the probability of recovery from a certain disease is 0.4. If 15 people are stricken with the disease what is the probability that

a. 3 or more will recover?

b. 4 or more will recover?

c. at least 5 will recover?

d. fewer than three recover?

SAQ 5.5 (Testing Learning Outcomes 5.5)

In the study of a certain aquatic organism, a large number of samples were taken from a pond, and the number of organisms in each sample was counted. The average number of organisms per sample was found to be two. Assuming the number of organisms to be Poisson distributed. Find the probability that:

i. The next sample taken will contain one or more organisms,

ii. The next sample taken will contain exactly three organisms,

iii. The next sample taken will contain fewer than five organisms.

SAQ 5.6 (Testing Learning Outcomes 5.6)

It has been observed that the number of particles emitted by a radioactive substance, which reach a given portion of space during time t, follows closely the Poisson distribution with parameter λ =100. Calculate the probability that:

i. No particles reach the portion of space under consideration during time t;

ii. Exactly 120 particles do so;

iii. At least 50 particles do so.

SAQ 5.7 (Testing Learning Outcomes 5.7)

The phone calls arriving at a given telephone exchange within one minute follow the Poisson distribution with a parameter value equal to ten. What is the probability that in a given minute:

i. No calls arrive?

ii. Exactly 10 calls arrive?

iii. At least 10 calls arrive?

References

Hogg R. V and Craig A. T (1995)**.***Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993)**.***Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982)**.***Introduction to Probability Theory and Statistical Inference.* Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). A First Course in Mathematical Statistics Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 6: Probability Distribution of Continuous Random Variables

Expected duration: 1 week or 2 contact hours

Introduction

In the last study session, we discussed the distributions of the discrete random variable. Other examples of the discrete distribution are; Multinomial distributed, Hyper geometric Distribution, Negative Binomial etc.

We will now consider the notion of a continuous random variable. In study session three, we defined a continuous random variable stating clearly its properties. We also demonstrated how the mean and variance of a continuous random variable can be obtained. In this study session, you shall look at some distributions of the continuous type.

Learning Outcome from Study Session 6

At the end of this study session , you should be able to:

1. Describe the theory of continuous distributions;
2. Explain the properties of continuous distributions; and
3. Evaluate the means and variances of any given function.

6.1 Uniform Distribution

Suppose that a continuous random variable X can assume values in a bounded interval only, say the open interval (a, b), and suppose the p.d.f. of X is given as

$$f(x;a,b) = f(x) = \frac{1}{b-a}, a < x < b$$

$$= 0, \text{ elsewhere.}$$

This distribution is referred to as the Uniform or Rectangular Distribution on the interval (a, b) and is simply written as $X \sim U(a,b)$, where 'a' and 'b' are the parameters of the distribution. It provides a probability model for selecting a point at random from the interval (a, b).

Properties

Mean $\mu = \frac{a+b}{2}$

Variance $\sigma^2 = \frac{(b-a)^2}{12}$

Standard Deviation $\sigma = \sqrt{\frac{(b-a)^2}{12}}$

Example 1: The hardness of a certain alloy (measured on Rockwell scale) is a random variable X. Assume that $X \sim U[50,75]$.

a. Find $P[60 < X < 70]$

b. Find E(X)

c. Find Var (X)

Solution:

i. $P[60 < X < 70] = \int_{60}^{70} \frac{1}{b-a} dx$

$$= \frac{1}{75-50}[x]_{60}^{70}$$

$$= \frac{2}{5}$$

ii. E(X) = $\frac{1}{b-a}\int_{50}^{75} x dx = \frac{125}{2}$

Or

E(X) = $\frac{b+a}{2} = \frac{75+50}{2} = \frac{125}{2}$

iii. $Var(X) = E(X^2) - E^2(X) = \frac{1}{25}\int_{50}^{75} x^2 dx - \left(\frac{125}{2}\right)^2 = \frac{625}{12}$

Or

$$Var(X) = \frac{(b-a)^2}{12} = \frac{625}{12}$$

6.2 Exponential Distribution

An exponential distribution is a continuous distribution related to the Poisson distribution. In the Poisson process, the number of changes occurring in a given interval is counted which results in discrete distribution.

But not only is the number of changes a random variable; the waiting times between successive changes are also random variables which are of the continuous type. The latter results into a distribution called the exponential distribution.A continuous random variable X has the Exponential Distribution with parameter $\theta > 0$ if it has a p.d.f. of the form

$$f(x;\theta) = f(x) = \frac{1}{\theta}\lambda^{-\frac{x}{\theta}}, x > 0$$

$$= 0, \text{ otherwise.}$$

The exponential distribution, which is an important probability distribution for lifetimes, is characterized by the following properties.

Properties

Mean $E(X) = \theta$

Variance $Var(X) = \theta^2$

Standard Deviation $\sigma = \theta$

So if λ is the mean of changes in the unit interval, then $\theta = \frac{1}{\lambda}$ is the mean waiting time for the first change.

Example 2:

Let the p.d.f. of X be $f(x) = \left(\frac{1}{2}\right)\lambda^{-x/2}, 0 \le x < \infty$.

i. What are the mean and variance of X?
ii. Calculate $P(X > 3)$
iii. Calculate $P(X > 5 | X > 2)$
iv. Calculate $P(X < 2)$

Solution

i. $E(X) = \theta = 2$ and $Var(X) = \theta^2 = 4$

ii. $P(X > 3) = \frac{1}{2}\int_3^{\infty} \lambda^{-\frac{x}{2}} dx = \lambda^{-\frac{3}{2}} = 0.2231$

iii. $P(X>5|X>2)=\dfrac{\frac{1}{2}\int\limits_{5}^{\infty}\lambda^{-\frac{x}{2}}}{\frac{1}{2}\int\limits_{2}^{\infty}\lambda^{-\frac{x}{2}}}=\dfrac{\lambda^{-\frac{5}{2}}}{\lambda^{-2/2}}=\lambda^{-\frac{3}{2}}=0.2231$

iv. $P(X<2)=\dfrac{1}{2}\int\limits_{0}^{2}\lambda^{-\frac{x}{2}}=1-\lambda^{-\frac{2}{2}}=1-\lambda^{-1}=0.6321$

6.2.1 The Normal Distribution

The normal distribution plays a central role in statistical theory and practice, particularly in the area of statistical inference. The normal distribution is perhaps the most important distribution in statistical applications since many measurements have (approximate) have normal distributions. The main reason for this is its role in the Central Limit Theorem (CLT).

The random variable X has a normal distribution if its p.d.f. is defined by

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

In this equation, the mean and standard deviation, which determine the location and spread of the distribution, are denoted by μ and σ, respectively. These are said to be the two parameters of the normal distribution satisfying $-\infty<\mu<\infty, 0<\sigma<\infty$. Briefly, we say that X is $N(\mu,\sigma^2)$.

Theorem: If the random variable X is $N(\mu,\sigma^2), \sigma^2>0$ then the random variable $Z=(X-\mu)/\sigma$ is $N(0,1)$.

Proof: The distribution function of Z is

$$P(Z\le z)=P\left(\frac{X-\mu}{\sigma}\le z\right)=P(X\le z\sigma+\mu)$$

$$=\int\limits_{-\infty}^{z\sigma+\mu}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{(x-\mu)^2}{2\sigma^2}\right]dx$$

changing the variable of integration by writing $w=(x-\mu)/\sigma$, hence $x=w\sigma+\mu$. We then obtain

$$=P(Z\le z)=\int\limits_{=\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\lambda^{-w^2/2}\,dw$$

This is the expression for $\Phi(z)$, the distribution function of a standardized normal random variable. Hence, Z is $N(0,1)$.

This fact considerably simplifies the calculations of probabilities concerning normally distributed variables, as seen in the following illustration:

Suppose that X is $N(\mu, \sigma^2)$, let $c_1 < c_2$, and since $P(X = c_1) = 0$, then

$$P(c_1 < X < c_2) = P(X < c_2) - P(X < c_1)$$

$$= P\left(\frac{X-\mu}{\sigma} < \frac{c_2-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} < \frac{c_1-\mu}{\sigma}\right)$$

$$= \Phi\left(\frac{c_2-\mu}{\sigma}\right) - \Phi\left(\frac{c_1-\mu}{\sigma}\right)$$

Note that $\Phi(-x) = 1 - \Phi(x)$.

The normal distribution possesses the following properties.

Properties

Mean $E(X) = \mu$

Variance $Var(X) = \sigma^2$

Standard Deviation $= \sigma$

Example 3 :

1. If Z is $N(0,1)$, find;
   i. $P(0.53 < Z < 2.06)$
   ii. $P(Z > 2.89)$
   iii. $P(6 \le X \le 12)$
2. If X is $N(75,100)$, find $P(X < 60)$.
3. If X is normally distributed with a mean of 6 and a variance 25, find $P(6 \le X \le 12)$

Solution

1.
   i. $P(0.53 < Z < 2.06) = \Phi(2.06) - \Phi(0.53) = 0.9803 – 0.7019 = 0.2784$
   ii. $P(Z > 2.89) = 1 - \Phi(2.89) = 1 - 0.9981 = 0.0019$
2. $P(X < 60) = P\left(\frac{X-75}{10} < \frac{60-75}{10}\right) = P(Z < -1.5) = 0.0668$

3. $P(6 \le X \le 12) = P\left(\frac{6-6}{5} \le Z \le \frac{12-6}{5}\right) = P(0 \le Z \le 1.2)$

$$= \Phi(1.2) - \Phi(0) = 0.8849 - 0.5000 = 0.3849$$

Summary

In this study session you have learnt about:

1. Uniform Distribution

Suppose that a continuous random variable X can assume values in a bounded interval only, say the open interval (a, b), and suppose the p.d.f. of X is given as

$$f(x; a, b) = f(x) = \frac{1}{b-a}, a < x < b$$

= 0, elsewhere.

This distribution is referred to as the Uniform or Rectangular Distribution on the interval (a, b) and is simply written as $X \sim U(a,b)$, where 'a' and 'b' are the parameters of the distribution. It provides a probability model for selecting a point at random from the interval (a, b).

2. Exponential Distribution

An exponential distribution is a continuous distribution related to the Poisson distribution. In the Poisson process, the number of changes occurring in a given interval is counted which results in discrete distribution.

But not only is the number of changes a random variable; the waiting times between successive changes are also random variables which are of the continuous type. The latter results into a distribution called the exponential distribution.

Self-Assessment Questions (SAQs) for study session 6

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 6.1 (Testing Learning Outcomes 6.1)

Let X have an exponential distribution with a mean of $\theta = 20$. Compute

i. $P(10 < X < 30)$

ii. $P(0 < X < 30)$

iii. $P(X > 30)$

iv. $P(X > 40 | X > 10)$

SAQ 6.2 (Testing Learning Outcomes 6.2)

Telephone calls enter a college switchboard according to a Poisson process on the average of two every 3 minutes. Let X denote the waiting time until the first call that arrives after 10 A.M.

i. What is the p.d.f. of X?

ii. Find $P(X > 2)$

SAQ 6.3 (Testing Learning Outcomes 6.3)

Customers arrive randomly at a bank teller's window. Given that one customer arrived during a particular 10-minute period, let X equal the time within the 10 minutes that the customer arrived. If X is $U(0,10)$, find

i. The p.d.f. of X.

ii. $P(X \geq 8)$

iii. $P(2 \leq X < 8)$

iv. $E(X)$

v. $Var(X)$

SAQ 6.4 (Testing Learning Outcomes 6.4)

Explain the relationship that exists between the Poisson and the Exponential distributions.

SAQ 6.5 (Testing Learning Outcomes 6.5)

If X is $N(75,100)$, find $P(X < 35)$ and $P(70 < X < 100)$.

SAQ 6.6 (Testing Learning Outcomes 6.6)

If Z is $N(0,1)$, find values of c such that

i. $P(Z \geq c) = 0.025$

ii. $P(|Z| \leq c) = 0.95$

$P(Z > c) = 0.05$

SAQ 6.7 (Testing Learning Outcomes 6.7)

Let X be $N(\mu, \sigma^2)$, so that $P(X < 89) = 0.90$ and $P(X < 94) = 0.95$. Find μ and σ^2.

SAQ 6.8 (Testing Learning Outcomes 6.8)

Show that the random variable $Z = (X - \mu)/\sigma$ is distributed $N(0,1)$.

SAQ 6.9 (Testing Learning Outcomes 6.9)

Suppose that $Z \sim N(0,1)$. Find the following probabilities:

i. $P(Z \leq 1.53)$

ii. $P(Z > -0.48)$

iii. $P(0.35 < Z < 2.01)$

iv. $P(|Z| > 1.28)$

Find the value of 'a' and 'b' such that

v. $P(Z \leq a) = 0.648$

vi. $P(|Z| \leq b) = 0.95$

References

Hogg R. V and Craig A. T (1995).*Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993).*Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982).*Introduction to Probability Theory and Statistical Inference.* Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics Reading* Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 7: Derivation of Means and Variances of Some Distributions

Expected duration: 1 week or 2 contact hours

Introduction

In study session three, we discussed the concept of mathematical expectation and we also obtained the means, variances and standard deviations of any given random variable. In the last three study session, we discussed some distributions where we gave the properties, that is, the means and variances of the distributions discussed. In this study session, we shall discuss the procedure for obtaining means and variances of some selected distributions.

