COURSE MANUAL

# Probability III

## STA 311

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

ISBN 978-021-692-2

*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.

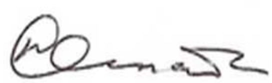
Apart from the responsibility for you to read this course material with others, you are also advised to seek assistance from your course facilitators especially academic advisors during your study even before the interactive session which is by design for revision. Your academic advisors will assist 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 | Alawode, O.A. & Shittu, O.I. |
| 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

## Course Introduction

Probability III is a course designed to build on the concept of Probability that the students have been exposed to from their first two years in the University.

Probability is a very important concept in Statistics as it helps statisticians in drawing valid inferences about their researches or experiments. Topics covered in this course have been carefully treated to ensure that students have a thorough understanding of the subject so that a sound foundation for advanced statistical inference can be established.

The course covers definitions and rules of probability, conditional probability and independence, Bayes' theorem, Urn models, Principle of Inclusion and Exclusion, Allocation and Matching problems, Probability Generating Functions, Bernoulli Trials, Binomial, Poisson, Hypergeometric, Negative Binomial and Geometric distributions, Multinomial Distributions and Poisson process.

## Objectives

The objectives of this course include the following:

1. to exhibit a proficiency in the topics to be covered in this course.
2. to understand the laws of probability
3. to engage in critical thinking and problem solving
4. to offer a sound foundation for statistical inference

# Study Session 1: Definitions and Rules of Probability

## Expected duration: 1 week or 2 contact hours

## Introduction

Since participation in games and gambling is as old as mankind, it then suggest to us that the idea of probability is almost as old. But, the realization that one could predict an outcome to a certain degree of accuracy could not be done until the 16th and 17th centuries. This idea that one could determine the chance of a future event was prompted from the necessity to achieve a predictable balance between risks taken and the potential for gain.

In order to make a profit, underwriters were in need of dependable guidelines by which a profit could be expected, while the gambler was interested in predicting the possibility of gain. In this study session, your focus will be directed to the definitions and rules of probability.

## Learning Outcomes for Study Session 1

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

1.1 Discuss the basics of probability theory stating the types and their definitions
1.2 State the relevant theorems and definitions that are central to the theory of probability

## 1.1 Probability

The theory of probability, in its early stages, was closely linked to the games of chance. Dice are commonly associated with games of chance. For instance, in a dice game, one is interested in only the numbers that rise to the top. If a single 'fair' die is rolled, there are six outcomes and each of them is equally likely.

If the probability of a number greater than 4 is required from a single fair die, then the denominator $D = 6$. Since, there are $N = 2$ salient outcomes, namely 5 and 6, then the probability we want is $\frac{2}{6} = \frac{1}{3}$

Probability can be defined as a measure of uncertainty. It can also be defined as a real value that measures the degree of belief that we have in the occurrence of an event. The following definitions of probability, based on different approaches to the concept, are:

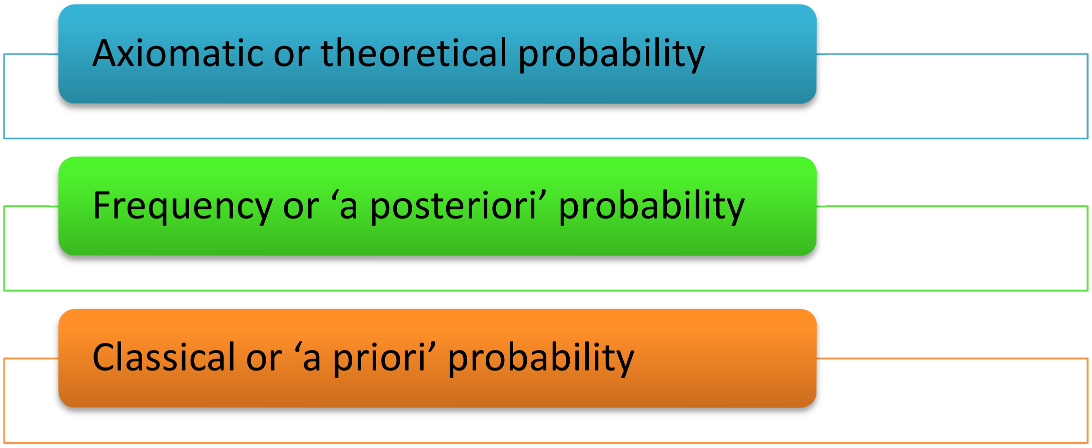



**Figure 1.1**: Different approaches to the concept of Probability

**1 Classical or ‘a priori’ probability:**

If there are $n$ number of exhaustive, mutually exclusive and equally likely cases of an event and suppose that $n_A$ of them are favourable to the happenings of an event $A$ under the given set of conditions, then $(A) = \frac{n_A}{n}$ . An example is the toss of a die once. The six possible outcomes are 1,2,3,4,5,6. The probability of occurrence of a 2 is $\frac{1}{6}$. The probability is ‘a priori’, that is it can be determined before carrying out the experiment.

| Box 1.1: Note |
| --- |
| The word ‘exhaustive’ assures the happening of an event either for or against and disallows the possibility of happening of neither for nor against in any trial. |

**2 Frequency or ‘a posteriori’ probability:**

This probability can only be determined after the experiment has been performed. The experiment is repeated many times, under similar conditions and observations are taken. The probability $p$ is then approximated by the relative frequency with which the repeated observations satisfy the event, i.e. $\lim_{N \to \infty} \frac{n(A)}{N}$

**3 Axiomatic or theoretical probability:**

This is based on available information. The probability of an event A in relation to an experiment with sample space S is a real valued function $P(A)$ which satisfies the following axioms:

$$P(A) \geq 0$$

$$P(S) = 1$$

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

If $A_1,\ A_2,\ A_3 \ldots\ldots$ are elementary events in a sample space S, then the probability of the union $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Example 1: Consider $S = \{HH, HT, TH, TT\}$, then

$$P\left(\bigcup_{i=1}^{4} A_i\right) = [P(HH) + P(HT) + P(HT) + P(TT)]$$

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

### 1.1.1 Probability Axioms

The function $P(.) \colon F \rightarrow [0,1]$ is a probability function if the following axioms are satisfied:

$$P(A) \geq 0$$

$$P(S) = 1$$

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

### In-Text Question

Frequency or ‘a posteriori’ approach believes that probability can only be determined after the experiment has been performed. True/False?

### In-Text Answer

True

## 1.2 Definitions of terms

The following are the definition of some of the relevant terms in probability:

1. **Sample Space ($\boldsymbol{S}$)**: This is the totality of possible outcomes of a random experiment. For instance, in the experiment of tossing a coin, $S = (H, T)$
2. **Equiprobable Sample Space**: This is the sample space to which we assign each of the sample elements equal probability. The sample elements are said to be equally likely or equiprobable.
3. **Event**: This is a subset of a sample space. Suppose $S = (HH, HT, TH, TT)$, the event $E$ of 2 heads occurring in the toss of a coin twice is $E = (HH)$
4. **Sure/Certain Event**: The sample space $S$ is the only sure event. The probability of a certain event E is one $(P(E) = 1)$
5. **Mutually Exclusive Events**: Two or more events are said to be mutually exclusive if the happening or occurrence of any one of them excludes the happening of the others.
6. **Union of $\boldsymbol{n}$ events**: The union of events $A_1, A_2, A_3, \ldots . A_n$ is denoted by $A_1 \cup A_2 \cup A_3 \cup \ldots . \cup A_n$
7. **Intersection/product of two events**: Consider two events $A$ and $B$. The intersection between the two events is usually denoted by $A \cap B$ or $AB$. For $n$ events, we have $A_1 \cap A_2 \cap A_3 \cap \ldots . \cap A_n$
8. **Dependent and independent events**: Two or more events are said to be dependent or independent according as the occurrence of one does or does not affect the occurrence of the other or others.
9. **Compound Events (Joint Occurrence)**: This is the simultaneous occurrence of two or more events.
10. **Impossible Event**: This is the complement of the sure event. It is an empty set $\emptyset$.

### 1.2.1 Rules of Probability

The following are the rules of probability:

**Theorem 1**: Let $S$ be a sample space and $P(.)$ be a probability function on $S$; then the probability that the event $A$ does not happen is $1 - P(A)$ i.e. $P(A') = 1 - P(A)$.

Proof:

From definition, $A \cap A' = \emptyset$

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

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

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

**Theorem 2**: Let $S$ be a sample space with probability function $P(.)$; then $0 \leq P(A) \leq 1$ for any event $A$ in $S$.

**Proof**:

By property (1), $P(A) \geq 0$

We need to show that $P(A) \leq 1$

From theorem (1), $P(A) + P(A') = 1$

But $P(A') \geq 0$

So, $P(A) = 1 - P(A') \leq 1$

**Theorem 3**: Let $S$ be a sample space with a probability function $P(.)$. If $\emptyset$ is the impossible event, then $P(\emptyset) = 0$.

**Proof**:

Observe that $\emptyset = S'$

From property (3), we get $P(S \cup S') = P(S) + P(S')$

$P(S) + P(\emptyset)$

But $S \cup S' = S$ and $P(S) = 1$

Therefore $P(\emptyset) = 0$

**Theorem 4**: If $A_1$ and $A_2$ are subsets of $S$ such that $A_1 \subset A_2$, then $P(A_1) \leq P(A_2)$.

**Theorem 5**: Commutative laws:

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

**Theorem 6**: Associative laws:

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

**Theorem 7**: Distributive laws:

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

**Theorem 8**:

$(A')' = A$

$$A' = S \setminus A$$

**Theorem 9**:

$A \cap S = A$

$$A \cup S = S$$
$$A \cap \emptyset = \emptyset$$
$$A \cup \emptyset = A$$

**Theorem 10**:

$$A \cap A' = \emptyset$$
$$A \cup A' = S$$
$$A \cap A = A$$

$$A \cup A = A$$

**Theorem 11**: De Morgan's laws:

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

**Theorem 12**:

$A - B = A \cap B' = A \setminus B$

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

**Theorem 13**:

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

If $A$ and $B$ are disjoint, that is $P(A \cap B) = \emptyset$,
then $P(A \cup B) = P(A) + P(B)$

**Theorem 14**:

$P(\emptyset) = 0$

**Theorem 15**: Multiplicative Law of Probability

If there are two events A and B, probabilities of their happening being $P(A)$ and $P(B)$ respectively, then the probability $P(AB)$ of the simultaneous occurrence of the events A and B is equal to the probability of A multiplied by the conditional probability of B(i. e. the probability of B when A has occurred) or the probability of B multiplied by the conditional probability of A, i.e.

$$P(AB) = P(A)P(B/A)$$

$$= P(B)P(A/B$$

## Summary of Study Session 1

In this study session, you have learnt about:

1) The different approaches to the concept of Probability
2) The definition of some of the relevant terms in probability

## Self-Assessment Question for Study Session 1(SAQs) for Study Session 1

Having studied this session, you can now assess how well you have learnt its learning outcomes by answering the following questions. It is advised that you answer the questions thoroughly and discuss your answers with your tutor in the next Study Support Meeting. Also, brief answers to the Self-Assessment Questions are provided at the end of this module as a guide.

**SAQ (Test of learning outcome)**

Using set theory, show that:

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

$$A \cap S = A$$

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

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Gupta, B. D. (2001). *Mathematical Physics*. Vikas Publishing House PVT Ltd.

Hogg, R. V. and Craig, A. T. (1970). *Introduction to Mathematical Statistics*. New York: Macmillan Publishing Co.

# Study Session 2: Conditional Probability and Independence

## Expected duration: 1 week or 2 contact hours

## Introduction

One of the most important concepts in the theory of probability is based on the question: How do we modify the probability of an event in light of the fact that something new is known? What is the chance that we will win the game now that we lost our first point? What is the chance that I am a carrier of a diabetic disease now that my first child does not have the genetic condition?

What is the chance that a child will be a smoker if the household has two parents who smoke? This question leads us to the concept of conditional probability. In study sessions one, you learnt about three major approaches to the concept of probability. Some rules governing probability were also highlighted. Here you will learn the concepts of conditional probability and independence.

## Learning Outcomes for Study Session 2

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

2.1 Explain the concept of conditional probability
2.2 Explain the concept of independence
2.3 Solve problems related to these concepts.

## 2.1 Conditional Probability

Let S be a sample space, A, a collection events and P(.), a probability function. Then, we have a probability space [S, A, P(.)].

Definition 1: Let A and B be two events in A of the given probability space [S, A, P(.)].

The conditional probability of event A given event B, denoted by P[A/B], is defined by

$$P[A/B] = \frac{P[A \cap B]}{P(B)}, P(B) > 0 \text{ ...........(1)}$$

**Example 1**: Two students are chosen at random from a class consisting of 18 boys and 12 girls. What is the probability that the two students selected are:

(a) both boys (b) both girls (c) of the same sex (d) a boy and a girl.

**Solution**: Let $B_1$ to be the event that the first student selected is a boy.

Let $B_2$ be the event that the second student selected is a boy.

Let $B_1 \cap B_2$ denote the event that the two students selected are both boys.

(i) $P(B_1 \cap B_2) = P(B_1)\ ,\ P(B_2 / B_1)$ where

$$P(B_1) = \frac{18}{30} = 3/5$$

$$P(B_2 / B_1) = 17/29$$

Therefore, $P(B_1 \cap B_2) = \frac{3}{5}\ x\ \frac{17}{29}$

$$= \frac{51}{145}$$

(ii) Let $G_1 \cap G_2$ denote the event that the two students selected are both girls.

$$\text{P}\, G_1 \cap G_2 = P(G_1)\ \ P(G_2 / G_1)$$

$$= \frac{12}{30}\ x\ \frac{11}{29}$$

$$= \frac{132}{870} = \frac{22}{145}$$

(iii) $B_1B_2 \cup G_1G_2$ is the event that both students selected are of the same sex.

$$P(B_1B_2 \cup G_1G_2) = P(B_1B_2) + P(G_1G_2)$$

Since $B_1 \cap B_2$ *and* $G_1G_2$ are mutually exclusive

$$\therefore\ P(B_1B_2 \cup G_1G_2) = \frac{51}{145} + \frac{22}{145}$$

$$= \frac{73}{145}$$

(iv) $B_1G_2 \cup G_1B_2$ is the event that the two students selected are a boy and a girl.

$$(B_1G_2 \cup G_1B_2) = P(B_1G_2) + P(G_1B_2)$$

$$= P(B_1) \,.\, P(G_2 / B_1) + P(G_1) \,.\, P(B_2 / G_1)$$

$$= \frac{18}{30}\ x\ \frac{12}{29} + \frac{12}{30}\ x\ \frac{18}{29}$$

$$= \frac{3}{5}\ x\ {}^{12}\!/_{29} + \frac{2}{5}\ x\ {}^{18}\!/_{29} = {}^{72}\!/_{145}$$

**Example 2**: A boy has 10 identical marbles in a container consisting of 6 red and 4 blue marbles. He draws two marbles at random one after the other from the container without replacement. Find the probability that:

(a) the first draw is red while the second is blue

(b) both draws are of the same colour

(c) both draws are of different colours.

**Solution**:

(a) Let $R_1$ be the event that the first draw is red

Let $B_2$ be the event that the second draw is blue.

The event $R_1 \cap B_2$ is the event that the first draw is red while the second draw is blue.

$$P(R_1 \cap B_2) = P(R_1) \,.\, P(B_2 / R_1) \text{ where}$$

$$P(R_1) = {}^{6}\!/_{10}$$

$$P(B_2 / R_1) = {}^{4}\!/_{9}$$

$$\therefore P(R_1 \cap B_2) = \frac{6}{10} \; x \; \frac{4}{9}$$

$$= \frac{4}{15}$$

(b) Let $R_1$ be the event that the first draw is red.

Let $R_2$ be the event that the second draw is red.

Let $B_1$ be the event that the first draw is blue

Let $B_2$ be the event that the second draw is blue.

Therefore

$P(R_1R_2 \cup B_1B_2) = P(R_1R_2) + P(B_1B_2)$ since $R_1R_2$ and $B_1B_2$ are mutually exclusive.

$$P(R_1R_2) = P(R_1) \; P(R_2 / R_1)$$

$$= \frac{6}{10} \; x \; \frac{5}{9}$$

$$= 1/3$$

$$P(B_1B_2) = P(B_1) \; x \; P(B_2 / B_1)$$

$$= \frac{4}{10} \; x \; \frac{3}{9}$$

$$= \frac{12}{15}$$

Therefore, $P(R_1R_2 \cap B_1B_2) = \frac{1}{3} + \frac{2}{15} = \frac{7}{15}$

## 2.2 Independence

Recall that P [A/B] $= \dfrac{P[A \cap B]}{P(B)} \quad P(B) > 0$

Definition 2: Two events $A$ and $B$ are said to be stochastically or statistically independent if and only if any one of the following conditions is satisfied:

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

(ii) $P(A/B) = P(A)$ if $P(B) > 0$

(iii) $P(B/A) = P(B)$ if $P(A) > 0$

It is easily shown that (i) implies (ii), (ii) implies (iii) and (iii) implies (i). See Post-test (2).

Therefore, $P(A \cap B) = P(A/B)\,P(B) = P(B/A)\,P(A)$ if P(A) and P(B) are non-zero. This implies that one of the events is independent of the other. In fact,

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

So, if P(A), P(B) > 0 and one of the events is independent of the other, then the second event is also independent of the first. Thus, independence is a symmetric relation.

Remark: Two mutually exclusive events A and B are independent if and only if P(A) P(B) = 0 which is true if and only if either P(A) or P(B) = 0

Also, if $P(A) \neq 0$ and $P(B) \neq 0$, then A and B independent implies that they are not mutually exclusive.

Definition 2: Events $A_1, A_2, \ldots, A_n$ from $A$ in the probability space [$S$, $A$, $P(.)$] are said to be completely independent if and only if

(i) $P(A_i \cap A_j) = P(A_i)\,P(A_j)$ for $i \neq j$

(ii) $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$ for $i \neq j$, $j \neq k$, $i \neq k$

(iii) $P\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} P(A_i)$

Note: (i) These events are said to be pairwise independent if

$$P(A_i \cap A_j) = P(A_i)\,P(A_j) \text{ for all } i \neq j$$

Pairwise independence does not imply independence

*A* and *B* mutually exclusive implies that they are not independent.

**Example 3**: Suppose two dice are tossed. Let A denote the event of an odd total, B, the event of an ace on the first die, and C the event of a total of seven.

Are A and B independent?

Are A and C independent?

Are B and C independent?

**Solution**:

P[A/B] = $\frac{1}{2}$ = P(A)

P[A/C] = 1 ≠ P[A] = ½

P[C/B] = $\frac{1}{6}$ = P(C) = $^1/_6$

So, A and B are independent

A is not independent of C

B and C are independent

**Example 4**: Let $A_1$ denote the event of an odd face on the first die,

Let $A_2$ denote the event of an odd face on the second die,

Let $A_3$ denote the event of an odd total in the random experiment consisting of two dice. Then,

$$P(A_1)\,P(A_2) = \tfrac{1}{2}\ x\ \tfrac{1}{2} = P(A_1 \cap A_2)$$

$$P(A_1)\,P(A_3) = \tfrac{1}{2}\ x\ \tfrac{1}{2} = P[A_3 / A_1]\,P(A_1) = P(A_1 \cap A_3)$$

$$P(A_2 \cap A_3) = \tfrac{1}{4} = P(A_2)\ P(A_3)$$

Therefore, $A_1$, $A_2$ and $A_3$ are pairwise independent.

But $P(A_1 \cap A_2 \cap A_3) = 0 \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3)$

So, $A_1$, $A_2$ and $A_3$ are not independent.

## Summary of Study Session 2

In this study sessions, you have learnt:

1. The concepts of conditional probability and independence.
   Conditional probability is given by:
   P[A/B] =
2. Two events A and B are said to be stochastically or statistically independent if and only if any one of the following conditions is satisfied:
   a. $P(A \cap B) = P(A)P(B)$

## Self-Assessment Questions(SAQs)for Study Session 2

A random sample of 60 candidates who sat for Part I and II of an examination in 1984 is taken. The table below shows the number of candidates who passed or failed each part of the examination.

| | | Part I | | Total |
|---|---|---|---|---|
| Part II | Pass | Pass | Fail | |
| | Fail | | | |
| | | | 20 | 35 |
| Total | | 24 | | 60 |

copy and complete the table

if a candidate is chosen at random from the sample, use the table to find the probability that the candidate:

passed part II

passed parts I and II

passed part II but failed part I.

if a candidate is chosen at random from the subgroup of those who failed Part I, find the probability that the candidate passed Part II.

Given that:

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

$P(A/B) = P(A)$ if $P(B) > 0$

$P(B/A) = P(B)$ if $P(A) > 0$

Show that (i) implies (ii), (ii) implies (iii) and (iii) implies (i)

Consider the experiment of tossing 2 coins. Let the sample space S = {(H,H), (H,T), (T,H), (T,H)} and assume that each point is equally likely. Find:

the probability of two heads given a head on the first coin

the probability of two heads given at least one head.

Given that two dices are tossed. What is the probability that their sum will be 6 given that one face shows 2?

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Gupta, B. D. (2001). *Mathematical Physics*. Vikas Publishing House PVT Ltd.

Hogg, R. V. and Craig, A. T. (1970). *Introduction to Mathematical Statistics*. New York: Macmillan Publishing Co.

# Study Sessions 3: Bayes Theorem and Total Probability

## Expected duration: 1 week or 2 contact hours

## Introduction

When the ideas of probability are applied to Market a product, there are occasions when you need to calculate conditional probabilities other than those already known. For example, if the product purchase in a city is rated at 100%, what is the probability that the next city will also purchase on a high scale? You might need to know the probability of getting a product outcome ranging from 10% to 90%, there comes Bayes Theorem and Total Probability.

In study sessions two, our discussion centred on conditional probability and independence of occurrence of events. In this study sessions, we proceed further by building on the concepts through the concepts of Bayes theorem and total probability.

## Learning Outcomes for Study Session 3

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

3.1 Explain the concepts of Bayes theorem and total probability.
3.2 Solve problems on the twin concepts of Bayes theorem and total probability.

## 3.1 Bayes theorem

Given that P(A/B) = $\frac{P(A \cap B)}{P(B)}$ - (1)

P(B/A) = $\frac{P(B \cap A)}{P(A)}$ - (2)

This implies that $P(A \cap B) = P(B \cap A) = P(B/A)P(A)$ - (3)

Therefore P(A/B) = $\frac{P(B/A)P(A)}{P(B)}$ -(4)

Equation (4) is known as Bayes theorem.

## 3.2 Total Probability

But $P(B) = P(B \cap A) + P(B \cap A^1)$ [from $P(B) - P(B \cap A) = P(B \cap A^1)$]

$= P(B/A)P(A) + P(B/A^1)P(A^1)$

Generally, $P(B) = \sum_{i=1}^{n} P(B \cap Ai) = \sum_{i=1}^{n} P(B/Ai)P(Ai)$ - (5)

Equation (5) is the Total probability

Therefore, Bayes' rule can be written as

$$P[A_k / B] = \frac{P[B/A_k]P[A_k]}{\sum_{i=1}^{n} P[B/A_i]P[A_i]} \qquad \text{-(6)}$$

**Example 1**: The contents of 3 identical baskets $B_i$ (i = 1, 2, 3) are:

$B_1$: 4 apples and 1 orange

$B_2$: 1 apple and 4 oranges

$B_3$: 2 apples and 3 oranges

A basket is selected at random and from it, a fruit is picked. The fruit picked turns out to be an apple on inspection. What is the probability that it come from the first basket

**Solution**:

Let E be the event of picking an apple.