Learning Outcomes from Study Session 7

At the end of this study session, you should be able to:

2.1 Explain Derivation of means and variances of some Distribution

7.1 Derivation of Means and Variances of Some Distributions

1.Discrete Case

The following are discrete case variance distributions:

- **Binomial Distribution**

Given that X has a Binomial distribution with parameters n and p. Obtain the mean and variance of X

Solution: The pdf of a Binomial distribution is

$$f(x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, ---- n$$

$$E(X) = \sum_{x=0}^{n} x f(x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= \sum x.\frac{n!}{x!(n-x)!} p^x q^{n-x}$$ Let x – 1 = y and n– 1 = m

= $\sum_{x} x.\frac{n(n-1)!}{x(x-1)!(n-x)!}P.p^{x-}q^{(n-1)-(x-1)}$

= $\sum \frac{m!}{y!(m-y)!}p^{y}q^{m-y}$

= $np\sum_{y=0}^{n}\binom{m.}{y}p^{y}q^{m-y}$

= np

$E(X^2) = \sum x^2 f(x)$

= $\sum x^2 \binom{n}{x} p^x q^{n-x}$

= $\sum x^2 \frac{n!}{x!(n-x)!}p^x q^{n-x}$

= $\sum x^2 \frac{n(n-1)!}{x(x-1)!(n-x)!}P.p^{x-1}q^{(n-1)-(x-1)}$ Let x – 1 = y n– 1 = m

= $np\sum \frac{x(n-1)!}{(x-1)!(n-x)!}p^{x-1}q^{(n-1)-(x-1)}$

= $np\sum_{y=0}^{m} y+1\frac{m!}{y!(m-y)!}p^{y}q^{m-y}$

= $np\left[\sum_{y=0}^{m} y\binom{m}{y}p^{y}q^{m-y} + \sum_{y=0}^{m}\binom{m}{y}p^{y}q^{m-y}\right]$

= np[mp + 1]

= np[(n – 1)p + 1]

= np[np – p + 1]

= $n^2p^2 - np^2 + np$

Var (X) = $E(X^2) - (E(X))^2$

= $n^2p^2 - np^2 + np - n^2p^2$

= $np - np^2$

= np(1 – p)

= npq

2. Poisson Distribution

Given that X has a Poisson distribution with parameter λ. Obtain the mean and variance of X.

Solution The p.d.f. of X is given by

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, ----\infty, \quad \lambda > 0$$

$$E(X) = \sum_{x=0}^{\infty} x \frac{\lambda^x \lambda^{-\lambda}}{x!} = \lambda^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^x}{x(x-1)!} = \lambda^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}$$

Let $k = x - 1$, then

$$E(X) = \lambda^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k+1}}{k!} = \lambda\lambda^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda\lambda^{-\lambda}\lambda^{\lambda} = \lambda,$$

Since from Maclaurin's series expansion $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \lambda^{\lambda}$

$\therefore E(X) = \lambda$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{\lambda^x \lambda^{-\lambda}}{x!} = \lambda^{-\lambda} \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x(x-1)!} = \lambda^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda^x}{(x-1)!}$$

Let $k = x - 1$, then

$$= \lambda^{-\lambda} \sum_{k=1}^{\infty} k + 1 \frac{\lambda^{k+1}}{k!} = \lambda\lambda^{-\lambda} \left[\sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \right] = \lambda \left[\sum_{k=0}^{\infty} k \frac{\lambda^k \lambda^{-\lambda}}{k!} + \sum_{k=0}^{\infty} \frac{\lambda^k \lambda^{-\lambda}}{k!} \right]$$

$$= \lambda[\lambda + 1]$$

Note that $\sum_{k=0}^{\infty} \frac{\lambda^k \lambda^{-\lambda}}{k!} = 1$

$$= \lambda^2 + \lambda$$

$$\text{var}(X) = E(X^2) - E^2(X)$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

B. Continuous Case

Uniform Distribution

Let X have a uniform distribution U (a, b) with p.d.f

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

Obtain the mean and variance of X.

Solution: $E(X) = \int_a^b xf(x)dx = \int_a^b x\frac{1}{b-a}dx$

$$= \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a}\left[\frac{b^2-a^2}{2}\right]$$

$$= \frac{1}{b-a}\left[\frac{(b+a)(b-a)}{2}\right]$$

$$= \frac{b+a}{2}$$

$$E(X^2) = \int_a^b x^2 f(x)dx$$

$$= \frac{1}{b-a}\int_a^b x^2 dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b = \frac{1}{b-a}\left[\frac{b^3-a^3}{3}\right]$$

$$= \frac{1}{b-a}\left[\frac{(b-a)(b^2+ab+a^2)}{3}\right]$$

$$= \frac{b^2+ab+a^2}{3}$$

$$V(X) = \frac{b^2+ab+a^2}{3} - \frac{(b^2+2ab+a^2)}{4}$$

$$= \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12}$$

$$= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

Summary

In this study session you have learnt about:

1. Derivation of Means and Variances of Some Distributions

- Discrete Case

The following are discrete case variance distributions:

- **Binomial Distribution**

Given that X has a Binomial distribution with parameters n and p. Obtain the mean and variance of X

Self-Assessment Questions (SAQs) for study session 7

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 7.1 (Testing Learning Outcomes 7.1)

The p.d.f of X is $f(x) = d/x^3$, 1 < x <∞, zero elsewhere.

i. Calculate the value of d so that f(x) is a p.d.f
ii. Find E(X)
iii. Show that Var(X) does not exist.

SAQ 7.2 (Testing Learning Outcomes 7.2)

Find the mean and variance of the following distributions.

i. $f(x) = \left(3/2\right)x^2, \quad -1 < x < 1$

ii. $f(x) = \frac{1}{2}, \qquad -1 < x < 1$

iii. $f(x) = \begin{cases} x+1 & -1 < x < 0 \\ 1-x & 0 \le x < 1 \end{cases}$

SAQ 7.3 (Testing Learning Outcomes 7.3)

Obtain the mean and variance of the exponential distribution.

SAQ 7.4 (Testing Learning Outcomes 7.4)

Derive the mean and variance of the normal distribution.

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995)**.***Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993)**.***Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory,* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading, Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and* Methods Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 8:Moment Generating Functions

Expected duration: 1 week or 2 contact hours

Introduction

In the preceding study session we have seen the importance of the mean, standard deviation and variance of a random variable X. For some distributions, it can be fairly difficult to obtain directly $E(X)$ and $E(X^2)$, the first and second moments.

You will discuss here a function of a real variable t that can be used to find $E(X)$ and $E(X^2)$ as well as other moments of X. Although, the moment-generating function (m.g.f.), if it exists, is a useful tool for determining moments, its major importance is in the fact that it uniquely determines the distribution.

Learning Outcomes from Study Session 8

At the end of this study session, you should be able to:

8.1 Explain Moment – Generating Functions

8.1 Moment – Generating Functions

The k^{th} moment about the origin of a random variable X is $\mu'_k = E(X^k)$ and the k^{th} moment about the mean is $\mu_k = E[X - E(X)]^k$

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

Eample 1: Let X be a random variable of the discrete type with p.d.f f(x) and space R. If there is a positive number h such that

$$E(e^{tx}) = \sum_{x \in R} e^{tx} f(x)$$

exists for $-h < t < h$, then the function of t defined by

$$M(t) = E(e^{tx})$$

is called the Moment – generating function (MGF) of X.

The derivatives of M (t) of all orders exist at t = 0. Thus

$$M'(t) = \sum_{x \in R} x e^{tx} f(x)$$

$$M''(t) = \sum_{x \in R} x^2 e^{tx} f(x)$$

and, for each positive integer r,

$$M^{(r)}(t) = \sum_{x \in R} x^r e^{tx} f(x)$$

Setting t = 0

$$M'(0) = \sum x f(x) = E(X)$$

$$M''(0) = \sum x^2 f(x) = E(X^2)$$

and in general

$$M^{(r)}(0) = \sum_{x \in R} x^r e^{tx} f(x) = E(X^r)$$

In particular, if the MGF exists,

$\mu = M'(0)$ and $\sigma^2 = M''(0) - [M'(0)]^2$

For continuous random variable X

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx \qquad -h < t < h$$

The moment-generating functions have the following properties:

1. $M_x(0) = 1$.
2. $|M_x(t)| \le 1$
3. M_x is uniformly continuous.
4. $M_{x+d}(t) = \lambda^{td} M_x(t)$, where d is a constant.
5. $M_{cx}(t) = M_x(ct)$, where c is a constant.
6. $M_{cx+d}(t) = \lambda^{td} M_x(ct)$

$$\left. \frac{d^n}{dt^n} M_x(t) \right|_{t=0} = E(X^n), \qquad n = 1,2,\ldots, \text{if } E\left|X^n\right| < \infty$$

The proofs of selected properties are as follows;

Property 1: $M_x(t) = E(\lambda^{tx})$, for $t = 0$, $M_x(t) = E(\lambda^0) = E(1) = 1$

Property 4: $M_{x+d}(t) = E(\lambda^{t(x+d)}) = E(\lambda^{tx}\lambda^{td}) = \lambda^{td} E(\lambda^{tx})$

$$= \lambda^{td} M_x(t)$$

Property 5: $M_{cx}(t) = E(\lambda^{ctx}) = E(\lambda^{(ct)x}) = M_x(ct)$

Example 2:Binomial Distribution.

Let X has a binomial distribution $b(n, p)$ with p.d.f

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2, ----, n$$

The MGF of X is

$$M(t) = E[e^{tX}] = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (pe^t)^x (1-p)^{n-x}$$

Using the formula for the binomial expansion with a = 1 – p and b = pet

$$= \sum \binom{n}{x} b^x a^{n-x}$$

$$= (a + b)^n$$

$\therefore$ $M(t) = [(1-p) + pe^t]^n \forall$ real values of t

The mean and variance are,

$$M'(t) = n[(1-p) + pe^t]^{n-1}(pe^t)$$

and $M''(t) = n(n-1)[(1-p) + pe^t]^{n-2}(pe^t)^2 + n[(1-p) + pe^t]^n(pe^t)$

Thus $\mu = E(X) = M'(0) = np$

and $\sigma^2 = E(X^2) - E^2(X) = M''(0) - [M'(0)]^2$

$$= n(n-1)p^2 + np - (np)^2$$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np(1-p)$$

Example 2:Poisson Distribution

Let X has a Poisson distribution with p.d.f

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, 2, ----$$

The MGF of X can be obtained as follows;

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} f(x)$$

$$= \frac{\sum e^{tx} \lambda^x e^{-\lambda}}{x!}$$

$$= e^{-\lambda} \sum \frac{(\lambda e^t)^x}{x!}$$

Note that, $e^X = \sum_{X=0}^{\infty} \frac{x^n}{n!}$

$$M(t) = e^{-\lambda} e^{\lambda e^t}$$

$$= e^{\lambda(e^t - 1)} \qquad \forall \text{ real values of t}$$

The first and second moments are;

$$E(X) = M'(0)$$

$$M'(t) = \lambda\lambda^t e^{\lambda(\lambda^t - 1)} = M'(0) = \lambda$$

$$M''(t) = (\lambda\lambda^t)e^{\lambda(\lambda^t - 1)} + (\lambda\lambda^t)^2 e^{\lambda(\lambda^t - 1)} = M''(0) = \lambda + \lambda^2$$

Then $Var(X) = M''(0) - [M'(0)]^2$

$$= \lambda + \lambda^2 - \lambda^2$$

$$= \lambda$$

Example 3: Normal Distribution

Let X have a normal distribution with p.d.f.

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

= 0 elsewhere

The moment-generating function of X can be calculated as follows;

$$M(t) = E(\lambda^{tx}) = \int_{-\infty}^{\infty} \lambda^{tx} \frac{1}{\sigma\sqrt{2\pi}} \lambda^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \lambda^{tx - \frac{1}{2\sigma^2}} \left(x^2 - 2\mu x + \mu^2\right) dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \lambda^{-\frac{\mu^2}{2\sigma^2}} \int_{-\infty}^{\infty} \lambda^{-\frac{1}{2\sigma^2}\left[x-(\mu+\sigma^2 t)\right]^2 - \left[(\mu+\sigma^2 t)^2 + \mu^2\right]} dx$$

$$= \lambda^{-\frac{1}{2\sigma^2}\left[\mu^2 - (\mu+\sigma^2 t)^2\right]} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \lambda^{-\frac{1}{2\sigma^2}\left[x-(\mu+\sigma^2 t)\right]^2} dx$$

$$= \lambda^{-\frac{1}{2\sigma^2}\left[\mu^2 - \mu^2 - 2\mu\sigma^2 t - (\sigma^2)^2 t^2\right]}$$

Since $\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \lambda^{-\frac{1}{2\sigma^2}\left[x-(\mu+\sigma^2 t)\right]^2} dx = 1$

$$M(t) = \lambda^{\mu t + \frac{\sigma^2 t^2}{2}}$$

The first and second moments are given as follows:

$$M(t) = \lambda^{\mu t + \frac{\sigma^2 t^2}{2}} = M'(t) = (\mu + \sigma^2 t)\lambda^{\mu t + \frac{\sigma^2 t^2}{2}}$$

The mean is therefore

$$M'(0) = \mu$$

To get the variance, we proceed as follows:

$$M''(t) = \sigma^2 \lambda^{\mu t + \frac{\sigma^2 t^2}{2}} + \left[(\mu + \sigma^2 t)(\mu + \sigma^2 t)\lambda^{\mu t + \frac{\sigma^2 t^2}{2}}\right]$$

$$M''(0) = \sigma^2 + \mu^2$$

Thus,

$$Var(X) = M''(0) - [M'(0)]^2 = \sigma^2 + \mu^2 - \mu^2$$

$$= \sigma^2$$

Example 4: Uniform Distribution.

Let X have a uniform distribution with p.d.f

$$f(x) = \frac{1}{b-a} \qquad a < x < b$$

Obtain the moment-generating function of X. Hence find its mean and variance.

$$M(t) = \int_a^b e^{tx} \frac{1}{b-a} dx$$

$$= \frac{1}{b-a} \int_a^b e^{tx} dx = \frac{1}{b-a} \left[\frac{e^{tx}}{t} \right]_a^b$$

$$= \frac{1}{b-a} \left[\frac{e^{tb} - e^{ta}}{t} \right]$$

$$= \frac{e^{tb} - e^{ta}}{t(b-a)}$$

$$= \frac{1}{t(b-a)} \left[e^{tb} - e^{ta} \right]$$

Note that $e^X = 1 + \frac{X}{1!} + \frac{X^2}{2!} + \frac{X^3}{3!} ------$

$$e^{tb} = 1 + \frac{tb}{1!} + \frac{t^2 b^2}{2!} + \frac{t^3 b^3}{3!} ------$$

$$e^{ta} = 1 + \frac{ta}{1!} + \frac{t^2 a^2}{2!} + \frac{t^3 a^3}{3!} ------$$

$$M(t) = \frac{1}{t(b-a)} \left[1 + tb + \frac{t^2 b^2}{2} + \frac{t^3 b^3}{6} - 1 - ta - \frac{t^2 a^2}{2} - \frac{t^3 a^3}{6} \right]$$

$$= \frac{t}{t(b-a)} \left[(b-a) + \frac{tb^2}{2} - \frac{ta^2}{2} + \frac{t^2 a^3}{6} - \frac{t^3 a^3}{6} \right]$$

$$M'(t) = \frac{1}{b-a} \left[\frac{b^2}{2} - \frac{a^2}{2} + \frac{2tb^3}{6} - \frac{2ta^3}{6} \right]$$

The mean is thus,

$$M'(0) = \frac{1}{b-a} \left[\frac{b^2 - a^2}{2} \right]$$

And the variance is

$$M''(0) = \frac{1}{b-a} \left[\frac{2b^3}{6} - \frac{2a^3}{6} \right]$$

$$= \frac{1}{b-a} \left[\frac{b^3 - a^3}{3} \right]$$

$$Var(X) = M''(0) - [M'(0)]^2 = \left[\frac{b^3 - a^3}{3(b-a)} \right] - \left[\frac{b^2 - a^2}{2(b-a)} \right]^2$$

$$\frac{(b-a)^2}{12}$$

Summary

In this study session, we have learned the following:

1. The importance of moment- generating function.
2. How to obtain the means and variances of distributions from m.g.f.

Self-Assessment Questions (SAQs) for study session 8

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 8.1 (Testing Learning Outcomes 8.1)

Define the r^{th} moment about the mean.

SAQ 8.2 (Testing Learning Outcomes 8.2)

Define the moment-generating function of a discrete random variable X.

SAQ 8.3 (Testing Learning Outcomes 8.3)

Define the moment-generating function of a discrete random variable X.

SAQ 8.4 (Testing Learning Outcomes 8.4)

Give the proof of the remaining properties given above.

SAQ 8.5 (Testing Learning Outcomes 8.5)

Obtain the m.g.f. of the exponential distribution and hence or otherwise find the mean and variance.

SAQ 8.6 (Testing Learning Outcomes 8.6)

Find the m.g.f. when the p.d.f. of X is defined by

i. $f(x)=\frac{1}{5},\ x=1,2,3$

ii. $f(x)=1,\ x=5$

iii. $f(x)=\frac{5-x}{10},\ x=1,2,3,4.$

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995)**.***Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993)**.***Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 9: Cumulant Generating Functions

Expected duration: 1 week or 2 contact hours

Introduction

You saw in the last study session that the moments of any given random variable can be computed using the moment-generating function approach. For many random variables the cumulant generating function proves easier to use in evaluating the mean and variance.

The reason for this simplicity is that the first two derivatives of c.g.f. of X written as $C_x(t)$evaluated at $t=0$directly give the mean and the variance of X for many of the standard random variables that we will discuss, these two derivatives are very easy to compute. Either the cumulant generating function or the moment generating function can be used to evaluate means and variances (and other moments).

Learning outcomes from Study Session 9

At the end of this study session, you should be able to:

9.1 Discuss the concept of the cumulant generating function;

9.1 Cumulant Generating Functions

The cumulant generating function is defined to be the natural log of the moment generating function (assuming it exists). That is, M(t) is the moment generating function of X, then the cumulant generating function for X is

$$C_x(t) = \ln M(t)$$

i.e $$M(t) = \lambda^{C_x(t)}$$

So if $C_x(t)$ were known, it is easy to find $M_x(t)$. Then

$$\frac{d}{dt}C(t) = \frac{M'(t)}{M(t)}$$

$$\frac{d^2}{dt^2}C(t) = \frac{M''(t)M(t) - (M'(t))^2}{(M(t))^2}$$

where

$$\frac{d}{dt}M(t) = M'(t), \frac{d^2}{dt^2}M(t) = M''(t)$$

Since M(0) = 1 i.e M(t) = E(e^{tx}), then M(0) = E(e^0) = 1

$$\left.\frac{d}{dt}C(t)\right|_{t=0} = \frac{M'(0)}{M(0)} = \frac{M_1}{1} = \mu$$

$$\left.\frac{d^2}{dt^2}C(t)\right|_{t=0} = \frac{M''(0)M(0) - (M'(0))^2}{M^2(0)}$$

$$= \frac{M_2 - (M_1)^2}{1} = \sigma^2$$

The first two derivatives of $C_x(t)$ evaluated at t = 0 directly give the mean and variance of X.

Example 1: Let X be a random variable with M(t) = ¼(1 + e^t)^2. Find the cumulant generating function.

Solution: C(t) = ln M(t) = In ¼ + 21n(1 + e^t)

$$\frac{d}{dt}C(t) = \frac{2e^t}{1+e^t}$$

$$\frac{d^2}{dt^2}C(t) = \frac{2e^t}{(1+e^t)^2}$$

$$\left.\frac{d}{dt}C(t)\right|_{t=0} = \frac{2}{2} = 1 = \mu$$

$$\left.\frac{d^2}{dt^2}C(t)\right|_{t=0} = \frac{2}{2^2} = \frac{1}{2} = \sigma^2$$

Example 2: A discrete random variable X has mgf

M(t) = exp[2(e^t – 1)]

Find the c.g.f of X and use it to evaluate μ and σ^2.

Solution: C(t) = In M(t) = In exp[2(e^t – 1)] = 2[e^t – 1]

$$\frac{d}{dt}C(t) = 2e^t$$

$$\frac{d^2}{dt^2}C(t) = 2e^t$$

$$\frac{d}{dt}C(t)\bigg|_{t=0} = 2 = \mu$$

$$\frac{d^2}{dt^2}C(t)\bigg|_{t=0} = 2 = \sigma^2$$

Example 3: 1. The m.g.f. of a normal random variable X is

$M(t) = e^{\mu t} + \frac{1}{2}\sigma^2 t^2$

Find the c.g.f and hence it's mean and variance

2. i. The mgf of a binomial random variable X is

$M(t) = (q + Pe^t)^n$

ii. For a Poisson random variable X is

$M(t) = e^{\lambda(e^t - 1)}$

Find the means and variances of X.

Solution:

1. $M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$$C(t) = InM(t) = \mu t + \frac{1}{2}\sigma^2 t^2$$

$$\frac{d}{dt}C(t) = \mu + \sigma^2 t$$

$$\frac{d^2}{dt^2}C(t) = \sigma^2$$

$$\frac{d}{dt}C(t)\Big/_{t=0} = \mu$$

2i. $M(t) = (q + Pe^t)^n$

$C(t) = \text{In } M(t) = n \text{ In } (q + Pe^t)$

$$\frac{d}{dt}C(t) = \frac{nPe^t}{q + Pe^t}$$

$$\frac{d^2}{dt^2}C(t) = \frac{(q + Pe^t)nPe^t - nPe^t(Pe^t)}{(q + Pe^t)^2}$$

$$\frac{d}{dt}C(t)\bigg|_{t=0} = nP$$

$$\frac{d^2}{dt^2}C(t)\Big|_{t=0} = \frac{nPqe^t - nP^2e^{2t} - nP^2e^{2t}}{(q+Pe^t)^2} = nPq$$

ii. C(t) = In M(t) = $\lambda(e^t - 1)$

$$\frac{d}{dt}C(t) = \lambda e^t$$

$$\frac{d^2}{dt^2}C(t) = \lambda e^t$$

$$\frac{d}{dt}C(t)\Big|_{t=0} = \lambda$$

$$\frac{d^2}{dt^2}C(t)\Big|_{t=0} = \lambda$$

Example 4: Let X have an exponential distribution with p.d.f

$$f(x) = \frac{1}{\theta}e^{-x/\theta} \qquad 0 \leq x < \infty, \quad \theta > 0$$

Obtain the MGF of X, hence or otherwise find μ and σ^2.

Solution:

$$M(t) = \int_0^\infty e^{tx}\left(\frac{1}{\theta}\right)e^{-x/\theta}dx$$

$$= \lim_{b\to\infty}\int_0^b\left(\frac{1}{\theta}\right)e^{-(1-\theta t)x/\theta}dx$$

$$= \lim_{b\to\infty}\left[-\frac{e^{-(1-\theta t)x/\theta}}{1-\theta t}\right]_0^b$$

$$= \frac{1}{1-\theta t}, t < \frac{1}{\theta}$$

$$M'(t) = \frac{\theta}{(1-\theta t)^2}$$

$$M''(t) = \frac{2\theta^2}{(1-\theta t)^2}$$

$$\mu = M'(0) = \theta$$

$$\sigma^2 = M''(0) - [M'(0)]^2$$

$$= \theta^2$$

Example 5: Let X have the pdf

$$f(x) = \begin{cases} xe^{-x}, & 0 \le x < \infty \\ 0 \quad, & \text{elsewhere} \end{cases}$$

$$M(t) = \int_0^\infty e^{tx} x e^{-x} dx$$

$$= \quad \lim_{b\to\infty} \int_0^b x e^{-(1-t)x} dx$$

$$= \quad \lim_{b\to\infty} \left[-\frac{xe^{-(1-t)x}}{1-t} - \frac{e^{-(1-t)x}}{(1-t)^2} \right]_0^b$$

$$= \quad \lim_{b\to\infty} \left[-\frac{be^{-(1-t)b}}{1-t} - \frac{e^{-(1-t)b}}{(1-t)^2} \right] + \frac{1}{(1-t)^2}$$

$$M(t) = \frac{1}{(1-t)^2}, \text{ provided } \quad t < 1$$

Summary

In this study session you have learnt about:

1.Cumulant Generating Functions

The cumulant generating function is defined to be the natural log of the moment generating function (assuming it exists). That is, M(t) is the moment generating function of X, then the cumulant generating function for X is

$$C_x(t) = \ln M(t)$$

1. The notion of cumulant generating function.
2. How moment-generating function relates to the cumulant generating function.
3. How to derive cumulant generating function from the moment-generating function.

Evaluating the means and variances of random variables using c.g.f. from a given m.g.f.

Self-Assessment Questions (SAQs) for study session 9

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 9.1 (Testing Learning Outcomes 9.1)

A random variable X has moment-generating function

$$M_x(t) = \left(0.25 + 0.75\lambda^t\right)^{12},$$

Find the cumulant generating function of X and use it to obtain the mean and variance of X.

SAQ 9.2 (Testing Learning Outcomes 9.2)

Find the moment-generating function, mean and variance of X if the p.d.f of X is

$$f(x) = \left(\frac{1}{2}\right)\left(\frac{2}{3}\right)^x,\ x = 1,2,3,4,...$$

From the m.g.f. obtained, find the c.g.f of X.