Using the table below:

| State of Nature | $P(B_i)$ | $P(E/B_i)$ | $P(B_i)P(E/B_i)$ | $P(B_i/E)$ |
|---|---|---|---|---|
| $B_1$ (4A, 10) | $1/3$ | $4/5$ | $4/15$ | $4/7$ |
| $B_2$ (1A, 40) | $1/3$ | $1/5$ | $1/15$ | $1/7$ |
| $B_3$ (2A, 30) | $1/3$ | $2/5$ | $2/15$ | $2/7$ |
| Total | 1 | - | $7/15$ | 1 |

The required probability

$$P(B_1/E) = 4/7$$

i.e. $$P(B_1/E) = \frac{P(E/B_1)\,P(B_1)}{\sum_{i=1}^{3} P(B_i)\,P(E/B_i)}$$

$$= \frac{\frac{4}{5}\cdot\frac{1}{3}}{7/15} = 4/7$$

**Example 2**: In a certain town, there are only two brands of hamburgers available, Brand A and Brand B. It is known that people who eat Brand A hamburger have a 30% probability of suffering stomach pain and those who eat Brand B hamburger have a 25% probability of suffering stomach pain.

Twice as many people eat Brand B compared to Brand A hamburgers. However, no one eats both varieties. Supposing one day, you meet someone suffering from stomach pain who has just eaten a hamburger what is the probability that they have eaten Brand A and what is the probability that they have eaten, Brand B?

**Solution**: Let S denote people who have just eaten a hamburger

Let A denote people who have eaten Brand A hamburger

Let B denote people who have eaten Brand B hamburger

Let C denote people who are suffering stomach pains

We are given that

P(A) = $\frac{1}{3}$

P(B) = $\frac{2}{3}$

P(C/A) = 0.3

P(C/B) = 0.25

S = $A \cup B$

As those who have stomach pain have either eaten Brand A or B. then $A \cap B = \phi$

P(C) = $P(C \cap S) = P(C \cap A) + P(C \cap B)$

= P(C/A)P(A) + P(C/B)P(B)

$0.3 \times \frac{1}{3} + 0.25 \times \frac{2}{3}$

= $\frac{8}{30}$

Then P(A/C) = $\frac{P(C/A)P(A)}{P(C)}$

$$= \frac{0.3\,X\,(\frac{1}{3})}{\frac{8}{30}} = \frac{3}{8}$$

And P(B/C) = $\frac{P(C/B)P(B)}{P(C)} = \frac{0.25\,x\,(\frac{2}{3})}{\frac{8}{30}}$

$$= \frac{5}{8}$$

So, if someone has stomach pain, the probability that they have eaten Brand A hamburger is

$3/8$ and the probability that they have eaten Brand B is $5/8$.

## Summary of Study Session 3

In this study sessions, you have been able to;

1. Understand how to explain, define and prove the twin concepts of Bayes theorem and total probability.
2. Given any two events A and B in a sample space S, Bayes theorem is defined as:

P(A/B) =

Also, the denominator in the definition of Bayes theorem i.e. P(B) is defined as:SP(B)= ∑_(i=1)^n▒〖P(B∩A_i )= ∑_(i=1)^n▒〖P(B/A_i)P(A_i)〗 〗

## Self-Assessment Question(SAQs) for study Session 3

1. A certain brand of compact disc (CD) player has an unreliable integrated circuit$[IC]$, which fails to function on 1% of the models as soon as the player is connected. On 20% of these occasions, the light displays fail and the buttons fail to respond, so that it appears exactly the same as if the power connection is faulty. No other component failure causes that symptom. However, 2% of people who buy the CD player fail to fit the plug correctly, in such a way that they also experience a complete loss of power. A customer rings the supplier of the CD players saying that the light displays and buttons are not functioning on the CD. What is the probability that the fault is due to the IC failing as opposed to the poorly fitted plug?

An electronic has 3 components and the failure of any one of them may or may not cause the device to shut off automatically. Furthermore, these failures are the only possible causes for a shut-off and the probability that two of the components will fail simultaneously is negligible. At any time, component $B_1$ will fail with probability 0.1, component $B_2$ will fail with probability 0.3 and component $B_3$ will fail with probability 0.6. Also, if component $B_1$ fails, the device will shut off with probability 0.2; if component $B_2$ fails, the device will shut off with probability 0.5, if component $B_3$ fails, the device will shut off with probability 0.1. The device suddenly shuts off, what is the probability that the shut off was caused by the failure of component $B_1$.

Stores X, Y, Z sell brands A, B and C of men's shirts. A customer buys 50% of his shirts at X, 20% at Y and 30% at Z. Store X sells 25% brand A, 40% brand B and 25% brand C. Store Y sells 40% brand A, and 20% brand B and 30% brand C. Store Z sells 20% brand A, 40% brand B and 20% brand C. The customer comes home one day with a new shirt of brand C. What is the probability that it was purchased at store X?

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Amahia, G. N.(2007). *STA 211- Probability II. Ibadan Distance Learning Centre Series.* Distance Learning Centre, University of Ibadan.

Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Newnes. Elsevier. Linacre House, Jordan Hill, Oxford.

Olubusoye, O. E. (2000). *Study sessions notes on STA 311*.

# Study Sessions 4: Urn Models

## Expected duration: 1 week or 2 contact hours

## Introduction

You should have studied through the concept of conditional probability and proved the twin concepts Bayes theorem and total probability in the previous study session. Here in study session 4, you will learn how to discuss issues of counting, partitioning, selection and distribution of objects.

## Learning Outcomes for Study Session 4

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

4.1 State the fundamental principles of counting
4.2 Define the Stirling's number of the second kind
4.3 State the Stirling's identity
4.4 Apply (i)- (iv) in solving related problems

## 4.1 Sampling With and Without Replacement

Fundamental Principle of Counting

Consider a finite sequence of decisions. Suppose the number of choices for each individual decision is independent of decisions made previously in the sequence. Then, the number of ways to make the whole sequence of decisions is the product of these numbers of choices, i.e. n!

**Example 1**: The number of four-letter words that can be formed by rearranging the letters in the word PLAN is 4! =24.

| | | | | | |
|---|---|---|---|---|---|
| PLAN | PLNA | PALN | PANL | PNLA | PNAL |
| LPAN | LPNA | LAPN | LANP | LNPA | LNAP |
| APLN | APNL | ALPN | ALNP | ANPL | ANLP |
| NPLA | NPAL | NLPA | NLAP | NAPL | NALP |

Definition 1: The number of ordered selections of $r$ elements chosen from an $n$-element set is $P(n, r)$.

$$P(n,r) = n(n-1)(n-2)\ldots.(n-[r-1])$$

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

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

$$= r!\,C(n,r)$$

**Example 2**: Suppose 6 members of Adeola School's Parent Teachers Association meet to select a 3-member delegation to represent the association at a state-wide convention. If the laws stipulate that each delegation be consisted of a delegate, a first alternate and a second alternate. How many ways can the 6 members comply from among themselves?

Solution: $P(6,3) = 120$ ways or $3!\,C(6,3) = 120$ ways

Definition 2: The number of ways of making a sequence of $r$ decisions each of which has n choices is $n^r$ if order matters and the selection is with replacement.

Definition 3: The number of different ways to choose $r$ things from $n$ things with replacement if order does not matter is $C(r+n-1,\ r)$

Example 3: How many different three letter "words" can be produced using the "alphabet" ALEXY?

Solution: Since there are no restrictions on the number of times a letter can be used, $5^3 = 125$ such words can be formed.

Example 4: At a fundraising luncheon, each of 50 patrons is given a numbered ticket, while its duplicate is placed in a bowl from which prize-winning numbers will be drawn.

If the prizes are #50, #100, and #150, how many outcomes are possible assuming that winning tickets are not returned to the bowl.

If the prizes are the same, say, #70 each for the 3 prizes, how many outcomes are possible assuming that winning tickets are not returned to the bowl?

If the winning tickets are returned to the bowl, how many outcomes are possible when the prizes are as in (i) and (ii) respectively?

Solution:

$P(50,3) = 117600$ different outcomes are possible

Since the prizes are the same, then order is not important implying that there are $C(50,3) = 19600$ different possible outcomes.

a) Different prizes, with replacement, order matters: $50^3 = 125000$

b) Same prizes, with replacement, order does not matter:

$C(3 + 50 - 1, 3) = 22100$

Note: In choosing with replacement, elements may be chosen more than once. If order does not matter, then we are only concerned with the multiplicity with which each element is chosen.

**Table 5.1** summarizes the four scenarios that we have considered.

Table 5.1

| | **Order Matters** | **Order does not matter** |
|---|---|---|
| Without replacement | $P(n,r)$ | $C(n,r)$ |
| With replacement | $n^r$ | $C(r + n - 1, r)$ |

## 4.2 Stirling Numbers of the Second Kind

Definition 4: Let $S$ be a set. A partition of $S$ is an ordered collection of pairwise, disjoint, non-empty subsets of $S$ whose union is all of $S$. The subsets of a partition are called blocks.

For $S = A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_k$ to be a partition of $S$:

$A_i \cap A_j = \emptyset$ whenever $i \neq j$

$A_j \neq \emptyset$, $1 \leq j \leq k$

Two partitions are equal if and only if they have the same blocks.

For instance, $\{1\} \cup \{2,3\}$, $\{1\} \cup \{3,2\}$, $\{2,3\} \cup \{1\}$ and $\{3,2\} \cup \{1\}$ are 4 different looking ways of writing the same two-block partition of $S = \{1,2,3\}$

The other partition of $S = \{1,2,3\}$ are

$\{1\} \cup \{2\} \cup \{3\}$- 3 blocks

$\{1,2\} \cup \{3\}$- 2 blocks

$\{1,3\} \cup \{2\}$- 2 blocks

$\{1,2,3\}$- 1 block

Thus, S has a total of 5 different partitions made up of:

One of $\{1\} \cup \{2,3\}$, $\{1\} \cup \{3,2\}$, $\{2,3\} \cup \{1\}$ and $\{3,2\} \cup \{1\}$

$\{1\} \cup \{2\} \cup \{3\}$

$\{1,2\} \cup \{3\}$

$\{1,3\} \cup \{2\}$

$\{1,2,3\}$

Definition 5: The number partitions of $\{1,2,3,\ldots m\}$ into $n$ blocks is denoted by $S(m,n)$ and this is known as the Stirling number of the second kind.

Note: $S(m,n) = 0$ if $n < 1$ or $n > m$.

Also, $S(m,1) = 1 = S(m,m)$. This is because there is just one way to partition $\{1,2,3,\ldots m\}$ into a single block and
$\{1\} \cup \{2\} \cup \{3\} \cup \ldots\ldots \cup \{m\}$ is the unique unordered way of expressing $\{1,2,3,\ldots m\}$ as the disjoint union of $m$ non-empty subsets.

### 4.3 Stirling's Identity: For any two positive integers $m$ and $r$,

$$r!\,S(m,r) = \sum_{t=1}^{r} (-1)^{r+t} C(r,t)\, t^m$$

Therefore $S(m,r)=\frac{1}{r!}\sum_{t=1}^{r}(-1)^{r+t}C(r,t)\,t^m$

Example 5: $S(4,1)=C(1,1)1^4$

$=1$

$$S(4,2)=\frac{1}{2}[-C(2,1)1^4+C(2,2)2^4]$$

$=\frac{1}{2}[-2+16]=7$

$$S(4,3)=\frac{1}{6}[C(3,1)1^4-C(3,2)2^4+C(3,3)3^4]$$

$=\frac{1}{6}[3-48+81]=6$

$$S(4,4)=\frac{1}{24}[-C(4,1)1^4+C(4,2)2^4-C(4,3)3^4+C(4,4)4^4]$$

$=\frac{1}{24}[-4+96-324+256]=1$

## 4.4 Application of Stirling's number of the second kind to distribution of objects into urns

We are interested in the question "In how many different ways can $m$ balls be distributed among $n$ urns?" We are going to answer this question by considering whether the balls and urns are labelled or not and whether a particular urn can be left empty?

We will consider 4 variations:

Variation 1: In how many ways can $m$ labelled balls be distributed among $n$ unlabelled urns if no urn is left empty? This is the same as "In how many ways can the set $\{1,2,3,\ldots m\}$ be partitioned into $n$ blocks. This is $S(m,n)$.

Example 6: In how many ways can 4 labelled balls be distributed among 2 unlabelled urns if no urn is left empty?