SAQ 9.3 (Testing Learning Outcomes 9.3)

1. Given the following m.g.fs of X, find the c.g.fs and hence or otherwise obtain μ and σ^2.

   i. $M(t) = \lambda^{4\left(\lambda^t - 1\right)}$

   ii. $M(t) = \dfrac{0.3\lambda^t}{1 - 0.7\lambda^t},\ t < -\ln(0.7)$

   iii. $M(t) = 0.3\lambda^t + 0.4\lambda^{2t} + 0.2\lambda^{3t} + 0.1\lambda^{4t}$

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995).*Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993).*Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 10: Chebyshev's Inequality

Expected duration: 1 week or 2 contact hours

Introduction

Limit theorems basically is concerned with finding approximations to distributions of statistics and / or statistics themselves. We cannot attain finding limits without looking at some important inequalities. One of the most important inequalities used in probability is the Chebyshev's inequality.

Learning outcomes from Study Session 10

At the end of this study session, you should be able to:

10.1 Explain chebyshev's Inequality

10.1 Chebyshev's Inequality

Many important inequalities exist which relate expectations and probabilities. A lot of these are variations on the basic inequality called Markov's inequality.

Chebyshev's inequality will be used to show that the sample mean, $\bar{x}$, is a good statistic to use to estimate a population mean, μ; the relative frequency of success in n Bernoulli trials, x/n, is a good statistic for estimating P; the empirical distribution function, $F_n(x)$, can be used to estimate the theoretical distribution function F(x).

The effect of the sample size n on these estimates is discussed.First show that Chebyshev's inequality gives added significance to the standard deviation in terms of bounding certain probabilities. The inequality is valid for all distributions for which the standard deviation exists.

Theorem 1: (Markov's Inequality) If X is a random variable and $U(X)$ is a non-negative real-valued function, then for any positive constant c > 0,

$$P[U(X) \geq c] \leq \frac{E[U(X)]}{c}$$

Proof: If A = $\{x | U(x) \geq c\}$, then for a continuous random variable,

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

$$= \int_A U(x)f(x)dx + \int_{A^{c^c}} U(x)f(x)dx$$

$$\geq \int_A U(x)f(x)dx$$

$$\geq \int_A cf(x)dx$$

$$= cP[X \in A]$$

$$= cP[U(X) \geq c]$$

Theorem 2: (Chebyshev's Inequality) If the random variable X has a mean μ and variance σ^2, then for every k $\geq$ 1

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}$$

Proof: Let f(x) denote the pdf of X. then

$$\sigma^2 = E[(X-\mu)^2] = \sum_{x \in R} (x-\mu)^2 f(x)$$

$$= \sum_{x \in A} (x-\mu)^2 f(x) + \sum_{x \in A'} (x-\mu)^2 f(x)$$

where $A = \{x; |x-\mu| \geq k\sigma\}$

Hence $\sigma^2 \geq \sum_{x \in A} (x-\mu)^2 f(x)$

However, in A, $|x - \mu| \geq k\sigma$; so

$$\sigma^2 \geq \sum_{x \in A} (k\sigma)^2 f(x) = k^2\sigma^2 \sum_{x \in A} f(x)$$

But $\sum_{x \in A} f(x) = P(X \in A)$ and thus

$$\sigma^2 \geq k^2\sigma^2 P(X \in A) = k^2\sigma^2 P(|X - \mu| \geq k\sigma)$$

That is $P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$

Corollary If $\varepsilon = k\sigma$, then

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

In words, Chebyshev's inequality states that the probability that X differs from its mean by at least k standard deviations is less than or equal to $1/k^2$. It follows that the probability that X differs from its mean by less than k standard deviations is at least $1 - 1/k^2$.

That is

$$P(|X-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

From the corollary, it also follows that

$$P(|X-\mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

Thus, Chebyshev's inequality can be used as a bound for certain probabilities. However, in many instances, the bound is not very close to the true probability.

Example 1: If it is known that X has a mean of 25 and a variance of 16, then, since $\sigma = 4$, a lower bound for P (17 < X < 33) is given by

$$P(17 < X < 33) = P(|X-25| < 8)$$

$$= \quad P(|X-\mu| < 2\sigma) \geq 1 - \frac{1}{4} = 0.75$$

and an upper bound for P(|X – 25|≥ 12) is found to be

$$P(|X-25| \geq 12) = P(|X-\mu| \geq 3\sigma) \leq \frac{1}{9}$$

Example 2: If X is a random variable with mean 33 and variance 16, use Chebyshev's inequality to find

a. A lower bound for P(23 < X < 43)

b. An upper bound for P(|X – 33|≥ 14)

Solution:

a. $P(|X-33| < 10) = P(|X-\mu| < 2.5\sigma) \geq 1 - \frac{1}{2.5^2}$

$$= \quad 1 - 0.16$$

$$= \quad 0.84$$

b. $P(|X-33| \geq 14) = P(|X-\mu| \geq 3.5\sigma) \leq \frac{1}{3.5^2} = \frac{1}{12.25}$

$$= \quad 0.082$$

Example 3: Let X denote the outcome when rolling a fair die. Then $\mu = 7/2$ and $\sigma^2 = 35/12$. Note that the maximum deviation of X from μ equals $5/2$. Express this deviation in terms of the number of standard deviations; that is find k where $k\sigma = 5/2$. Determine a lower bound for P(|X – 3.5|< 2.5).

Solution: $k\sigma = 5/2$

$$k = \frac{2.5}{\sigma}$$

but $\sigma = \sqrt{35/12}$

$$\therefore k = \frac{2.5}{\sqrt{35/12}} = 1.464$$

$$P(|X - 3.5| < 2.5) = P(|X - \mu| < 1.4645\sigma) \geq 1 - \frac{1}{1.464^2}$$

$$= \quad 1.0.467$$

$$= \quad 0.5328 \simeq 0.533$$

If Y is the number of successes in n Bernoulli trials with probability p of success on each trial, then Y is b(n, p). Furthermore, Y/n gives the relative frequency of success, and when p is unknown, Y/n can be used as an estimate of p. To gain some insight into the closeness of Y/n to p, we shall use Chebyshev's inequality with $\varepsilon > 0$.

$$p\left(\left|\frac{Y}{n} - p\right| \geq \varepsilon\right) = p(|Y - np| \geq n\varepsilon)$$

$$= \quad p\left(|Y - np| \geq \frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\sqrt{npq}\right)$$

However, μ = np and $\sigma = \sqrt{npq}$ are the mean and the standard deviation of Y so that, with $k = \sqrt{n}\varepsilon / \sqrt{pq}$, we have

$$p\left(\left|\frac{Y}{n} - p\right| \geq \varepsilon\right) = p(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2} = \frac{pq}{n\varepsilon^2} \quad \text{--------*}$$

or, equivalently,

$$p\left(\left|\frac{Y}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{pq}{n\varepsilon^2}$$

when p is completely unknown, pq = p(1 – p) is a maximum when p = ½ in order to find a lower bound for the probability in equation (*). That is

$$1 - \frac{pq}{n\varepsilon^2} \geq 1 - \frac{(1/2)(1/2)}{n\varepsilon^2}$$

For example, if ε = 0.05 and n = 200

$$p\left(\left|\frac{Y}{400} - p\right| < 0.05\right) \geq 1 - \frac{(1/2)(1/2)}{200(0.0025)} = 0.75$$

Note that Chebyshev's inequality is applicable to all distributions with a finite variance, thus the bound is not always a tight one; i.e, the bound is not necessarily close to the true probability.

In general, it should be noted that with fixed $\varepsilon > 0$ and $0 < p < 1$,

$$\lim_{n\to\infty} p\left(\left|\frac{Y}{n} - p\right| < \varepsilon\right) \geq \lim_{n\to\infty}\left(1 - \frac{pq}{n\varepsilon^2}\right) = 1$$

Since the probability of every event is less than or equal to 1, then

$$\lim_{n\to\infty} p\left(\left|\frac{Y}{n} - p\right| < \varepsilon\right) = 1$$

That is, the probability that the relative frequency Y/n is within ε of p is close to 1 when n is large enough. Thus is one form of the law of large numbers.

Summary

In this study session you have learnt about:

1. Chebyshev's Inequality

Many important inequalities exist which relate expectations and probabilities. A lot of these are variations on the basic inequality called Markov's inequality.

Chebyshev's inequality will be used to show that the sample mean, $\bar{x}$, is a good statistic to use to estimate a population mean, μ; the relative frequency of success in n Bernoulli trials, x/n, is a good statistic for estimating P; the empirical distribution function, $F_n(x)$, can be used to estimate the theoretical distribution function F(x).

Self-Assessment Questions (SAQs) for study session 10

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 10.1 (Testing Learning Outcomes 10.1)

1. If E(X) = 17 and $E(X^2) = 298$. Use Chebyshev's inequality to determine
   i. A lower bound for P (10 < X < 24).
   ii. An upper bound for P(|X – 17|≥ 16).

SAQ 10.2 (Testing Learning Outcomes 10.2)

State and prove the Chebyshev's inequality.

SAQ 10.3 (Testing Learning Outcomes 10.3)

If X is a random variable such that E(X) = 3 and $E(X^2)$ =13, use Chebyshev's inequality to determine a lower bound for the probability $P(-2 < X < 8)$.

SAQ 10.4 (Testing Learning Outcomes 10.4)

If Y is $b(n, 0.5)$, give a lower bound for $P\left(\left|\frac{Y}{n} - 0.5\right| < 0.08\right)$ when

i. n = 100

ii. n = 500

iii. n = 1000

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995).*Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993).*Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 11: Central Limit Theorem

Expected duration: 1 week or 2 contact hours

Introduction

The relationship between the shapes of the population distribution and the sampling distribution of the mean can be summarized in what is often referred to as the most important theorem in statistics, namely, the central limit theorem.

The central limit theorem is concerned with the probability distribution of sums of random variables as n, the number of terms in the sum, increases without bound. The central limit theorem is frequently relied on to justify the assumption of a normal probability distribution for any random variable whose value can be thought as the accumulation of a large number of independent quantities.

Learning from Study Session 11

At the end of this study session, you should be able to:

11.1 Discuss Central Limit Theorem and approximate distribution in cases where the exact distribution is unknown or intractable.

11.1 Central Limit Theorem

Theorem: If $\bar{X}$ is the mean of a random sample $X_1, X_2, ---, X_n$ of size n from a distribution with a finite mean μ and a finite positive variance σ^2 then the distribution of

$$W = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$$

is N(0, 1) in the limit as $n \to \infty$

Proof: Let $M_X(t) = M(t)$

$M_S(t) = (M(t))^n$ where $S = \sum X_i$

$$M_w(t) = E[e^{tw}] = E\left[e^{t\left(\frac{S - n\mu}{\sqrt{n}\sigma}\right)} \right]$$

$$= \quad e^{-\frac{\sqrt{n}\mu}{\sigma}t} E\left[e^{\frac{t}{\sqrt{n}\sigma}S} \right]$$

$$= \quad e^{-\left(\frac{\sqrt{n}\mu}{\sigma}t\right)} \left(E\left(e^{\frac{t}{\sqrt{n}\sigma}X} \right) \right)^n$$

$$= \quad e^{-\frac{\sqrt{n}\mu}{\sigma}t} \left(M\left(\frac{t}{\sqrt{n}\sigma} \right) \right)^n$$

Loge $\quad M_w(t) = -\dfrac{\sqrt{n}\mu}{\sigma}t + n\log_e M\left(\dfrac{t}{\sqrt{n}\sigma}\right)$

Note $\quad \log_e X = 1 + X + \dfrac{X^2}{2!} + \dfrac{X^3}{3!} + ----$

$$\log_e M_w(t) = 1 + M'(0)t + \frac{M''(0)t^2}{2} + ----$$

$$= \quad 1 + \mu t + \frac{(\mu^2 + \sigma^2)}{2}$$

$$In\ M_w(t) = -\frac{\sqrt{n}\mu t}{\sigma} + nIn\left(1 + \frac{\mu t}{\sqrt{n}\sigma} + \frac{\mu^2 + \sigma^2}{2}\frac{t^2}{n\sigma^2} ----\right)$$

$$\log_e(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3}$$

$$\Rightarrow InM_w(t) = -\frac{\sqrt{n}\mu t}{\sigma} + n\left[\frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)t^2}{2n\sigma^2} - \frac{1}{2}\left(\frac{\mu t}{\sqrt{n}\sigma} + \frac{(\mu^2 + \sigma^2)t^2}{2n\sigma^2}\right)^2\right]$$

$$= \frac{\sqrt{n}\mu t}{\sigma} + \frac{\sqrt{n}\mu t}{\sigma} + \frac{(\mu^2 + \sigma^2)t^2}{2\sigma^2} - \frac{\mu^2 t^2}{2\sigma^2} - \frac{1}{2}\frac{\mu t^3(\mu^2 + \sigma^2)}{\sqrt{n}\sigma^3}$$

$$= t^2/2 - \frac{\frac{1}{2}\mu t^3(\mu^2 + \sigma^2)}{\sqrt{n}\sigma^3}$$

$= t^2/2$, since N(0, 1)

Example 1: Let $\overline{X}$ denote the mean of a random sample of size n = 15 from the distribution whose pdf is $f(x) = \left(3/2\right)x^2$, - 1 <$x$ 1. Approximate $P\left(0.03 \le \overline{X} \le 0.15\right)$.