Solution: $S(4,\ 2)=7$ that is if the balls are labelled $1,2,3,4$ then the 7 possibilities are

$\{1\}\&\{2,3,4\}$

$\{2\}\&\ \{1,3,4\}$

$\{3\}\&\ \{1,2,4\}$

$\{4\}\&\ \{1,2,3\}$

$\{1,2\}\&\ \{3,4\}$

$\{1,3\}\&\{2,4\}$

$\{1,4\}\&\{2,3\}$

Because the urns are unlabelled,

$\{2\}\ \&\ \{1,3,4\} = \{1,3,4\}\ \&\ \{2\}$ etc.

Variation 2: In how many ways can $m$ labelled balls be distributed among $n$ unlabelled urns?

Solution: This is $S(m,1) + S(m,2) + \cdots + S(m,n)$. This is the same as finding the number of ways in which $\{1,2,\ldots,m\}$ can be partitioned into $n$ or fewer blocks since it is no longer a requirement that no urn be left empty.

Example 7: The number of ways to distribute four labelled balls among two unlabelled urns is $S(4,1) + S(4,2) = 1 + 7 = 8$, i. e. $S(4,1) = \{1,2,3,4\}\&\{\ \ \}$ $S(4,2) = \{1\}\&\{2,3,4\}$, $\{2\}\&\{1,3,4\}$, $\{3\}\&\{1,2,4\}$, $\{4\}\&\{1,2,3\}$, $\{1,2\}\&\{3,4\}$,$\{1,3\}\&\{2,4\}$, $\{1,4\}\&\{2,3\}$

Variation 3: In how many ways can $m$ labelled balls be distributed among $n$ labelled urns? This is $n^m$.

Example 8: Five labelled balls can be distributed among 3 labelled urns in $3^5 = 243$ ways.

Variation 4: In how many ways can $m$ labelled balls be distributed among $n$ labelled urns if no urn is left empty? This is $n!\,S(m,n)$.

There are $S(m,n)$ ways to distribute $m$ labelled balls among $n$ unlabelled urns using variation 1. After the distribution of the balls, there are $n!$ ways to label the urns. By the fundamental principle of counting, the answer is $n!\,S(m,n)$.

Example 9: In how many ways can 5 labelled balls be distributed among 3 labelled urns if no urn is left empty?

Solution: $3!\,S(5,3)$

## Summary of Study Session 4

In this study sessions, you have learnt about:

1. The concept of sampling with and without replacement.
2. Suppose the number of choices for each individual decision is independent of decisions made previously in the sequence. Then, the number of ways to make the whole sequence of decisions is the product of these numbers of choices, i.e. n!. This is known as the fundamental principle of counting.
3. Stirling number of the second kind is the number partitions of {1,2,3,…m} into n blocks and it is denoted by S(m,n)
4. Stirling's identity is defined as r!S(m,r)=∑_(t=1)^r▒〖(-1)^(r+t) C(r,t) 〗 t^m application of Stirling's number of the second kind to distribution of objects into urns

## Self-Assessment Questions(SAQs) for study Session 4

1. Show that:

a. $nP(n-1,r) = P(n,r+1)$

b. $P(n+1,r) = rP(n,r-1) + P(n,r)$

2. In how many ways can four elements be chosen from a ten-element set:

a. with replacement if order matters?

b. with replacement if order does not matter?

c. without replacement if order does not matter?

d. without replacement if order matters?

3. In how many ways can six balls be distributed among four urns if :

a. the urns are labelled but the balls are not?

b. the balls are labelled but the urns are not?

c. both balls and urns are labelled?

d. neither balls nor urns are labelled?

## References


Merris, R. (2003). Combinatorics. Wiley-Interscience.

Rosen, K. H., Michaels, J. G., Gross, J. L., Grossman, J. W. and Shier, D. R. (2000). *Handbook of Discrete and Combinatorial Mathematics*. CRC Press.

3. Brualdi, R. A. (1999). *Introductory Combinatorics*. Pearson Education Asia Limited and China Machine Press.

4. Grimaldi, R. P. (1999). *Discrete and Combinatorial Mathematics*. Pearson Addison-Wesley.

5. Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

# Study Sessions 5: Principle of Inclusion and Exclusion

## Expected duration: 1 week or 2 contact hours

## Introduction

In the previous study session, we discussed the fundamenta5tprinciple of counting. In this study sessions, we shall discuss the second counting principle; solve set problems using both the Venn diagram and the Principle of inclusion and exclusion.

## Learning Outcomes for Study Session 5

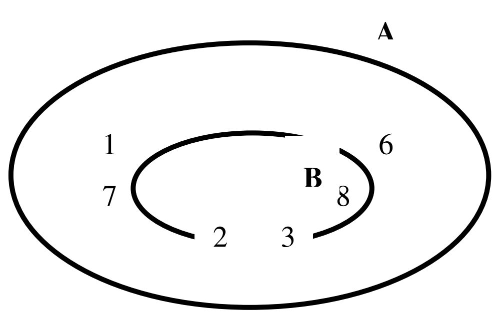
At the end of the study sessions, students should be able to:

5.1 Use the Venn diagrams and the Principle of Inclusion and Exclusion to solve problems in set theory.
5.2 State the Principle of Inclusion and Exclusion

## 5.1 Venn Diagrams

A set is a collection of objects, which can be distinguished from each other. The objects comprising the set are called the elements of the set and they may be finite or infinite in number.

Venn diagrams are diagrammatical representation of sets. For instance, consider the set $A = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$, suppose that $A$ has a subset $B = \{2, 3, 4, 5\}$. The diagrammatic representation of this is shown below.

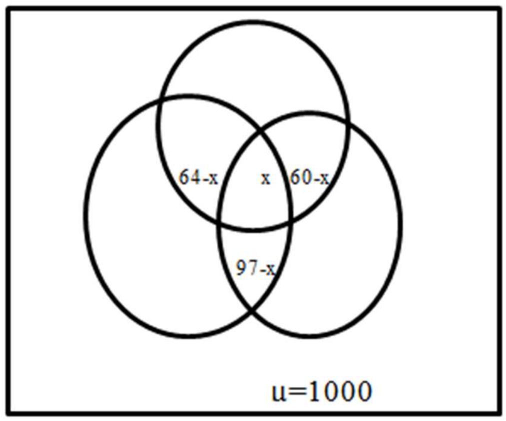


### 5.1.1 Solving Problems using Venn diagrams

Example 1: In a sample of 1000 foodstuff stores taken at an Ibadan market, the following facts emerged:

200 of them stock rice, 240 stock beans, 250 stock garri, 64 stock both beans and rice, 97 stock both rice and garri, while 60 stock beans and garri. If 430 do not stock rice, do not stock beans and do not stock garri, how many of the stores stock rice, beans and garri?

Solution:




Let: R represent rice stores

B represents beans stores

G represents garri stores

Let $x$ represents those that stock all the 3 food items

Those that stock garri alone are $250 - [(97 - x) + x + (60 - x)] = 93 + x$

Those that stock beans alone are $240 - [(60 - x) + (x) + (64 - x)] = 116 + x$

Those that stock rice alone are $200 - [(64 - x) + x + (97 - x)] = 39 + x$

430 did not stock any of the food items

Therefore,$1000 = (39 + x) + (93 + x) + (116 + x) + x + (64 - x) + (60 - x) + (97 - x) + 430$

And $x = 1000 - 899 = 101$

Therefore 101 stores stock rice, beans and garri.

## 5.2 The Principle of Inclusion and Exclusion

5.2.1 The Second Counting Principle

If a set is the disjoint union of two (or more) subsets, then the number of elements in the set is the sum of the numbers of elements in the subsets, i.e.

$n(A \cup B) = n(A) + n(B)$ implying that $|A \cup B| = |A| + |B|$ if $A$ and $B$ are disjoint.

Theorem 1: $|A \cup B| < |A| + |B|$ if A and B are not disjoint.

This is because $|A| + |B|$ counts every element of $A \cap B$ twice. Let us illustrate this with the following example.

Example 2: If $A = (2, 3, 4, 5, 6)$, $|A| = 5$ and $B = (3, 4, 5, 6, 7)$, $|B| = 5$

then, $|A| + |B| = 10$

$A \cup B = 2, 3, 4, 5, 6, 7$

$|A \cup B| = 6$

Since $A$ and $B$ are not disjoint, $|A \cup B| < |A| + |B|$

Compensating for this double counting yields the formula

$|A \cup B| = |A| + |B| - |A \cap B|$............eqn.(1)

From our example, $A \cap B = 3, 4, 5, 6$

$|A \cap B| = 4$

$|A \cup B| = 5 + 5 - 4$

$= 6$

thus proving equation (1)

Theorem 2: $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$ for three sets $A$, $B$ and $C$.

Proof:

We know from equation (1) above that $|A \cup B| = |A| + |B| - |A \cap B|$

Then, for 3 sets, $|A \cup B \cup C| = |A \cup [B \cup C]|$

$= |A| + |B \cup C| - |A \cap [B \cup C]|$

Applying equation (1) to $|B \cup C|$ gives

$|A \cup B \cup C| = |A| + [|B| + |C| - |B \cap C|] - |A \cap [B \cup C]|$………eqn (2)

Because $A \cap [B \cup C] = (A \cap B) \cup (A \cap C)$, we can apply equation(1) again to obtain

$|A \cap [B \cup C]| = |A \cap B| + |A \cap C| - |A \cap B \cap C|$………eqn (3)

Finally, a combination of equations (2) and (3) yields

$|A \cup B \cup C| = [|A| + |B| + |C|] - [|A \cap B| + |A \cap C| + |B \cap C|] + |A \cap B \cap C|$………eqn (4)

Thus proving theorem 2.

From this derivation, we notice that an element of $A \cap B \cap C$ is counted 7 times in equation(4), the first 3 times with a plus sign, then 3 times with a minus sign and then once more with a plus sign.

Example 3: If $A = \{1, 2, 3, 4\}$ $B = \{3, 4, 5, 6\}$ $C = \{2, 4, 6, 7\}$ then

$A \cup B \cup C = \{1, 2, 3, 4, 5, 6, 7\}$

$|A \cup B \cup C| = 7$………………..(a)

$|A| = 4$

$|B| = 4$

$|C| = 4$

$|A| + |B| + |C| = 12$

$A \cap B = 3, 4, \quad A \cap C = 2, 4, \quad B \cap C = 4, 6$

In this example, $|A \cap B| = |A \cap C| = |B \cap C| = 2$ so that

$|A \cap B| + |A \cap C| + |B \cap C| = 6$ and

$A \cap B \cap C = 4, \quad |A \cap B \cap C| = 1$

Therefore, $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$

$= 12 - 6 + 1 = 7$……………… (b)

Thus, (a) = (b) thus establishing theorem 2.

Generally, the Principle of Inclusion and Exclusion (PIE) states that:

If $A_1, A_2, \ldots, A_n$ are finite sets, the cardinality of their union

$|A_1 \cup A_2 \cup \ldots \cup A_n| =$

$$\sum_{i=1}^{n} |A_i| - \sum_{1 \le i < j \le n} |A_i \cap A_j| + \sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |\cap A_i|$$

Proof:

On the left is the number of elements in the union of n sets. On the right, we first count elements in each of the sets separately and add them up. If the sets $A_i$ are not disjoint, the elements that belong to at least two of the sets $A_i$, or the intersections $A_i \cap A_j$, are counted more than once. We wish to consider every such intersection, but each only once. Since $A_i \cap A_j = A_j \cap A_i$, we should consider only pairs $(A_i, A_j)$ with $i < j$.

When we subtract the sum of the number of elements in such pairwise intersections, some elements may have been subtracted more than once. Those are the elements that belong to at least three of the sets $A_i$. We add the sum of the elements of intersections of the sets taken three at a time. (Note: the condition $i < j < k$ ensures that every intersection is counted only once)

The process continues with sums being alternately added or subtracted until we come to the last term which is the intersection of all sets $A_i$ thus proving the theorem.

Let $S = A_1 \cup A_2 \cup \ldots .\cup A_n$ and $A_i{}^c = S\backslash A_i$ then the PIE principle can also be expressed as

$|A_1{}^c \cap \ldots ..\cap A_n{}^c| =$

$$|S| - \sum_{i=1}^{n} |A_i| + \sum\nolimits_{1\le i<j\le n} |A_i \cap A_j| - \sum\nolimits_{1\le i<j<k\le n} |A_i \cap A_j \cap A_k| + \cdots - (-1)^{n+1} |\cap A_i|$$

Example 4: Let $A$ be the subset of the first $700$ hundred numbers $S = \{1, 2, \ldots ..,700\}$ that are divisible by 7. Find the number of elements in $S$ that are not divisible by 7.

Solution:

$$A = \{7, 14, 21, 28, 35, 42, 49, \ldots ..\}$$

$$|A| = 100$$

$$|A^c| = |S| - |A|$$

$$= 700 - 100$$

$$= 600$$

Example 5: Find the number of integers from 1 to $1000$ that are not divisible by 5, 6 and 8

Solution: Let $A_1, A_2, A_3$ be the subset consisting of those integers that are divisible by 5, 6 and 8. The number we are interested in is

$$\left|A_1^1 \cap A_2^1 \cap A_3^1\right| = 1000 - |A_1| - |A_2| - |A_3| + |A_1 \cap A_2| + |A_1 \cap A_3| + |A_2 \cap A_3| - |A_1 \cap A_2 \cap A_3|$$

$$|A_1| = \left\lfloor \frac{1000}{5} \right\rfloor = 200 \quad |A_2| = \left\lfloor \frac{1000}{6} \right\rfloor = 166 \quad |A_3| = \left\lfloor \frac{100}{8} \right\rfloor = 125$$

Note: The results for $|A_1|, |A_2|$ and $|A_3|$ were achieved by using the round down, notation $\lfloor \ \rfloor$ which involves the dropping of the fractional part.

To compute the number in a 2 and 3 – set interaction, we use the least common multiple (LCM), i.e.

$$|A_1 \cap A_2| = \left\lfloor \frac{1000}{30} \right\rfloor = 33$$

$$|A_1 \cap A_3| = \left\lfloor \frac{1000}{40} \right\rfloor = 25$$

$$|A_2 \cap A_3| = \left\lfloor \frac{1000}{24} \right\rfloor = 41$$

$$|A_1 \cap A_2 \cap A_3| = \left\lfloor \frac{1000}{12} \right\rfloor = 8$$

Thus, $|A_1^1 \cap A_2^1 \cap A_3^1| = 1000 - 200 - 166 - 125 + 33 + 25 + 41 - 8 = 600$

## Summary of Study session 5

In this study session, you have learnt about:

1. Studied Venn diagrams
2. Learnt about the second counting principle: If a set is the disjoint union of two (or more) subsets, then the number of elements in the set is the sum of the numbers of elements in the subsets, i.e.

n(A∪B)=n(A)+n(B) implying that |A∪B|=|A|+|B| if A and B are disjoint.

stated the Principle of Inclusion and Exclusion:

|A_1∪A_2∪….∪A_n |=

∑_(i=1)^n▒〖|A_i |-∑_(1≤i<j≤n)▒〖|A_i∩A_j |+〗〗 ∑_(1≤i<j<k≤n)▒〖|A_i∩A_j∩A_k |-…+(-1)^(n+1) 〗 |∩A_i |

Used Venn diagrams and the Principle of Inclusion and Exclusion to solve problems in set theory.

## Self-Assessment Questions(SAQs) for Study Session 5

Among the mathematics courses offered at Atise High School are Algebra, Geometry and Trigonometry. To be in the Mathematics club, a student must have completed at least one of these three courses. The Mathematics Club has 56 students members altogether. Of these, 28 have taken Algebra and 28 have taken Geometry, 11 have taken both Algebra and Geometry, 12 have taken both Algebra and Trigonometry and 13 have taken both Geometry and Trigonometry. If 5 of the students have taken all 3 courses, how many have taken Trigonometry?

Solve the problem using Venn diagrams

Solve the problem using the principle of Inclusion and Exclusion

Which is easier?

Let $S = \{1, 2, \ldots\ldots, 700\}$ and $A_1$ be the subset of all elements divisible by $7$ and $A_2$ be the subset of all elements divisible by $10$. Find the number of elements that are not divisible by $7$ nor $10$.

Find the number of integers from $1$ to $2000$ that are not divisible by $5$, $6$ and $8$.

## References


Merris, R. (2003). Combinatorics. Wiley-Interscience.

Rosen, K. H., Michaels, J. G., Gross, J. L., Grossman, J. W. and Shier, D. R. (2000). *Handbook of Discrete and Combinatorial Mathematics*. CRC Press.

Brualdi, R. A. (1999). *Introductory Combinatorics*. Pearson Education Asia Limited and China Machine Press.

Grimaldi, R. P. (1999). *Discrete and Combinatorial Mathematics*. Pearson Addison-Wesley.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Olubusoye, O. E. (2010). *Study sessions notes on STA 311*.

# Study Sessions 6: Allocation and Matching Problems

## Expected duration: 1 week or 2 contact hours

## Introduction

In study sessions five, you considered the Principle of Inclusion and Exclusion (PIE). In this study sessions, we shall utilize the PIE in considering problems relating to allocation and matching of objects or persons. The probability aspects of these issues shall also be considered.

## Learning Outcomes for Study Session 6

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

6.1 Explain the concept of derangement
6.2 Solve allocation and matching problems

## 6.1 Derangements

Consider the following example as an illustration:

Example 1: At a party, 10 gentlemen check their hats. In how many ways can their hats be returned so that no gentleman gets the hat with which he arrived?

This problem consists of an n-element set X in which each element has a specified location. We are required/asked to find the number of permutations of the set X in which no element is in its specified location.

Here, the set X is the set of 10 hats and the specified location of a hat is (the head of) the gentlemen to which it belongs.

Let us take X to be the set {1, 2,……., 10} in which the location of each of the integers is that specified by its position in the sequence 1, 2,…….10.

Definition 1: A derangement of (1, 2,….n) is a permutation $i_1$, $i_2$,….., $i_n$ of (1, 2, …,n) such that $i_1 \neq 1$, $i_2 \neq 2$, …, $i_n \neq n$.

Thus, a derangement of (1, 2,…,n) is a permutation $i_1, i_2, \ldots, i_n$ of (1, 2,…,n) in which no integer is in its natural position: $i_1 \neq 1$, $i_2 \neq 2$, …, $i_n \neq n$.

Denote by $D_n$ the number of derangement of (1, 2, …., n)

Theorem: For n ≥ 1, $D_n = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots\ldots + (-1)^n \frac{1}{n!}\right]$

Proof: Let S be the set of all n! permutations of (1, 2, …..,n). For j = 1, 2,…, n, let $p_j$ be the property that in a permutation, j is in its natural position. Thus, the permutation $i_1, i_2, \ldots, i_n$ of (1, 2, ….,n) has property $p_j$ provided $i_j = j$. A permutation of (1, 2,…n) is a derangement if and only if it has none of the properties $p_1, p_2, \ldots, p_n$.

Let $A_j$ denote the set of permutations of (1, 2,….n) with property $p_j$, (j = 1, 2, ….., n). The derangements of (1,2,…,n) are those permutations in $A_1^c \cap A_2^c \cap \ldots\ldots \cap A_n^c$.

Thus, $D_n = \left|A_1^C \cap A_2^c \cap \ldots\ldots \cap A_n^c\right|$

The PIE is used to evaluate $D_n$ as follows:

The permutation in $A_1$ are of the form $1, i_2 \ldots i_n$, where $i_2 \ldots i_n$ is a permutation of (2,……,n). Thus, $|A_1| = (n-1)!$ And more generally for $|A_j| = (n-1)!$ for j = 1, 2, ……,n.

The permutations in $A_1 \cap A_2$ are of the form $1, 2, i_3 \ldots i_n$ where $i_3 \ldots i_n$ is a permutation of (3, …….,n). Thus, $|A_1 \cap A_2| = (n-2)!$

Generally, $|A_i \cap A_j| = (n-2)!$ for any 2 combinations $(i, j)$ of (1, 2,….,n).

For any integer k, with 1 ≤ k ≤ n, the permutations in $A_1 \cap A_2 \cap \ldots \cap A_k$ are of the form 1, 2, …,k, $i_{k+1} \ldots i_n$, where $i_{k+1} \ldots, i_n$ is a permutation of (k+1,…..,n). Thus, $|A_1 \cap A_2 \cap \ldots \cap A_k| = (n-k)!$.

Generally, $|A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}| = (n-k)!$ for any k-combination $(i_1, i_2, \ldots i_k)$ of (1, 2,….n):

Since there are $\binom{n}{k}$ k – combinations of (1, 2, …,n), applying the inclusion-exclusion principle, we obtain:

$$D_n = n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \binom{n}{3}(n-3)! + \ldots\ldots + (-1)^n \binom{n}{n}(n-n)!$$

$$= n! - \frac{n!}{1!} + \frac{n!}{2!} - \frac{n!}{3!} + \dots\dots\dots + (-1)^n \frac{n!}{n!}$$

$$= n! \left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \dots\dots\dots + (-1)^n \frac{1}{n!}\right]$$

Thus, from example 1 above,

$$D_{10} = 10! \left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \frac{1}{10!}\right]$$

You should be able to supply the final answer for $D_{10}$

Note: (i) The series expansion for $e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \dots\dots$

(ii) $\frac{Dn}{n!}$ is the ratio of the number of derangement of (1, 2,….,n) to the total number of permutations of (1, 2,….n).

Thus, $\frac{Dn}{n!}$ is the probability that it is a derangement if we select a permutation of (1, 2,…,n) at random.

## 6.2 The Matching Problem

Suppose that an absent minded secretary prepares *n* letters and envelopes to send to *n* different people. If the letters were randomly stuffed into the envelopes, a match occurs if a letter is inserted in the proper envelope.

**Example 2**: Suppose that each of *N* men in a room throws his shirt into the centre of the room. The shirts are first mixed up and then each man randomly selects a shirt.

What is the probability that none of the men selects his own shirt?

What is the probability that at least one of the men selects his own shirt?

What is the probability that exactly *k* of the men select their own shirt?

**Solution**:

From our discussion on derangement, the probability that none of the men selects his own shirt is

$$\frac{D_N}{N!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \text{........} + (-1)^N \frac{1}{N!}$$

The probability that at least one of the men selects his own shirt is

1 – Prob [None selects his own shirt]

$$= 1 - \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \text{........} + \frac{(-1)^N}{N!}\right]$$

$$= 1 - 1 + 1 - \frac{1}{2!} + \frac{1}{3!} \text{...........} - \frac{(-1)^N}{N!}$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} \text{...................} - \frac{(-1)^N}{N!}$$

The probability that exactly k of the men select their own shirt is as follows: First fix attention on a particular set of *k* men. The number of ways in which this and only this k men can select their own shirt is equal to the number of ways in which the other N-k men can select among their shirts in such a way that none of them selects his own shirt.

The probability that none of the N-K men, (selecting among their shirts), selects his own shirt is $1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \text{..........} + \frac{(-1)^{N-K}}{(N-K)!}$

It follows that the number of ways in which the set of men selecting their own shirts corresponds to the set of k men under consideration is

$$(N-K)! \left[1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \text{..........} + \frac{(-1)^{N-K}}{(N-K)!}\right]$$

Also, as there are $\binom{N}{K}$ possible selections of a group of K men, it follows that there are

$$\binom{N}{K}(N-K)!\left[1-1+\frac{1}{2!}-\frac{1}{3!}+\ldots+\frac{(-1)^{N-K}}{(N-K)!}\right]$$

ways in which exactly K of the men select their own shirts.

The probability required is thus

$$\frac{\binom{N}{K}(N-K)!\left[1-1+\frac{1}{2!}-\frac{1}{3!}+\ldots+\frac{(-1)^{N-K}}{(N-K)!}\right]}{N!}$$

$$\frac{1-1+\frac{1}{2!}-\frac{1}{3!}+\ldots\ldots+\frac{(-1)^{N-K}}{(N-K)!}}{K!}$$

This result is approximately $\frac{e^{-1}}{K!}$, for large N, k = 0,1, ……….