Solution:

$$f(x) = \left(3/2\right)x^2$$

$$E(X) = 3/2\int_{-1}^{1} x^3 dx = 3/2\left[\frac{x^4}{4}\right]_{-1}^{1} = 0$$

$$E(X^2) = 3/2\int_{-1}^{1} x^4 dx = 3/2\left[\frac{x^5}{5}\right]_{-1}^{1} = 3/10[1+1] = 6/10 = 3/5$$

$$V(X) = E(X^2) - \mu^2 = 3/5$$

$$P\left(0.03 \le \overline{X} \le 0.15\right) = P\left[\frac{0.03-0}{\sqrt{3/5}/\sqrt{15}} \le \frac{\overline{X}-0}{\sqrt{3/5}/\sqrt{15}} \le \frac{0.15-0}{\sqrt{3/5}/\sqrt{15}}\right]$$

$= \quad P[0.15 \le W \le 0.75]$

$= \quad \phi(0.75) - \phi(0.15)$

$= \quad 0.7734 – 0.5596$

$= \quad 0.2138$

OR $n\overline{X} = \sum X_i = Y \sim N\left(0, 3/5\, n\right)$

$= \quad Y \sim N(0, 9)$

$$P(0.03 \le \overline{X} \le 0.15) = P(0.03n \le n\overline{X} \le 0.15)$$

$= \quad P(0.45 \le Y \le 2.25)$

$= \quad P\left[\frac{0.45}{\sqrt{9}} \le W \le \frac{2.25}{\sqrt{9}}\right]$

$= \quad P[0.15 \le W \le 0.75]$

$= \quad \phi(0.75) - \phi(0.15)$

$= \quad 0.2138$

Example 2: Uniform: Let $\overline{X}$ be the mean of a random sample of size 12 from the uniform distribution on the interval (0, 1). Approximate $P\left(\frac{1}{2} \leq \overline{X} \leq \frac{2}{3}\right)$

Solution: $E(X_i) = \frac{1}{2}$

$$V(X_i) = \frac{1}{12}$$

$$Y = \sum_{i=1}^{n} X_i \text{ ; then } Y \sim N\left(\frac{n}{2}, \frac{n}{12}\right)$$

$$P(6 \leq Y \leq 8) = P\left(\frac{6-6}{1} \leq \frac{Y-6}{1} \leq \frac{8-6}{1}\right)$$

$$= P(0 \leq W \leq 2)$$

$$= \Phi(2) - \Phi(0)$$

$$= 0.9772 - 0.5$$

$$= 0.4772$$

Example 3: Exponential: Let $\overline{X}$ be the mean of a random sample of size 36, from an exponential distribution with mean 3. Approximate, $P(2.5 \leq \overline{X} \leq 4)$.

$$E(X_i) = \theta, \ V(X_i) = \theta^2$$

$$Y = \sum X_i, \qquad Y \sim N(108, \ 324)$$

$$P(2.5 \leq \overline{X} \leq 4) \equiv P(90 \leq Y \leq 144)$$

$$= P\left[\frac{90-108}{18} \leq \frac{Y-108}{18} \leq \frac{144-108}{18}\right]$$

$$= P[-1 \leq W \leq 2]$$

$$= \Phi(2) - \Phi(-0)$$

$$= 0.9772 - 0.158$$

$$= 0.8185$$

Example 4: Binomial: Let X_1, X_2, --- X_n be a random sample from a Bernoulli distribution with mean μ = P and a variance σ^2 = P(1 – P), 0 < P < 1. Then $Y = \sum_{1}^{n} X_i$ is b(n, P)

The CLT states that the distribution of

$$W = \frac{Y - nP}{\sqrt{nP(1-P)}} = \frac{\overline{X} - nP}{\sqrt{P(1-P)/n}} \quad \sim N(0,1) \text{ as } n \to \infty \text{ is } \left(K - \tfrac{1}{2} < y < k + \tfrac{1}{2}\right) \text{ if}$$

$nP \geq 5, n(1-P) \geq 5$

Example 5: Poisson: If Y has a Poisson distribution with mean λ, then the distribution of

$$W = \frac{Y - \lambda}{\sqrt{\lambda}} \sim \text{N(0, 1) when } \lambda \text{ is sufficiently large.}$$

Example 6: Let Y be b(36, ½). Then np = 36 x ½ = 18 > 5 and n(1 – P) > 5

$$P(12 < Y \leq 18) = P(12.5 \leq Y \leq 18.5)$$

$$= P\left(\frac{12.5 - 18}{\sqrt{9}} \leq \frac{Y - 18}{\sqrt{9}} \leq \frac{18.5 - 18}{\sqrt{9}}\right)$$

$$= \Phi(0.167) - \Phi(-1.833)$$

$$= 0.5329$$

Note that 12 was increased to 12.5 because P(Y = 12) is not included in the desired probability.

Example 7: Let Y have the binomial distribution of b(10, ½). Approximate

P(3 ≤ Y < 6)

P(3 ≤ Y < 6) = P(2.5 ≤ Y ≤ 5.5) because P(Y = 6) is not included in the probability.

$$P\left(\frac{2.5 - 5}{\sqrt{10/4}} \leq \frac{Y - 5}{\sqrt{10/4}} \leq \frac{5.5 - 5}{\sqrt{10/4}}\right)$$

$$= \Phi(0.316) - \Phi(-1.581)$$

$$= 0.6240 - 0.0570$$

$$= 0.5670$$

Summary

In this study session you have learnt about:

1. The theory of central limit theory.
2. To approximate exact probability distributions for sums of independent random variables.The central limit theorem also gives a good approximation when the number of random variables summed together is large

Theorem: If $\bar{X}$ is the mean of a random sample X_1, X_2. ---, X_n of size n from a distribution with a finite mean μ and a finite positive variance σ^2 then the distribution of

$$W = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$$

Self-Assessment Questions (SAQs) for study session 11

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 11.1 (Testing Learning Outcomes 11.1)

A nurseryman plants 115 cuttings of ivy in every flat he prepares. Assume the probability that an individual cutting will develop roots is 0.9 and approximate the probability that the average number of rooted cuttings (per flat) in 50 flats is less than 100.

SAQ 11.2 (Testing Learning Outcomes 11.2)

1. A big city car dealer opens 365 days per year; the number of sales he makes per day is a Poisson random variable with parameter $\mu = 2$, independently from one day to another. Let Y be the number of sales he makes in a year, approximate
   i. $P(Y \geq 700)$
   ii. $P(\leq 800)$
   iii. $P(700 \leq Y \leq 800)$

SAQ 11.3 (Testing Learning Outcomes 11.3)

Let $Y = X_1 + X_2 + ... + X_{15}$ be the sum of a random sample of size 15 from the distribution whose p.d.f. is $f(x) = \frac{3}{2}x^2$, $-1 < x < 1$. Approximate $P(-0.3 \le \overline{X} \le 1.5)$

SAQ 11.4 (Testing Learning Outcomes 11.4)

Let $\overline{X}$ be the mean of a random sample of size 36 from an exponential distribution with mean 3. Approximate $P(2.5 \le \overline{X} \le 4.0)$.

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995).*Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993).*Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 12: Joint Probability Density Function

Expected duration: 1 week or 2 contact hours

Introduction

Up till now, we have only examined the probabilities (outcomes) as a function of one variable. We have also spent time studying the concept of a random variable and have studied some simple models that led to several other frequently used probability distributions.

The random variables considered so far are called one dimensional, because the observed value for a random variable can be thought of as a single point on a real line. In almost all applications, random variables do not occur singly. We need to develop tools necessary to describe the behavior of two, three or more random variables simultaneously.

For example, the hardness and tensile strength of a manufactured piece of steel may be of interest and so a p.d.f. May be necessary for experimental outcome. In order to deal with situations such as these, we will extend certain definitions as well as give new ones.

Learning Outcomes From Study Session 12

At the end of this study session, you should be able to:

12.1 Highlight on Joint Probability Density Function
12.2 State the Marginal Probability Density Functions

12.1 Joint Probability Density Function

Definition 1: Let X and Y be two random variable defined on a discrete probability space. Let R denote the corresponding two-dimensional space of X and Y, the two random variables of the discrete type. The probability that X = x, and Y = x is denoted by f(x, y) = P(X = *x*, Y = y). The function f(*x*, y) is called the joint p.d.f. of X and Y and has the following properties:

a. $0 \le f(x, y) \le 1$
b. $\sum\sum_{(x,y)\leftarrow R} f(x, y) = 1$
c. $P[(X,Y) \in A] = \sum_{(x,y)\in A} f(x, y)$, where A is a subset of the space R.

Example 1 : Roll a pair of unbiased dice. For each of the 36 sample points with probability $1/36$, Let X denotes the smaller and Y the larger outcome on the dice.

Definition 2: The joint probability density function of two continuous-type random variables is an integrable function $f(x, y)$ with the following properties:

a. $f(x, y) \geq 0$.

b. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1$.

c. $P[(x, y) \in A] = \iint_A f(x, y)dxdy$,

where $\{(X,Y) \in A\}$ is an event defined in the plane. Property (c) implies that $P[(X,Y) \in A]$ is the volume of the solid over the region A in the $xy-plane$ and bounded by the surface $z = f(x, y)$.

12.2 The Marginal Probability Density Functions

Definition 3: Let X and Y have the joint p.d.f $f(x, y)$ with space R. The p.d.f of X alone, called the marginal p.d.f of X is defined by

$$f_1(x) = \sum_y f(x, y) \qquad x \in R_1$$

the summation is taken over all possible y values for each given x in the space R_1.

The marginal p.d.f of Y is given by

$$f_2(y) = \sum_x f(x, y) \qquad x \in R_2$$

where the summation is taken over all possible x values for each given y in the space R_2.

The random variables X and Y are independent if and only if

$$f(x, y) \equiv f_1(x) f_2(y) \qquad x \in R_1, y \in R_2$$

otherwise X and Y are said to be dependent.

Example 1: Let the joint p.d.f of X and Y be defined by

$$f(x, y) = \frac{x+y}{21}, \quad x = 1, 2, 3, \quad y = 1, 2$$

Obtain the marginal p.d.f.'s of X and Y.

Solution:

$$f_1(x) = \sum_y f(x, y) = \sum_{y=1}^{2} \frac{x+y}{21}$$

$$= \quad \frac{x+1}{21} + \frac{x+2}{21} = \frac{2x+3}{21}, \quad x \;=\; 1, 2, 3;$$

$$f_2(y) = \sum_x f(x, y) = \sum_{x=1}^{3} \frac{x+y}{21} = \frac{6+3y}{21}, \quad \text{y} \;=\; 1, 2.$$

Definition 4: Let $f(x, y)$ be the p.d.f. of X and Y then the respective marginal p.d.f's of continuous-type random variables X and Y are given by

$$f(x) = \int_{-\infty}^{\infty} f(x, y)dy, \; x \in R_1$$

$$f(y) = \int_{-\infty}^{\infty} f(x, y)dx, \; y \in R_2$$

where R_1 and R_2 are the spaces of X and Y.

Example 2: Let $f(x, y) = 2\lambda^{-x-y}$, 0 $\leq x \leq \infty$ be the joint p.d.f of X and Y. Find f(x) and f(y), the marginal p.d.f's of X and Y respectively.

Solution:

$$f(x) = \int_x^{\infty} f(x, y)dy = \int_x^{\infty} 2e^{-x-y}dy = \left.\frac{2e^{-x-y}}{-1}\right|_x^{\infty} = 2e^{-x-y} = 2e^{-2x} x < \text{y} < \infty$$

$$f(y) = \int_0^{y} f(x, y)dx = \int_0^{y} 2e^{-x-y}dx = \left.\frac{2e^{-x-y}}{-1}\right|_0^{y} - 2e^{-2y} + 2e^{-y}$$

$$= \quad 2e^{-y} - 2e^{-2y}$$

$$= \quad 2e^{-y}(1-e^2) \qquad 0 < \text{y} < \infty$$

Example 3: Let f(x, y) = ¼, 0 $\leq x \leq$ 2, 0 $\leq$ y $\leq$ 2 be the joint p.d.f of X and Y. Find;

i. f(x)
ii. f(y); the marginal probability density functions.
iii. Are the two random variables independent?

Solution:

i. $f(x)=\int_0^2 f(x,y)dy=\int_0^2 \frac{1}{4}dy=\frac{1}{4}y\Big|_0^2=\frac{2}{4}=\frac{1}{2} \qquad 0\le x\le 2$

ii. $f(y)=\int_0^2 f(x,y)dx=\int_0^2 \frac{1}{4}dx=\frac{1}{4}x\Big|_0^2=\frac{2}{4}=\frac{1}{2} \qquad 0\le y\le 2$

iii. $f(x)f(y)=\frac{1}{2} \text{ x } \frac{1}{2}=\frac{1}{4}$, Yes, the two random variables are independent.

Summary

In this study session, you have learned the following:

1. We discussed that in many situations, we may be interested in observing two characteristics simultaneously.
2. How to solve problems involving two-dimensional random variables.

The concepts of joint and marginal probability density functions.

Self-Assessment Questions (SAQs) for study session 12

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 12.1 (Testing Learning Outcomes 12.1)

Let the joint p.d.f of X and Y be defined by

$$f(x,y)=\frac{x+y}{32}, \qquad x = 1, 2, \quad y = 1, 2, 3, 4$$

Find

a. $f(x)$, the marginal p.d.f of X

b. $f(y)$, the marginal p.d.f of Y

c. P(X > Y)

d. P(Y = 2X)

e. P(X + Y = 3)

f. P(X ≤ 3 – Y)

g. Are X and Y independent or dependent?

SAQ 12.2 (Testing Learning Outcomes 12.2)

Let $f(x, y) = \lambda^{-x-y}$, $0 < x < \infty, 0 < y < \infty$, de the joint p.d.f. of X and Y. Argue that X and Y are independent and compute

i. $P(X < Y)$

ii. $P(X = Y)$

iii. $P(X < 2)$

iv. $P(0 < X < \infty, X/3 < Y < 3X)$

v. $P(0 < X < \infty, 3X < Y < \infty)$

References

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995).*Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993).*Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Study Session 13: Conditional Distributions and Mathematical Expectations

Expected duration: 1 week or 2 contact hours

Introduction

You have discussed how expected values can be used to summarize or describe various aspects of one-dimensional probability distributions. These concepts can be extended to the case of two-dimensional variables. Just as we did in study session four, in this study session, we will discuss measures that describe the "middle" or the "spread" of probability distribution involving two random variables.

Learning Outcomes from Study Session 13

At the end of this study session, you should be able to:

13.1 Compute Conditional Distributions and Mathematical Expectations

13.1 Conditional Distributions and Mathematical Expectations

Let X and Y have a joint discrete distribution with p.d.f $f(x, y)$ on space R. Also let $f_1(x)$ and $f_2(y)$ be the marginal probability density functions with spaces R_1 and R_2 respectively. Let event $A = \{X = x\}$ and event $B = \{Y = y\}$, $(x, y \in R)$. Thus $A \cap B = \{X = x, y = y\}$. Since

$$P(A \cap B) = P(X = x, Y = y) = f(x, y)$$

and

$$P(B) = P(Y = y) = f_2(y) > 0 \text{ (since } y \in R_2\text{)},$$

The conditional probability of event A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{f(x, y)}{f_2(y)}.$$

The following definition is then apparent.

Definition 1: The conditional probability density function of X, given that $Y = y$, is defined by

$$g(x|y)=\frac{f(x,y)}{f_2(y)}, \text{provided that } f_2(y)>0$$

Similarly, the conditional probability density function of Y, given $X = x,$ is given by

$$h(y|x)=\frac{f(x,y)}{f_1(x)}, \text{provided that } f_1(x)>0$$

Moreover, since $h(y|x)\geq 0$, if we sum $h(y|x)$ over y for fixed x, we obtain

$$\sum_y h(y|x)=\sum_y \frac{f(x,y)}{f_1(x)}=\frac{f_1(x)}{f_1(x)}=1$$

Thus $h(y|x)$ satisfies the conditions of a probability density function, so the conditional probability can be computed as

$$P(a<Y<b|X=x)=\sum_{\{y:a<y<b\}} h(y|x)$$

and the corresponding conditional expectations are given as

$$E[u(Y)|X=x]=\sum_y u(y)h(y|x)$$

Written compactly, the conditional mean and conditional variance of Y given $X=x$ are given respectively as

$$\mu_{Y|x}=E(Y|x)=\sum_y yh(y|x),$$

and

$$\sigma^2_{Y|x}=E\{[Y-E(Y|x)]^2|x\}=\sum_y [Y-E(Y|x)]^2 h(y|x),$$

which is alternatively written as

$$\sigma^2_{Y|x}=E(Y^2|x)-[E(Y|x)]^2.$$

Similar expressions can be used for conditional mean and conditional variance for X given $Y = y$.

The above definitions also hold for continuous random variables. For continuous random variables, X and Y, with joint p.d.f. $f(x,y)$ and marginal p.d.f's $f_1(x)$ and $f_2(y)$, respectively. The conditional p.d.f., mean and variance of Y, given $X = x,$ are, respectively,

$$h(y|x) = \frac{f(x, y)}{f_1(x)}, \text{ provided } f_1(x) > 0$$

$$E(Y|x) = \int_{-\infty}^{\infty} yh(y|x)dy,$$

and

$$Var(Y|x) = E\{[Y - E(Y|x)]^2 |x\}$$

$$= \int_{-\infty}^{\infty} [y - E(Y|x)]^2 h(y|x)dy$$

$$= E(Y^2|x) - [E(Y|x)]^2.$$

Expressions for conditional distribution of X, given $Y = y$ are similar.

Example 1: Let X and Y have the joint p.d.f.

$$f(x, y) = \frac{x + y}{32}, \qquad x = 1, 2, \quad y = 1, 2, 3, 4$$

Find

i. $g(x|y)$

ii. $h(y|x)$

iii. $P(1 \le Y \le 3|X = 1)$,

iv. $P(Y \le 2|X = 2)$

v. $P(X = 2|Y = 3)$

vi. $E(Y|X = 1)$

vii. $Var(Y|X = 1)$.

Solution

i. $g(x|y) = \dfrac{f(x, y)}{f(y)} = \dfrac{(x + y)/32}{(2y + 3)/32} = \dfrac{x + y}{2y + 3}, x = 1,2$

ii. $h(y|x) = \dfrac{f(x, y)}{f(x)} = \dfrac{(x + y)/32}{(4x + 10)/32} = \dfrac{x + y}{4x + 10}, y = 1,2,3,4$

iii. $P(1 \le Y \le 3|X = 1) = \displaystyle\sum_{y=1}^{3} \left(\frac{1 + y}{14}\right) = \frac{1}{14}(2 + 3 + 4) = \frac{9}{14}$

iv. $P(Y \le 2|X = 2) = \displaystyle\sum_{1}^{2} \frac{2 + y}{18} = \frac{3}{18} + \frac{4}{18} = \frac{7}{18}$

v. $P(X=2|Y=3)=\frac{5}{9}$

vi. $E(Y|X=1)=\sum_{1}^{4} yh(y|x=1)=\sum_{1}^{4} y\left(\frac{1+y}{14}\right)$

$$=\frac{1}{14}[1(2)+2(3)+3(4)+4(5)]=\frac{40}{14}=\frac{20}{7}$$

vii. $Var(Y|X=1)=E(Y^2|X=1)-[E(Y|X=1)]^2=\sum_{1}^{4} y^2\left(\frac{1+y}{14}\right)-\left(\frac{20}{7}\right)^2$

$$=\frac{130}{14}+\frac{400}{49}=\frac{55}{49}$$

Definition 2: Let $X_1, X_2,..., X_n$, be random variables of the discrete type having a joint distribution. If $U(X_1, X_2,..., X_n)$ is a function of n random variables of the discrete type that have a joint pdf. $f(x_1, x_2,..., x_n)$ and space R, then

$$E[U(X_1, X_2,..., X_n)]=\sum_{(x_1,...,x_n)}\cdots\sum U(x_1, x_2,..., x_n) f(x_1, x_2,..., x_n)$$

If it exists, is called the mathematical expectation of $U(X_1, X_2,..., X_n)$

a. If $U_i(X_1, X_2,..., X_n)=X_i$, then

$$E[U_i(X_1, X_2,..., X_n)]=E(X_i)=\mu_i$$

is called the mean of $X_i, i=1,2,..., n$

b. If $U_2(X_1, X_2,..., X_n)=(X_i-\mu_i)^2$, then

$$E[U_2(X_1, X_2,..., X_n)]=E[(X_i-\mu_i)^2]=\sigma_i^2$$

is called the variance of $X_i, i=1,2,..., n$

Example 2: Let the joint p.d.f of X_1 and X_2 be defined as

$$f(x_1, x_2)=\frac{3-x_1-x_2}{8},\ x_1=0,1 \text{ and } x_2=0,1$$

Find the $E(X_1+X_2)$

Solution:

$$E(X_1+X_2)=\sum_{x_2=0}^{1}\sum_{x_1=0}^{1}(x_1+x_2)\frac{3-x_1-x_2}{8}$$

$$= \quad 0\left(\frac{3}{8}\right)+1\left(\frac{2}{8}\right)+1\left(\frac{2}{8}\right)+2\left(\frac{1}{8}\right)=\frac{6}{8}=\frac{3}{4}$$

Example 3: Let X and Y have the joint p.d.f

$$f(x,y) = 2 \qquad 0 \le x \le y \le 1$$

Obtain

i. Marginal p.d.f's of X and Y.

ii. E(X), E(Y) and E(Y^2)

Solution:

i. $f(x) = \int_x^1 f(x,y)dy = \int_x^1 2dy = 2y\Big|_x^1 = 2(1-x)$

$$f(y) = \int_0^y f(x,y)dy = \int_0^y 2dx = 2y\Big|_0^y = 2y$$

ii. $E(X) = \int_0^1 \int_x^1 2xdydx = \int_0^1 2x(1-x)dx = \frac{2x^2}{2} - \frac{2x^3}{3}\Big|_0^1 = 1 - \frac{2}{3} = \frac{1}{3}$

$$E(Y) = \int_0^1 \int_0^y 2ydxdy = \int_0^1 2y^2dy = \frac{2y^3}{3}\Big|_0^1 = \frac{2}{3}$$

$$E(Y^2) = \int_0^1 \int_0^y 2y^2dxdy = \int_0^1 2y^3dy = \frac{2y^4}{4}\Big|_0^1 = \frac{1}{2}$$

Summary

In this study session you have learnt about:

1. The concepts of conditional probability density function of a random variable, given another random variable.
2. The methods used in obtaining the conditional mean and conditional variance.We have points given above are considered for both discrete and continuous random variables.

Self-Assessment Questions (SAQs) for study session 13

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 13.1 (Testing Learning Outcomes 13.1)

Let $f_1(x)=\frac{1}{10}, x=0,1,2,\ldots,9,$ and $h(y|x)=\frac{1}{(10-x)}, y=x,x+1,\ldots,9.$ Find

i. $f(x,y)$

ii. $f_2(y)$

iii. $E(Y|x)$

SAQ 13.2 (Testing Learning Outcomes 13.2)

Let $f(x,y)=1/8,\ 0\le y\le 4,\ 0\le x\le y+2$, be the joint p.d.f. of X and Y.

iv. Find $f_1(x)$, the marginal p.d.f. of X.

v. Determine $h(y|x)$, the conditional p.d.f. of Y, given $X=x$.

vi. Compute $E(Y|x)$, the conditional mean of Y, given $X=x$.

vii. Evaluate $Var(Y|x)$.

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995). *Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Study Session 14:Transformation of Random Variables

Expected duration: 1 week or 2 contact hours

Introduction

In this study session, another important method of constructing models shall be discussed. To understand the depth of theory and application of statistics, transformations of random variables must be taught.

An example of such transformations is that of the normal distribution that can be transformed into a standard normal distribution. We will first consider transformations of variables in one dimension in this current lecture. The joint transformations shall be discussed in the next lecture.

Let $u(x)$ be a real-valued function of a real variable x. If the equation $y = u(x)$ can be uniquely solved, say $x = w(y)$, then we have a one-to-one transformation. Discrete and continuous cases shall be considered separately.

Learning Outcomes for Study Session 14

At the end of this study session, you should be able to:

14.1 Transformation of Variables of the Discrete Type
14.2 Transformations of Variables of the Continuous Type

14.1 Transformation of Variables of the Discrete Type

An alternative method of finding the distribution of a function of one or more random variables is called the change-of-variable technique.

Let X have a Poisson p.d.f.

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

$$= \quad 0 \quad \text{elsewhere}$$

Let A denote the space A = $\{x_i : x = 0, 1, 2, ---\}$ so that f(x) > 0.

Define a new random variable Y by Y = 4X.

Suppose we wish to find the p.d.f of Y by the change-of-variable technique.

Let y = 4x, we call y = 4x a transformation from x to y, and we say that the transformation maps the space A on to the space B = {y : y = 0, 4, 8, 12, ----}. The space B is obtained by transforming each point in A in accordance with y = 4x. (i.e every point in A corresponds to one and only one point in B and vice-versa).

Therefore any function y = U(x) that maps a space A onto a space B such that there is a one-to-one correspondence between the points of A and those of B is called one-to-one transformation.

$$y = 4x \Rightarrow x = \frac{1}{4}y$$

The problem is to find g(y) of the discrete type of the random variable Y = 4X.