## Summary of Study Session 6

In this study sessions, you have:

1. Learnt about the problems involving allocation and matching of objects/persons from various perspectives including inability to match, exactly one match etc.
2. Also, derangement of (1, 2,….n) as a permutation i1, i2,….., in of (1, 2, …,n) such that i1 □ 1, i2 □ 2, …, in □ n that is no object is in its natural position (no match).
3. You also considered cases of exactly one object matches, exactly k object matches etc.
4. And the probabilities for cases considered were also calculated.

## Self-Assessment Questions(SAQs) for Study Session 6

1. Show that $D_5 = 44$

2. Seven gentlemen check their hats at a party. How many different ways can their hats be returned so that:

a) No gentleman receives his own hat?

b) At least one gentleman receives his own hat?

c) At least two gentlemen receive their own hat?

## References

Merris, R. (2003). Combinatorics. Wiley-Interscience.

Rosen, K. H., Michaels, J. G., Gross, J. L., Grossman, J. W. and Shier, D. R. (2000). *Handbook of Discrete and Combinatorial Mathematics*. CRC Press.

Brualdi, R. A. (1999). *Introductory Combinatorics*. Pearson Education Asia Limited and China Machine Press.

Grimaldi, R. P. (1999). *Discrete and Combinatorial Mathematics*. Pearson Addison-Wesley.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Olubusoye, O. E. (2010). *Study sessions notes on STA 311*.

# Study Sessions 7: Probability Generating Functions (PGF)

## Expected duration: 1 week or 2 contact hours

## Introduction

In your 200 level Probability course, you were taught the moment generating functions (MGF) and means and variances of some distributions were derived using the MGF. In this study sessions, while building upon the knowledge of MGF, you will be opened up to Probability Generating Functions (PGF).

## Learning Outcomes for Study Session 7

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

7.1 Define a probability generating function
7.2 State the properties of probability generating functions
7.3 Derive means and variances of some discrete probability distributions

## 7.1 Probability Generating Function (PGF)

The probability generating function (PGF) of a discrete random variable is a power series representation (the generating function) of the probability mass function of a random variable $X$.

PGFs are often employed for their succinct description of the sequence of probability $P[X = i]$ and to make available the well-developed theory of power series with non-negative co-efficients.

Definition 1: The probability generating function (PGF) of a random variable X is defined as:

$$G_x(t) \quad = \quad E\,[t^x]$$

$$= \sum_x t^x \quad P[X = x]$$

where:

$G_x(t)$ is defined only when X take values in the non-negative integers

P(X=$x$) is the probability mass function of X.

The notation $G_x$ is usually used to emphasize the dependence on X.

### 7.1.1 Properties of PGF

1. The probability mass function of X is recovered by taking derivatives of G.

$$P(k) = P(X = k) = \frac{G^{(K)}(0)}{K!}$$

2. If X and Y have identical PGFs, then they are identically distributed. i.e. if there are two random variables X and Y and $G_X = G_Y$, then $f_X = f_Y$.

3. The expectation of X is given by

$$E(X) = G^1(1)$$

Proof:

G(t) = E($t^x$) = $\sum_x t^x P(x)$

$$G'(t) \;=\; \frac{d\,G(t)}{dt}$$

$$= \sum_x \; x t^{x\,-\,1}\; P(x)$$

$$G'_{(1)} \;=\; \sum_x \; x\,P(x) \quad \Rightarrow\; G'_{(1)} = E(x)$$

The variance of X is given by:

$$\text{Var [X]} = G''_{(1)} + G'_{(1)} - [\,G'_{(1)}\,]^2$$

Proof: $G'_{(t)} = \sum_x x\;\, t^{x-1}\; P_{(x)}$

$$G''_{(t)} = \sum_x x(x-1)t^{x-2}\; P_{(x)}$$

$$G''_{(t)} = \sum_x (x^2 - x)\,P_{(x)}\,t^{x-2}$$

$$G''_{(t)} = \left[\sum_x x^2 P_{(x)} - \sum_x xP_{(x)}\right] t^{x-2}$$

$$G''_{(1)} = \sum_x x^2 P_{(x)} - \sum_x xP_{(x)}$$

$G''_{(1)}$ = E($x^2$) - E(x)

$G''_{(1)}$ = E($x^2$) - $G'_{(1)}$

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

$$V(x) = E(X^2) - \left[G_{(1)}{}^{1}\right]^2$$

But $E(x^2) = G''_{(1)} + G'_{(1)}$

Therefore Var [X] = $G''_{(1)} + G'_{(1)} - [G'_{(1)}]^2$

## 7.2 Probability Generating Functions Approach

Probability Generating Functions Approach for deriving means and variances of some discrete distributions:

### 1. Bernoulli Distribution

The probability density function (pdf) of a Bernoulli distribution is given by

$$P[X = x] = P^x q^{1-x}$$

Mean:

$G_x(t) = E[t^x]$

$$= \sum_{x=o}^{1} t^x P[X = x]$$

$$= t^0 p^0 q^{1-0} + t^1 p^1 q^{1-1}$$

$G_x(t) = q + pt$

$G'_{(t)} = p$

$G'_{(1)} = p = E(x)$

Variance:

$$G^{11}(t) = \frac{d^2 G(t)}{dt^2}$$

Therefore, $G^1(t) = p$

$G^{11}(t) = 0$

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

And Var(X) = $G^{11}(1) + G^1{}_{(1)} - [G^1{}_{(1)}]^2$

This implies that $0 + p - p^2 = p(1-p)$

Var(x) = pq

**2. Binomial Distribution**

The p.d.f of a binomial distribution is given by

$$\binom{n}{x} p^x q^{n-x}$$

where:

*n* is the number of observations

*p* is the probability of success

*q* is the probability of failure

*x* is the random variable

Mean:

G($t$) = E[$t^x$]

$$= \sum_{x=o}^{n} t^x P[X = x]$$

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

$$= \sum_{x=o}^{n} (pt)^x \binom{n}{x} q^{n-x}$$

$$= \sum_{x=o}^{n} (pt)^x \binom{n}{x} q^{n-x}$$

$$= \sum_{x=o}^{n} (pt)^x \binom{n}{x} (1-p)^{n-x}$$

$$= [pt + (1-p)]^n$$

$$G(t) = [pt + q]^n$$

$$G^1(t) = \frac{dG(t)}{dt}$$

$$= n[pt + q]^{n-1} p$$

$$G^1(t) = n(pt + q)^{n-1} p$$

$$G^1(1) = n(p + q)^{n-1} p$$

$G^1(1) = np = E(x) = \text{Mean}$

Variance :

$G^{11}(t) = n(n-1)(pt + q)^{n-2} p^2$

$G^{11}(1) = n(n-1)P^2$

$\therefore \text{Var}(x) = G^{11}{}_{(1)} + G^1{}_{(1)} - [G^1{}_{(1)}]^2$

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

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

$= np - np^2$

$= np(1-p)$

$\text{Var}(x) = npq$

**3. Poisson Distribution**

(i) Mean :

$G(t) = E(t^x)$

$$= \sum_{x-0}^{\infty} t^x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda t}$$

$$= e^{-\lambda + \lambda t}$$

$$G(t) = e^{-\lambda(1-t)}$$

$$G^1(t) = \lambda e^{-\lambda + \lambda t}$$

$$G^1(t) = \lambda e^{-\lambda + \lambda}$$

$$= \lambda e^0 = \lambda$$

$$G^{11}(t) = \lambda . \lambda e^{-\lambda + \lambda t}$$

$$= \lambda^2 e^{-\lambda + \lambda t}$$

$$G^{11}(1) = \lambda^2 e^{-\lambda + \lambda} = \lambda^2$$

(ii) Variance :

$\text{Var}(X) = G^{11}(1) + [G^{1}(1) – [G^{1}(1)]^{2}]$

$$= G^{11}(1) + G^{1}(1) – [G^{1}(1)]^{2}$$

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

$$= \lambda$$

## Summary of Study Session 7

In this study sessions, you learnt about:

1. The definition of probability generating function (PGF) as
2. The properties of probability generating functions
3. How to derive means and variances of some discrete probability distributions using the PGF approach.

## Self-Assessment Questions(SAQs)for Study Session 7

Obtain the PGF of:

Geometric distribution

Negative Binomial distribution

## References

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

*www.wikipedia.org*

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

# Study Sessions 8: Bernoulli Trials and Binomial Distribution

## Expected duration: 1 week or 2 contact hours

## Introduction

In the previous study sessions, probability was used in generating function approach to derive the means and variances of some discrete probability distributions. In this study session, you will be studying Bernoulli trials and Binomial distribution.

## Learning Outcomes for Study Session 8

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

8.1 Solve problems on Bernoulli trials
8.2 state the properties of a Binomial distribution
8.3 derive the mean and variance of a binomial distribution
8.4 solve problems on binomial distribution

## 8.1 Bernoulli Trials

In the theory of probability and statistics, a Bernoulli trial (or binomial trial) is a random experiment with exactly two possible outcomes, "success" and "failure", in which the probability of success is the same every time the experiment is conducted. It is named after Jacob Bernoulli, a Swiss mathematician of the 17th century.

The mathematical formalization of the Bernoulli trial is known as the Bernoulli process. This article offers an elementary introduction to the concept, whereas the article on the Bernoulli process offers a more advanced treatment.

### 8.1.1 Bernoulli Random Variable

A random variable $X$, that assumes only the value 0 or 1 is known as a Bernoulli random variable. The values 0, or 1 can be interpreted as events of failure and success respectively in an experiment usually referred to as *Bernoulli Trial.*

Definition 1: A random variable X is defined to have a Bernoulli distribution if the discrete density function of X is given by

$$f(x) = \begin{cases} p^x(1-p)^{1-x} \text{ for } x = 0 \text{ } or \text{ } 1 \\ 0 \quad otherwise \end{cases} = p^x(1-p)^{1-x} \, I_{\{0,\ 1\}}(x)$$

Where the $p$ satisfies $0 \le p \le 1$. $1-p$ is usually denoted by $q$

Theorem 1: If $X$ has a Bernoulli distribution, then

$E(X) = p, Var(X) = pq$

Proof:

$$E(X) = 0.q + 1.p = p$$

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

$$= 0^2.q + 1^2.p - p^2 = pq$$

Bernoulli distribution is a special type of discrete distribution sometimes referred to as Indictor function. This implies that for a given arbitrary probability space $[S, A, P(.)]$, let $A$ belong to $A$ , define the random variable $X$ to be the indicator function of $A$; that is $\chi(w) = I_A(w)$; then $X$ has a Bernoulli distribution with parameter $p = P[X = 1] = P[A]$.

## 8.2 Binomial Distribution

In Bernoulli distribution, there is just one trial that can result in either success or failure. But, in Binomial distribution, we have repeated an independent trials of an experiment with two outcomes resulting in either success or failure, yes or no, etc.

The probability of exactly $x$ successes in $n$ repeated trials is given by:

$$\text{f(x)} = \begin{cases} \binom{n}{x} p^x q^{n-x}; & x = 0, 1, 2, .........., n \\ 0 & elsewhere \end{cases}$$

where p is the probability of success

q = 1-p is the probability of failure

$x$ is the number of successes in repeated trials.

f($x$) is the probability density function (p.d.f).

### 8.2.1 Properties of Binomial Distribution

It has $n$ independent trials

It has constant probability of success $p$ and probability of failure $q = 1 - p$.

There is assigned probability to non-occurrence of events.

Each trial can result in one of only two possible outcomes called success or failure.

### 8.2.2 Mean and Variance of a Binomial Distribution

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

Mean:

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

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

$$= \sum_{x=0}^{n} x \frac{n!}{(n-x)!\,x!} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x \frac{n(n-1)!}{(n-x)!\,x(x-1)!} p^x q^{n-x}$$

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

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

Let s = $x$ – 1, $x$ = s + 1

$$= np \sum_{s=0}^{n-1} \frac{(n-1)!}{(n-s-1)!s!} p^{s} q^{n-s-1}$$

$$= np \sum_{s=0}^{n-1} \binom{n-1}{s} p^{s} q^{n-s-1}$$

$$= np\ (p+q)^{n-1} = np$$

Variance:

$Var(X) = E(X^2) – [E(X)]^2$

$E[X^2] = E[X(X-1)] + E(X)$

$$E[X(X-1)] = \sum x(x-1)\, f(x)$$

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

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

$$= \sum_{x=0}^{n} x(x-1) \frac{n(n-1)(n-2)!}{(n-x)!x(x-1)(x-2)!} p^{2} p^{x=2} q^{n-x}$$

$$= n(n-1)p^{2} \sum_{x=2}^{n-2} \frac{(n-2)!}{(n-x)!(x-2)!} p^{x=2} q^{n-x}$$

Let s = x-2 x = s + 2

$$= n(n-1)\, p^{2} \sum_{s=0}^{n-2} \frac{(n-2)!}{(n-s-2)!s!} p^{s} q^{n-s-2}$$

$$= n(n-1)\,p^2 \sum_{s=0}^{n-2} \binom{n-2}{s} p^s q^{n-s-2}$$

$$E[X(X-1)] = n(n-1)p^2$$

$$\therefore\ E(X^2) = E[X(X-1)] + E(X)$$

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

$$\therefore\ V(X) = E(X^2) - [E(X)]^2$$

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

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

$$= np - np^2$$

$$= np\,(1-p)$$

$$= npq$$

Remark: The binomial distribution reduces to the Bernoulli distribution when $n = 1$

**Example 1**:

It is known that screw produced by a certain company will be defective with probability 0.02 independently of each other. The company sells the screws in packages of 10 and offer a money back guarantee that at most 1 of 10 screws is defective. What proportion of packages sold must the company replace?

Solution

Let X be the number of defective screws thus n = 1=, p = 0.02

Pr (at most one defective) $\equiv 1 - P(X=0) - P(X=1)$

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

$$= 1 - \binom{10}{0}(0.2)^0 (0.8)^{10} - \binom{10}{1}(0.2)^1 (0.8)^9$$

What is the final answer?

**Example 2**:

A communication system consists of *n* components each of which will, independently function with probability p. The total system will be able to operate effectively if at least one-half of its components function.

For what value of p is the 7-component system likely to operate more effectively than a 5-component system.

Solution

A 7-component system will be effective

$$\textit{If}\ \mathrm{P(E_7 > 3) = P(E = 4) + P(E = 5) + (P(E = 6) + P(E = 7)}$$
$$= 1 - \mathrm{P(E \leq 3) = 1 - P(E = 0) - P(E = 1) - P(E = 2) - P(E = 3)}$$
$$= \binom{7}{4} \mathrm{P}^4 \mathrm{q}^3 \binom{7}{5} + \mathrm{p}^5\ \mathrm{q}^5 + \binom{7}{6}\ P^6 \mathrm{q}^1 + \mathrm{P}^7$$

A 5- component will be effective if

$$P(E_5 > 2) = P(E = 3) + P(E = 4) + P(E = 5)$$
$$= \binom{5}{3} P^3 q^2 + \binom{5}{4} P^4 q^1 + P^5$$

The 7-component will be better if

$$P(E_7 > 3) > \mathrm{P}(\mathrm{E}_5 > 2); \text{ for q} = 1 - \text{p}.$$

Complete this

Try for 5 and 3.

**Example 3**:

For what value of K will $P(X = K) \big/ P(X = K - 1)$ be greater or less than 1 if X is a

b (n, p) and 0<P<1.

Solution

$$\frac{P(X = K)}{P(X = K - 1)} = \frac{\begin{pmatrix} n \\ k \end{pmatrix} P(1 - P)^{n-k}}{\begin{pmatrix} n \\ k - 1 \end{pmatrix} P^{k-1}(1 - P)^{n-k+1}}$$

$$= \frac{(\text{n - k - 1})P}{k(1 - P)}$$

$$\therefore P(X = k) \geq P(X = k - 1) \quad \text{iff}$$

$$(\text{n - k} + 1)P \geq k(1 - P)$$

$$\text{i.e. K} \leq (\text{n} + 1)P$$

This implies that for b (n, p), as k goes from 0 to n, P (x=k) first increases monotonically and then decreases monotonically, reaching its largest value when k is the largest integral less than or equal to (n+1) p.

## Summary of Study Session 8

In this study sessions, you have learnt about:

1. The concept of Bernoulli trials. Its probability density function is given by [redacted](p^x (1-p)^(1-x) for x=0 or 1 @@0 otherwise)
2. How to solve problems on Bernoulli trials
3. How to state the properties of a Binomial distribution with density function

## Self-Assessment Questions (SAQs) for Study Session 8

1. An irregular six-faced die is thrown and the expectation that in 10 throws it will give 5 even number is twice the expectation that it will give 4 even numbers. How many times in 10,000 sets of 10 throws would you expect it to give no even numbers?

2. Four dice are thrown and the number of ones in each throw is recorded.

a. what is the probability that 0, 1, 2, 3,4 ones will occur?

b. Calculate the mean and variance of the distribution obtained in (a)

3. A file of data is stored on a magnetic tape with a parity bit stored with each byte (8 bits) making 9 bits in all. The parity bit is set so that the 9 bits add up to an even number. The parity bit allows error to be detected, but not corrected. However, if there are two errors in the 9 bits, then the error will go undetected, three errors will be detected, four errors undetected, etc. A very poor magnetic tape was tested for the reproduction of 1024 bits and 16 errors were found. If on one record on the tape, there are 4000 groups of 9 bits, estimate how many bytes will have undetected errors.

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Amahia, G. N. (2007). *STA 211- Probability II. Ibadan Distance Learning Centre Series.* Distance Learning Centre, University of Ibadan.

Hogg, R. V. and Craig, A. T. (1970). *Introduction to Mathematical Statistics.* New York: Macmillan Publishing Co.

Shittu, O. I. (2011). *Notes on Probability II.*

# Study Session 9: Poisson Distribution

## Expected duration: 1 week or 2 contact hours

## Introduction

In study sessions eight, we discussed the binomial distribution which gives the probability of $x$ successes in $n$ independent trials of an experiment. When n becomes large and p is fairly small, the use of the binomial distribution in calculating the various probabilities becomes cumbersome. To overcome this problem, we use another probability function which approximates the binomial distribution.

This probability function is known as the Poisson probability function which you shall be focusing on in this study session.

## Learning Outcomes for Study Session 9

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

9.1 Describe a Poisson random variable
9.2 List the properties of a Poisson experiment
9.3 Determine the mean and variance of a Poisson distribution
9.4 Derive the Poisson Distribution as an approximation of the binomial distribution

## 9.1 Poisson random variable

A random variable closely related to the binomial random variable is one whose possible values 0, 1, 2, 3,….. represent the number of occurrences of some outcomes not in a given number of trials but in a given period of time or region of space. This variable is called the Poisson variable.

### 9.1.1 Properties of a Poisson Experiment

A Poisson experiment is a statistical experiment that has the following properties:

1. The experiment results in outcomes that can be classified as success or failures.
2. The average number of success($\lambda$) that occurs in a specified region is known.
3. The probability that a success will occur is proportional to the size of the region.
4. The probability that a success will occur in an extremely small region is virtually zero.

| Box 9.1: Note |
| --- |
| The specified region may take many forms e.g. length, an area, a period of time, volume, etc. |

A Poisson random variable is the number of successes that result from a Poisson experiment. The probability distribution of a Poisson random variable is called a Poisson distribution. Given the mean number of successes $\lambda$ that occur in a specified region, the probability density function (pdf) of Poisson distribution is given by

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

where $x$ is the actual number of successes that result from the experiment.

$\lambda$= np ($n$ is the total number of observation in the experiment and $p$ is the probability of success).

| Box 9.2: Note |
| --- |
| Mean $\lambda$ and variance are equal, i.e. $\lambda$ = mean = variance. Also $\lambda$ is the parameter of the distribution, with e= 2. 71828 |

Some examples of random variables that obey the Poisson probability law are:

The number of customers entering a post office on a given day

The number of misprints on a page (or a group of pages) of a book.

The number of packages of instant noodles sold in a particular store on a given day.

## 9.2 Identities

$$e\lambda = \sum_{x=o}^{\infty} \frac{\lambda^x}{x!}$$

$$= 1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+....$$

Using the result, we have

$$\sum_{x=o}^{\infty} P_{(x)} = \sum_{x=o}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \quad e^{\lambda} = 1$$

## 9.3 Mean and variance of a Poisson Distribution

(i) Mean

$$E(X) \quad = \sum_{x=0}^{\infty} xP(x)$$

$$= \sum_{x=0}^{\infty} x \; \frac{e^{-\lambda}\lambda^x}{x!} \qquad \lambda^x = \lambda^{x-1}.\lambda$$

$$= \quad \sum_{x=0}^{\infty} x \; \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \quad \sum_{x=1}^{\infty} x\lambda \; \frac{\lambda^{x-1}e^{-\lambda}}{x(x-1)!}$$

$$= \quad \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}e^{-\lambda}}{(x-1)!}$$

Let s = $x - 1$

$$= \lambda \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^s}{s!}$$

$$= \lambda$$

(ii) Variance

$$\text{Var(X)} = E(X^2) - [E(x)]^2$$

$$E(x^2) = E[x(x-1)] + E(x)$$

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

$$= \lambda^2 \sum_{x-2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} \qquad \lambda^x = \lambda^{x-2}\lambda^2$$

Let $s = x - 2$

$$= \lambda^2 \sum_{s=0}^{\infty} \frac{e^{-\lambda}\lambda^{s}}{s!}$$

$$= \lambda^2$$

$$\therefore \text{Var}(X) = \lambda^2 + \lambda - [\lambda]^2$$

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

$$= \lambda$$

**Example 1**:

The average number of homes sold by Assurance Homes Company is 2 homes per day. What is the probability that exactly 3 homes will be sold tomorrow?

Solution: $\lambda$ = 2 since 2 homes are sold per day on the average

x = 3, e = 2.71828

$$\therefore P(x; \lambda = np) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \frac{2.71828^{-2}\ (2)^3}{3!}$$

$$P(3; 2) = 0.180$$

Thus the probability of selling 3 homes is 0.180

## 9.4 The Poisson Distribution as an Approximation to the Binomial Distribution

Let n and p be the parameters of a binomial distribution.

Therefore mean $\lambda$ = np

Variance $\sigma^2$ = np(1 – p)

If n $\to \infty$ and p $\to$ 0 simultaneously, in such a way that $\lambda$= np is fixed, then we can say that p = $\lambda/n$ where $\lambda$ is a fixed value.

Then as *n* increases, the binomial probabilities.

$P(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x$ = 0, 1, 2,…. Get closer and closer to the Poisson probabilities.

Proof: Given that

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

$$P(x: n, p) = p(x : n, \lambda/n)$$

$$= \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1 - \frac{\lambda}{n}\right)^{n-x} \quad x = 0, 1, 2, \ldots n$$

$$= \frac{n(n-1)\ (n-2)\ldots\ldots(n-x+1)(n-x)!}{x!(n-x)!}\ \frac{\lambda^x}{n^x}\ \left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n(n-1)\,(n-2)....(n-x+1)}{x!\, n^x}\ \lambda^x \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{n(n-1)(n-2).......(n-x+1)}{n^x}\ \frac{\lambda^x}{x!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$$

but $\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \equiv e^{-\lambda}$

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-x} \equiv 1$$

$$\frac{n(n-1)(n-2)....(n-x+1)}{n^x} \equiv 1$$

Therefore, we have $\frac{e^{-\lambda}\lambda^x}{x!}$, x = 0, 1, 2….

Thus, the binomial pdf approaches the Poisson as *n* increases and *p* tends to zero.

## Summary of Study Session 9

In this study sessions, you learnt the following:

1. How to describe a Poisson random variable
2. The properties of a Poisson experiment
3. The probability density function of a Poisson random variable X as
4. How to obtain the mean and variance of a Poisson distribution as λ
5. How to derive the Poisson Distribution as an approximation of the binomial distribution.