Now $g(y) = P(Y = y) = \Pr\left(X = \frac{y}{4}\right) = \frac{\lambda^{y/4} e^{-\lambda}}{\left(y/4\right)!}$, y = 0, 4, 8, 12, ----

= 0, elsewhere

Example 1: Let X have the binomial p.d.f

$$f(x) = \frac{3!}{x!(3-x)!}\left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{3-x}, \quad x = 0, 1, 2, 3, ---$$

= 0 elsewhere

Find the p.d.f of Y = X^2.

Solution:

y = U(x) = x^2 maps A = {x : x = 0, 1, 2, 3} onto

B = {y: y = 0, 1, 4, 9,}

$$g(y)\frac{3!}{\sqrt{y}!(3-\sqrt{y})!}\left(\frac{2}{3}\right)^{\sqrt{y}}\left(\frac{1}{3}\right)^{3-\sqrt{y}} \qquad y = 0, 1, 4, 9$$

Example 2: Let $X \sim GEO(P)$, so that

$f(x) = pq^{x-1}$ x = 1, 2, 3 ----

Find p.d.f. of Y = X – 1,

Solution:

y = U(x) = x – 1, and x = y + 1

g(y) = P(Y = y) = P(x = y + 1) = Pq^{y+1-1}

= Pq^y y = 0, 1, 2 ---

Example 3: Let X have a p.d.f. f(x) = $^1/_3$, x = 1, 2, 3

= 0 elsewhere

Find the p.d.f. of Y = 2X + 1

Solution: $x = {(y-1)}/{2}$

$$g(y) = P\left(X = {(y-1)}/{2}\right) = \frac{1}{3}, \quad y = 3, 5, 7$$

Example 4: Let X have the p.d.f. $f(x) = \left(\frac{1}{2}\right)^2$, x = 1, 2, 3 ----

= 0, elsewhere

Find the p.d.f of Y = X^3.

Solution:

Y = U(x) = x^3. Since $y = x^3 \Rightarrow x = 3\sqrt{y}$

$$g(y) = \left(\frac{1}{2}\right)^{3\sqrt{y}}, \quad y = 1, 8, 27 \text{ ----}$$

14.2 Transformations of Variables of the Continuous Type

Let X be a random variable of the continuous type having p.d.f f(x). Let A be the one-dimensional space where f(x) > 0. Consider the random variable Y = U(X), where y = U(x) defines a one-to-one transformation that maps the set A onto the set B.

Let the inverse of y = U(x) be denoted by x = w(y), and let the derivative $\frac{dx}{dy} = w'(y)$ be continuous and not equal to zero for all points y in B.

Then the p.d.f of the random variable Y = U(X) is given by

$$g(y) = f[w(y)]|w'(y)|, \quad y \in B$$

= 0 elsewhere

Where $\frac{dx}{dy} = w'(y)$ is the Jacobian (denoted by J) of the transformation.

Example 1: Let X be a random variable of the continuous type having p.d.f

f(x) = 2x 0 < x < 1

= 0

Define the random variable Y by Y = $8X^3$

Solution: g(y) = f(w(y))|J|

y = $8x^3$ is a transformation from x to y. $\Rightarrow x = \sqrt[3]{\frac{y}{8}} = \frac{1}{2}\sqrt[3]{y}$

$= \frac{1}{2} y^{1/3}$

$$|J| = \frac{dx}{dy} = \frac{1}{6} y^{-2/3}$$

$g(y) = y^{1/3}$ x $\frac{1}{6} y^{-2/3} = \frac{1}{6} y^{-1/3} = \frac{1}{6y^{1/3}}, 0 < y < 8$ i.e $0 < \frac{1}{2} y^{1/3} < 1$

$0 < y^{1/3} < 2$

0 < y < 8

Example 2: Let X have the p.d.f

f(x) = 1, 0 < x < 1

= 0 elsewhere

Find the p.d.f of Y = -2 ln X.

Solution: The transformation y = U(x) = - 2 ln x, so that $x = w(y) = e^{-y/2}$

$$J = \frac{dx}{dy} = w(y) = -\frac{1}{2} e^{-y/2}$$

The p.d.f. of Y is

$$g(y) = |J| f(w(y)) = \frac{1}{2} e^{-y/2}.1 = \frac{1}{2} e^{-y/2}, 0 < y < \infty$$

Example 3: Let the p.d.f of X be defined by $f(x) = x^3/4$, 0 < x < 2. Find the p.d.f of Y = X^2

Solution: The transformation y = U(x) = x^2, so that $x = \sqrt{y}$ or $y^{1/2} = w(y)$

$$g(y) = f(w(y))|J|$$

$J = \frac{dx}{dy} = 1/2\, y^{-1/2}$; $f(w(y)) = f(y^{1/2}) = \frac{y^{3/2}}{4}$

$$g(y) = \frac{y^{3/2}}{4} \text{ x } \frac{y^{-1/2}}{2}$$

$= y/8$, 0 < y < 4 since, 0 < x < 2

$0 < y^{1/2} < 2$

0 < y < 4

Example 4: Let the p.d.f of X be defined by $f(x) = \left(\frac{3}{2}\right)x^2$, $-1 < x < 1$. Find the p.d.f. of $Y = \frac{(X^3 + 1)}{2}$

Solution: The transformation $y = U(x) = \frac{(x^3 + 1)}{2}$, so that $x^3 = 2y - 1$

$$x = \sqrt[3]{2y-1} \quad or \quad (2y-1)^{1/3}$$

$$J = \frac{dx}{dy} = \frac{1}{3}(2y-1)^{-2/3} \text{ x } 2$$

$$= \quad \frac{2}{3}(2y-1)^{-2/3}$$

$$f(w(y)) = \frac{3}{2}(2y-1)^{2/3}$$

$$g(y) = f(w(y))|J| = \frac{2}{3} \text{ x } \frac{3}{2}(2y-1)^{-2/3+2/3} = 1$$

That is $-1 < x < 1 \Rightarrow -1 < (2y-1)^{1/3} < 1$

$-1 < 2y - 1 < 1$

$0 < 2y < 2$

$0 < y < 1$

Summary

In this study session you have learnt about

1. Finding the p.d.f. of a random variable from a p.d.f. of another random variable using the change-of-variable technique.
2. How to handle cases concerning either discrete or continuous random variable.

Self-Assessment Questions (SAQs) for study session 14

Now that you have completed this study session, you can assess how well you have achieved its Learning outcomes by answering the following questions. Write your answers in your study Diary and discuss them with your Tutor at the next study Support Meeting. You can check your Define School answers with the Notes on the Self-Assessment questions at the end of this Module.

SAQ 14.1 (Testing Learning Outcomes 14.1)

Let the p.d.f. of X be defined by $f(x)=\left(\frac{1}{2}\right)^x, x=1,2,3,...,$ zero elsewhere. Find the p.d.f. of $Y = X^3$.

SAQ 14.2 (Testing Learning Outcomes 14.2)

If the random variable X is distributed as $N(\mu,\sigma^2)$, show, by means of a transformation, that the random variable $Y = [(X-\mu)/\sigma^2]^2$ is distributed as χ_1^2.

SAQ 14.3 (Testing Learning Outcomes 14.3)

If Y has a uniform distribution on the interval (0, 1), find the p.d.f. of $X = (2Y-1)^{\frac{1}{3}}$.

SAQ 14.4 (Testing Learning Outcomes 14.4)

Let X have p.d.f. $f(x)=x^2/24, -2<x<4$ and zero elsewhere. Find the p.d.f. of $Y = X^2$.

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995). *Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Study Session 15: Joint Transformations

Expected duration: 1 week or 2 contact hours

Introduction

A one-dimensional case that was considered in the last lecture can be extended to a multi-dimensional case with appropriate modifications. Joint transformations of continuous random variables can be accomplished in a similar manner to those considered in our last lecture, but the notion of Jacobian must be generalized.

Learning Outcomes From Study Session 15

At the end of this study session, you should be able to:

15.1 Explain Joint Transformation

15.1 Joint Transformations

This method of finding the p.d.f of a function of one random variable of the continuous type will be extended to functions of two random variables of this type. Let $y_1 – U_1(x_1, x_2)$ and $y_2 – U_2(x_1, x_2)$ define a one-to-one transformation that maps a (two-dimensional) set A in the x_1, x_2 – plane unto a (two-dimensional) set B in the y_1y_2 – plane. If we express each of x_1 and x_2 in terms of y_1 and y_2 we can write $x_1 = w_1(y_1 y_2)$. The determinant is of order 2

$$\begin{vmatrix} \frac{dx_1}{dy_1} & \frac{dx_1}{dy_2} \\ \frac{dx_2}{dy_1} & \frac{dx_2}{dy_2} \end{vmatrix}$$ is called the Jacobian of the transformation, denoted by J.

Example:

Let X_1, X_2 denote a random sample from a distribution $\chi^2_{(2)}$. (note that p.d.f. of X is $f(x) = \frac{1}{\Gamma\left(\frac{r}{2}\right)2^{r/2}} X^{r/2-1} e^{-x/2} \qquad 0 \le x < \infty$, then X has a Chi-square distribution with r degrees of freedom abbreviated by saying X is χ^2_r). .

Find

i. The joint p.d.f of $Y_1 = X_1$ and $Y_2 = X_2 + X_1$, for $0 < y_1 < y_2 < \infty$.

ii. The marginal p.d.f of each of Y_1 and Y_2.

iii. Are Y_1 and Y_2 independent?

Solution:

i. $f(x_1) = \frac{1}{2} e^{-x_1/2}$

$f(x_2) = \frac{1}{2} e^{-x_2/2}$

The joint p.d.f of X_1 and X_2 is

$$f(x_1) f(x_2) = \frac{1}{4} e^{-\frac{x_1 + x_2}{2}}$$

The transformation $y_1 = U_1(x_1, x_2) = x_1$ and $y_2 = U_2(x_1, x_2) = x_2 + x_1$

So that $x_1 = w_1(y_1, y_2) = y_1$ and

$x_2 = w_2(y_1, y_2) = y_2 - y_1$

$g(y_1, y_2) = f[w_1(y_1, y_2)] f[w_2(y_1, y_2)] |J|$

$|J| = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$

The joint p.d.f of Y_1 and Y_2 is

$$g(y_1, y_2) = \frac{1}{4} e^{-\frac{(y_1 + y_2 - y_1)}{2}} = \frac{1}{4} e^{-y_2/2}$$

ii. The marginal p.d.fs are

$g_1(y_1) = \int g(y_1, y_2) dy_2 \qquad 0 < y_1 < y_2 < \infty$

$= \int_{y_1}^{\infty} \frac{1}{4} e^{-y_2/2} dy_2 \qquad y_1 < y_2 < \infty$

$= \frac{1}{4} \left[\frac{e^{-y_2/2}}{-\frac{1}{2}} \right]_{y_1}^{\infty}$

$= -\frac{1}{2} \left[-e^{-y_1/2} \right]$

$= \frac{1}{2} e^{-y_1/2}$

$g_2(y_2) = \int_0^{y_2} \frac{1}{4} e^{-y_2/2} dy_1 \qquad 0 < y_1 < y_2$

$= \frac{1}{4} \left[y_1 e^{-y_2/2} \right]_0^{y_2}$

$= \frac{1}{4} y_2 e^{-y_2/2}$

iii.No, since $\frac{1}{2}e^{-y_1/2}$ x $\frac{1}{4}y_2e^{-y_2/2} = \frac{1}{8}y_2e^{-(y_1+y_2)/2}$

Summary

This study session has discussed the followings:

1. Transformations of variables involving jointly distributed random variables.
2. The last lecture dealt with the notion of a one-to-one transformation and the mapping of one set to another set under that transformation; however, we built on these ideas in the current study session to help us find the distribution of a function of two variables of the discrete type.
3. We also looked at the same problem in raised in 2 above when the random variables are of the continuous type.
4. Examples of these transformations were given for clearer understanding of the concept.

Computation of marginal p.d.f.s and independent variables revisited.

SAQ 15.1 (Testing Learning Outcomes 15.1)

If X_1 and X_2 denote a random sample of size two from a Poisson distribution, $X_i \sim POI(\lambda)$, find the p.d.f. of $Y = X_1 + X_2$.

SAQ 15.2 (Testing Learning Outcomes 15.2)

Suppose that X_1 and X_2 denote a random sample of size two from a gamma distribution, $X_i \sim GAM(2,1/2)$.

i. Find the p.d.f. of $Y = \sqrt{X_1 + X_2}$

ii. Find the p.d.f. of $W = X_1/X_2$

SAQ 15.3 (Testing Learning Outcomes 15.3)

Let X_1 and X_2 denote a random sample of size two from a distribution that is $N(\mu,\sigma^2)$. Let $Y_1 = X_1 + X_2$ and $Y_2 = X_1 - X_2$. Find the joint p.d.f. of Y_1 and Y_2 and show that these random variables are independent.

SAQ 15.4 (Testing Learning Outcomes 15.4)

Let X_1 and X_2 have the joint p.d.f. $h(x_1,x_2) = 8x_1x_2$, $0 < x_1 < x_2 < 1$, zero elsewhere. Find the joint p.d.f. of $Y_1 = X_1/X_2$ and $Y_2 = X_2$ and argue that Y_1 and Y_2 are independent. (Hint: Use the inequalities $0 < y_1y_2 < y_2 < 1$ in considering the mapping from A to B).

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995). *Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference.*New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Study Session 16: The t and F Distributions

Expected duration: 1 week or 2 contact hours

Introduction

In this study session, we will consider two probability distributions that are used considerably in certain problems of statistical inference; the t distribution and the F distribution. Both of these probability distributions occur for certain functions of normal random variables, as will be seen shortly.

The t distribution is first discovered by W. S. Gosset when he was working for an Irish brewery. The distribution is often known as Student's t distribution. However, the F distribution is first proposed by George Snedecor to honor R. A. Fisher, who used a modification of this ratio in several statistical applications.

Learning Outcomes from Study Session 16

At the end of this Study Session, you should be able to:

16.1 Explain the t and F Distributions

16.1 The t and F Distributions

The t-distribution:

The t distribution with n degrees of freedom can be defined as that of a random variable symmetrically distributed about zero whose square has the F distribution with 1 and n degrees of freedom in the numerator and denominator, respectively. Let W denote a random variable that is N (0, 1); let V denote a random variable that is $\chi^2(r)$; and let W and V be independent.

Then

$$T = \frac{W}{\sqrt{V/r}}$$

has a t-distribution with r degrees of freedom. Its p.d.f is

$$g(t) = \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\,\Gamma(r/2)\left(1+t^2/r\right)^{(r+1)/2}}, \quad -\infty < t < \infty$$

Proof Since W and V are independent, the joint p.d.f of W and V, say h (w, v) is the product of the p.d.fs of W and V,

$$h(w,v) = \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} v^{r/2-1} e^{-V/2}$$

$$= \quad 0 \quad \text{elsewhere.} \qquad \begin{matrix} -\infty < w < \infty \\ 0 < v < \infty \end{matrix}$$

The charge-of-variable technique is used to obtain the p.d.f of g(t) of T.

$$t = \frac{w}{\sqrt{v/r}} \quad \text{and} \quad u = v$$

define a one-to-one transformation that maps A = {(w, v) : - ∞< w <∞, 0 < v <∞} onto B = {(t, u) : - ∞< t <∞, 0 < u <∞}.

Since $t = \frac{w}{\sqrt{v/r}} \Rightarrow w = t\sqrt{u}/\sqrt{r}$, v = u

$$|J| = \left|\frac{dw}{dt}\right| = \sqrt{u}/\sqrt{r}$$

The joint p.d.f of T and U is given by

$$g(t,u) = h\left(t\sqrt{u}/\sqrt{r}, u\right)|J|$$

$$= \frac{1}{\sqrt{2\pi}\Gamma(r/2)2^{r/2}} u^{r/2-1} \exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right]\frac{\sqrt{u}}{\sqrt{r}}$$

$$= \quad 0 \quad \text{elsewhere} \qquad -\infty < t < \infty, \quad 0 < u < \infty$$

The marginal p.d.f. of T is

$$g(t) = \int_{-\infty}^{\infty} g(t,u)du$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{(r+1)/2-1} \exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right]du$$

Let $z = u\left[1+\left(t^2/r\right)\right]/2$ then

$$g(t) = \int^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} \left(\frac{2z}{1+t^2/r}\right)^{\frac{(r+1)}{2}-1} e^{-Z\left(\frac{2}{1+t^2/r}\right)dZ}$$

$$= \frac{\Gamma\left[(r+1)/2\right]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{\left(1+t^2/r\right)^{(r+1)/2}} \qquad -\infty < t < \infty$$

Thus if W is N(0, 1), if V is $\chi^2(r)$, and W and V are independent, then

$$T = \frac{W}{\sqrt{V/r}}$$

Properties

1. The t distribution is symmetric about zero ($E(T)=0$ when $r \geq 2$) and its general shape is similar to that of the standard normal distribution.
2. The t distribution approaches the standard normal distribution as $r \to \infty$, for smaller v the t distribution is flatter with thicker tails and, in fact, $T \sim CAU(1,0)$ when $r = 1$ (the t distribution is the Cauchy distribution and the mean and thus the variance do not exist for the Cauchy distribution).
3. The t distribution has more variability than the standard normal distribution since more variation is noticed when Z is divided by another random variable.

16.2 The F-distribution

In the problem of comparing normal populations with respect to their variances, as well as in a variety of other problems, it will be necessary to know the distribution of the ratio of two chi-square random variables. Consider two independent Chi-square random variables U and V having r_1 and r_2 degrees of freedom respectively; then

$$F = \frac{U/r_1}{V/r_2}$$

has an F-distribution with r_1 and r_2 degrees of freedom. Its p.d.f is

$$h(u,v) = \frac{\Gamma[(r_1+r_2)/2](r_1/r_2)^{r_1/2} W^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1+r_1 w/r_2)^{(r_1+r_2)/2}}, \qquad 0 < w < \infty$$

Proof: The joint p.d.f of U and V is

$$h(u,v) = \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)2^{(r_1+r_2)}} U^{r_1/2-1} V^{r_2/2-1} e^{-(u+v)/2}$$

$$= \quad 0 \quad \text{elsewhere} \qquad 0 < u < \infty; \; 0 < v < \infty$$

To find the p.d.f g(w) of W, the equations

$$w = \frac{u/r_1}{v/r_2}$$ and

z = v

define a one-to-one transformation that maps the set

A = {(u, v) : 0 < u <∞, 0 < v <∞} onto the set B = {(w, z) : 0 < w <∞, 0 < z <∞}, Since $u = \left(r_1 / r_2 \right) zw$, v = z

$|J| = \left(r_1 / r_2 \right) z$. The joint p.d.f g(w, z) of the random variables W and Z=V is

$$g(w,z) = \frac{1}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right) 2^{(r_1+r_2)/2}} \left(\frac{r_1 zw}{r_2} \right)^{\frac{r_1}{2}-1} z^{\frac{r_2}{2}-1} \times \exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2} + 1 \right) \right] \frac{r_1 z}{r_2},$$ provided that (w, z) ∈ β and zero elsewhere,

The marginal p.d.f g(w) of W is

$$g(w) = \int_{-\infty}^{\infty} g(w,z)\,dz$$

$$= \int_0^{\infty} \frac{\left(r_1 / r_2 \right)^{r_1/2} (w)^{\frac{r_1}{2}-1}}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right) 2^{(r_1+r_2)/2}} z^{\frac{(r_1+r_2)}{2}-1} \times \exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2} + 1 \right) \right] dz$$

If we change the variable of integration by writing

$$y = \frac{z}{2}\left(\frac{r_1 w}{r_2} + 1 \right)$$

it can be seen that

$$g_1(w) = \int_0^{\infty} \frac{\left(r_1 / r_2 \right)^{r_1/2} (w)^{\frac{r_1}{2}-1}}{\Gamma\left(r_1/2 \right)\Gamma\left(r_2/2 \right) 2^{(r_1+r_2)/2}} \left(\frac{2_y}{r_1 w / r_2 + 1} \right)^{\frac{(r_1+r_2)}{2}-1} e^{-y} \text{ x } \left(\frac{2}{r_1 w / r_2 + 1} \right) dy$$

$$= \frac{\Gamma\left[(r_1+r_2)/2 \right] \left(r_1 / r_2 \right)^{\frac{r_1}{2}}}{\Gamma\left(r_1/2 \right)\Gamma\left(r_2/2 \right)} \frac{(w)^{\frac{r_1}{2}-1}}{\left(1 + r_1 w / r_2 \right)^{(r_1+r_2)/2}} \qquad 0 < w < \infty$$

∴ If U and V are independent Chi-square with r_1 and r_2 degrees of freedom respectively, then

$$W = \frac{U/r_1}{V/r_2}$$

has p.d.f g(w). The distribution of this random variable is usually called an F distribution.

If $X \sim F(r_1, r_2)$, then

$$E(X^r) = \frac{\left(\frac{r_2}{r_1}\right)^r \Gamma\left(\frac{r_1}{2}+r\right)\Gamma\left(\frac{r_2}{2}-r\right)}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)}, r_2 > 2r$$

For $r = 1$, the mean, $E(X) = \frac{r_2}{r_2 - 2}, 2 < r_2$

The variance is given by

$$Var(X) = \frac{2r_2^2(r_1 + r_2 - 2)}{r_1(r_2-2)^2(r_2-4)},\ 4 < r_2.$$

Properties

1. If F is distributed as F_{r_1,r_2}, then, $1/F$ is distributed as F_{r_2,r_1}.
2. If X is $N(0,1)$, Y is χ_r^2 and X, Y are independent, so that $T = X/\sqrt{Y/r}$ is distributed as t_r, then T^2 is distributed as $F_{1,r}$, since X^2 is χ_1^2.

Examples

1. Let X have a t distribution with r degrees of freedom. Find
   i. $P(X \ge 2.228)$ when $r = 10$
   ii. $P(X \le 2.228)$ when $r = 10$
   iii. $P(1.330 \le X \le 2.552)$ when $r = 18$
2. Let X have a t distribution with r=19. find c such that
   a. $P(X \ge c) = 0.025$
   b. $P(X \le c) = 0.95$

Solution

1. i. $P(X \ge 2.228) = 1 - 0.975 = 0.025$
   ii. $P(X \le 2.228) = 0.975$

iii $P(1.330 \le X \le 2.552) = 0.99 - 0.90 = 0.09$

2. i. $P(X \ge c) = 0.025$, c = 2.093

   ii. $P(X \le c) = 0.95$, c = 1.729

Summary

In this study session you have learnt about

1. The t distribution is completely determined by the number r, the degrees of freedom.
2. Because of the symmetry of the t distribution about t=0, the mean if it exists, must be equal to zero.
3. The F distribution depends on two parameters, r_1 and r_2 in that order. The first parameter is the number of degrees of freedom in the numerator, and the second is the number of degrees of freedom in the denominator.
4. The t distribution is very useful in testing hypothesis about the mean.

The F distribution on the other hand, is used to test hypothesis about ratio of two variances or sum of squares.

SAQ 16.1 (Testing Learning Outcomes 16.1)

Let X have an F distribution with r_1 and r_2 degrees of freedom. Find

i. $P(X \ge 3.02)$ when $r_1 = 9, r_2 = 10$

ii. $P(X \le 4.14)$ when $r_1 = 7, r_2 = 15$

iii. $P(X \le 0.1508)$ when $r_1 = 8, r_2 = 5$. Hint: $0.1508 = 1/6.63$

iv. $P(0.1323 \le X \le 2.79)$ when $r_1 = 6, r_2 = 15$

SAQ 16.2 (Testing Learning Outcomes 16.2)

Let X have an F distribution with r_1 and r_2 degrees of freedom. Find

i. $P(a \le X \le b) = 0.90$ when $r_1 = 8, r_2 = 6$

ii. $P(a \le X \le b) = 0.98$ when $r_1 = 8, r_2 = 6$

iii. Let X have a t distribution with r degrees of freedom. Find

iv. $P(|X| \ge 2.228)$ when $r = 10$

v. $P(-1.753 \le X \le 2.602)$ when $r = 15$

References

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995). *Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993). *Probability and Statistical Inference.* New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Bain L. J and Engelhardt Max (1989). *Introduction to Probability and Mathematical Statistics*. Boston: PWS-KENT Publishing Company.

Hogg R. V and Craig A. T (1970). *Introduction to Mathematical Statistics*. 3rd Edition. New York: Macmillan Publishing Co., Inc. London: Collier Macmillan Publishers.

Hogg R. V and Craig A. T (1995)**.***Introduction to Mathematical Statistics*. 5th Edition. London: Prentice-Hall, Inc.

Hogg R. V and Tanis E. A (1993)**.***Probability and Statistical Inference*. New York: Macmillan Publishing Company.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Lindgren Bernard W (1976). *Statistical Theory* Third Edition. New York: Macmillan Publishing Co., Inc.

Roussas George G. (1972). *A First Course in Mathematical Statistics*. Reading Massachusetts: Addisson-Wesley publishing Company.

Shangodoyin D. K, Olubusoye O. E, Shittu O. I and Adepoju A. A (2002). *Statistical Theory and Methods* Nigeria: Joytal Printing Press, Agbowo, Ibadan. ISBN: 978-2906-23-9.

Spiegel Murray R (1972). *Theory and Problems of Statistics*. SI (metric) Edition. New York: McGraw-Hill Book Company (UK) Limited.

Larson Harold J (1982). *Introduction to Probability Theory and Statistical Inference*. Third Edition. New York: John Wiley & Sons.

Hamburg Morris (1970).*Statistics Analysis for Decision Making*. New York: Harcourt, Brace & World, Inc.

Ross S (1994)**.***A First Course in Probability*. 4th Ed., New Jersey: Prentice-Hall, Inc. A Simon & Schuster Company, Englewood Cliffs.