## Self-Assessment Questions(SAQs) for Study Session 9

An electronic company manufactures a specific component for an electronic gadget. It finds out that out of every 1000 components produced, 8 are defective. If the components are packed in boxes of 250, find:

The probability of obtaining 0, 1, 2 defectives in the box

The probability that the box contains at least 2 defectives

Samples of 10 items are drawn from a manufacturing process in which the chance of any one item being defective is 0.1. Using Poisson approximation to the binomial, what are the chances of having:

Exactly two defectives

No defectives

Suppose that the average number of lions seen on a 1-day safari is 5. What is the probability that tourists will see fewer than four lions on the next 1-day safari.

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Amahia, G. N. (2007). *STA 211- Probability II. Ibadan Distance Learning Centre Series.* Distance Learning Centre, University of Ibadan.

Hogg, R. V. and Craig, A. T. (1970). *Introduction to Mathematical Statistics*. New York: Macmillan Publishing Co.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Newnes. Elsevier. Linacre House, Jordan Hill, Oxford.

# Study Sessions 10: Hypergeometric Distribution

## Expected duration: 1 week or 2 contact hours

## Introduction

In study sessions nine, we discussed the Binomial distribution where the sampling is with replacement and the probability of occurrence is thus constant from trial to trial. In this study sessions, we shall be considering the hypergeometric distribution which involves sampling without replacement.

## Learning Outcomes for Study Session 10

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

- 10.1 State its probability density function
- 10.2 Derive the mean and variance of the hypergeometric distribution
- 10.3 Binomial distribution as an approximation to the hypergeometric distribution

## 10.1 State its probability density function

Consider a lot consisting of $m + n$ items of which $m$ of them are defective and the remaining $n$ of them are non-defective. A sample of $r$ items is drawn randomly without replacement. Let $x$ denote the number of defective items that is observed in the sample.

The random variable $x$ is the hypergeometric random variable with parameters $m + n$ and $m$. Then, the number of ways selecting $x$ defective items from $m$ defective items is $\binom{m}{x}$; the number of ways of selecting $r - x$ non-defective items from $n$ non-defective items is $\binom{n}{r-x}$. Therefore, total number of ways of selecting $r$ items with $x$ defective and $r - x$ non-defective items is. $\binom{m}{x}\binom{n}{r-x}$

Finally, the number of ways one can select $r$ different items from a collection of $m + n$ different items is $\binom{m+n}{r}$. Thus, the probability of observing $x$ defective items in a sample of $r$ items (probability density function) is

$\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}$ for $x = 0, 1, 2, \ldots.., r, x \leq m$ and $r - x \leq n$

**Example 1**:

In a lottery, a player selects 6 different numbers from 1,2,…,44 by buying a ticket for 1 naira (N1.00). Later in the week, the winning numbers will be drawn randomly by a device. If the player matches all six winning numbers, then he or she will win the jackpot of the week. If the player matches 4 or 5 numbers, he or she will receive a lesser cash prize. If a player buys one ticket, what are the chances of matching: (a) all 6 numbers, (b) 4 numbers.

Solution: Let $x$ denote the number of winning numbers in the ticket. If we regard winning numbers as defective, then $x$ is a hypergeometric random variable with $m + n = 44, m = 6$ and $n$ = 38.

(a) $P(X = 6) = \frac{\binom{6}{6}\binom{38}{0}}{\binom{44}{6}} = \frac{1}{\binom{44}{6}} = \frac{6!38!}{44!} = \frac{1}{7059052}$

(b) $P(X = 4) = \frac{\binom{6}{4}\binom{38}{2}}{\binom{44}{6}} = 0.0014938$

**Example 2**:

As part of a health survey, a researcher decides to investigate prevalence of cholera in 8 sub-urban areas out of a city's 28 sub-urban areas. If 6 of the sub-urban areas have a very high prevalence rate, what is the probability that none of them will be included in the researcher's sample?

Solution:

Recall that we have $f(x) = \frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}$ as the p. d. f for hypergeometric distribution

Here, $x = 0, n = 22, m + n = 28$ $and$ $m = 6$

Then, we have $\dfrac{\binom{6}{0}\binom{28-6}{8-0}}{\binom{28}{8}}$

Complete this

## 10.2 Mean and Variance of Hypergeometric Distribution

Mean:

$$E(x) = \sum_{x=0}^{r} xf(x)$$

$$= \sum_{x=0}^{r} x \frac{m_{C_x} \; n_{C_{r-x}}}{\begin{matrix} m+n \\ C_r \end{matrix}}$$

$$= \sum_{x=0}^{r} x \frac{\frac{m!}{(m-x)!x!} \; {}^{n}C_{r-x}}{{}^{m+n}C_{r}}$$

$$= \sum_{x=1}^{r} \frac{\frac{x \; m(m-1)!}{(m-x)!x(x-1)!} \; {}^{n}C_{r-x}}{{}^{m+n}C_{r}}$$

$$= m \sum_{x=1}^{r} \frac{\frac{(m-1)!}{(m-x)!(x-1)!} \; {}^{n}C_{r-x}}{{}^{m+n}C_{r}}$$

let $= x - 1$ , $x = s + 1$

this implies $$\frac{m\sum_{s=0}^{r-1} \frac{(m-1)!}{(m-s-1)!s!} \quad {}^{n}C_{r-s-1}}{{}^{m+n}C_{r}}$$

$$= \frac{m}{{}^{m+n}C_{r}} \cdot {}^{m-1}C_{s} \quad {}^{n}C_{r-s-1}$$

$$= \frac{m}{{}^{m+n}C_{r}} \quad . \quad {}^{m+n-1}C_{r-1}$$

$$= \frac{\frac{m}{m+n!}}{(m+n-r)!r!} \quad x \quad \frac{(m+n-1)!}{[(m+n-1)-(r-1)]!\,(r-1)!}$$

$$= \frac{\frac{m}{m+n!}}{(m+n-r)!r!} \quad x \quad \frac{(m+n-1)!}{(m+n-r)!(r-1)!}$$

Simplifying gives

$$E(x) \quad = \quad \frac{mr}{m+n}$$

Variance:

$$E(x^2) \quad = \quad E[x(x-1) + x]$$

$$= \quad E[x(x-1)] \; + \; E(x)$$

$$E[x(x-1)] \qquad = \sum x(x-1)\,f(x)$$

$$= \sum_{x=0}^{r} \; x(x-1) \; \frac{m_{C_x} \;\; n_{C_{r-x}}}{\substack{m+n \\ C_r}}$$

Continuing gives

$$= \sum_{x=2}^{r} x(x-1) \frac{\frac{m(m-1)(m-2)!}{(m-x)!\,x(x-1)(x-2)!} \quad {}^{n}C_{r-x}}{{}^{m+n}C_{r}}$$

$$= m(m-1) \sum_{x=2}^{r} \frac{\frac{(m-2)!}{(m-x)!(x-2)!} \quad {}^{n}C_{r-x}}{{}^{m+n}C_{r}}$$

$$= \frac{m(m-1)}{{}^{m+n}C_{r}} \sum_{x=2}^{r} \frac{(m-2)!}{(m-x)!(x-2)!} \quad {}^{n}C_{r-x}$$

let s = x-2 x = s + 2

$$= \frac{m(m-1)}{{}^{m+n}C_{r}} \sum_{s=0}^{r-2} \frac{(m-2)!}{(m-s-2)!\,s!} \quad {}^{n}C_{r-s-2}$$

$$= \frac{m(m-1)}{{}^{m+n}C_{r}} \ x \ {}^{m-2}C_{s} \quad {}^{n}C_{r-s-2}$$

$$= \frac{m(m-1)}{{}^{m+n}C_{r}} \ x \ {}^{m+n-2}C_{r-2}$$

$$= \frac{\frac{m(m-1)}{m+n!}}{(m+n-r)!\,r!} \ x \ \frac{(m+n-2)!}{\left[(m+n-2)-(r-2)\right]!(r-2)!}$$

$$= \frac{m(m-1)(m+n-r)!\,r!}{(m+n)!} \ x \ \frac{(m+n-2)!}{(m+n-r)!(r-2)!}$$

$$= \frac{m(m-1)\,r(r-1)(r-2)!}{(m+n)(m+n-1)(m+n-2)!} \quad \frac{(m+n-2)!}{(m+n-2)!}$$

$$= \frac{rm(m-1)(r-1)}{(m+n)(m+n-1)}$$

$$E(x^2) = E[x(x-1)] + E(x)$$

$$= \frac{m(m-1)\,r(r-1)}{(m+n)(m+n-1)} + \frac{rm}{m+n}$$

Therefore, $V(x) = E(x^2) - [E(x)]^2$

$$= \frac{m(m-1)\,r(r-1)}{(m+n)(m+n-1)} + \frac{rm}{m+n} - \left(\frac{rm}{m+n}\right)^2$$

Simplify this last expression to obtain $\frac{rm}{m+n}\left(1-\frac{m}{m+n}\right)\left(\frac{m+n-r}{m+n-1}\right)$ (Post-Test 2)

Note: If the sampling was with replacement, $r$ and $p = \frac{m}{m+n}$ would be the appropriate binomial parameter and its respective variance would be $r\frac{m}{m+n}\left(1-\frac{m}{m+n}\right)$.

The binomial variance is slightly greater than the hypergeometric variance because of the factor $\left(\frac{m+n-r}{m+n-1}\right)$ in the hypergeometric variance.

As $m+n$ becomes very large compared to $r$, the hypergeometric distribution tends to the binomial distribution.

## 10.3 Binomial distribution as an approximation to the hypergeometric distribution

Suppose the p.d.f of a hypergeometric distribution is given by

$$f(x) = \frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}$$

then, we have the following theorem.

Theorem:

Let m, n $\rightarrow \infty$ and suppose that

$$\frac{m}{m+n} = P_{m,n} \to P,\ 0 < P < 1$$

then $$\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}} \to \binom{r}{x} P^x q^{n-x},\ x = 0,1,2,......,r$$

Proof:

We have

$$\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}$$

$$\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}} = \frac{\frac{m!}{(m-x)!x!} \quad \frac{n!}{[n-(r-x)]!}(r-x)!}{\frac{(m+n)!}{(m+n-r)!r!}}$$

$$= \frac{m!}{(m-x)!x!} \quad \frac{n!}{(n-r+x)!(r-x)!} \quad x \quad \frac{(m+n-r)!r!}{(m+n)!}$$

$$= \binom{r}{x} \frac{m!n!\,(m+n-r)!}{(m-x)!(n-r+x)!(m+n)!}$$

$$= \binom{r}{x} \frac{m(m-1)........(m-x+1)\,(m-x)!\ n(n-1)....(n-r+x+1)\,(n-r+x)!}{(m-x)!(n-r+x)!(m+n)......[(m+n)-(r-1)]}$$

$$= \frac{\binom{r}{x} m(m-1).......[m-(x-1)].\ n(n-1).......[n-(r-x-1)]}{(m+n).........[(m+n)\ -\ (r-1)]}$$

Divide through by m+n

$$\binom{r}{x}\left[\frac{\left(\frac{m}{m+n}\right)\left(\frac{m}{m+n}-\frac{1}{m+n}\right)\cdots\cdots\left(\frac{m}{m+n}-\frac{x-1}{m+n}\right)\times\left(\frac{n}{m+n}\right)\left(\frac{n}{m+n}-\frac{1}{m+n}\right)\cdots\left(\frac{n}{m+n}-\frac{r-x-1}{m+n}\right)}{\left(\frac{m+n}{m+n}\right)\cdots\cdots\left(\frac{m+n}{m+n}-\frac{r-1}{m+n}\right)}\right]$$

$$\Rightarrow\binom{r}{x}\left(\frac{m}{m+n}\right)\left(\frac{m}{m+n}-\frac{1}{m+n}\right)\cdots\left(\frac{m}{m+n}-\frac{x-1}{m+n}\right)\times\left(\frac{n}{m+n}\right)\left(\frac{n}{m+n}-\frac{1}{m+n}\right)\cdots\left(\frac{n}{m+n}-\frac{r-x-1}{m+n}\right)\times\frac{1}{1\cdots\left[1-\frac{(r-1)}{(m+n)}\right]}$$

Since $\frac{m}{m+n}\Rightarrow P,\ \ hence\ \frac{n}{m+n}\Rightarrow 1-p=q$

Therefore

$$\lim_{m+n\to\infty}\binom{r}{x}(p)\left(p-\frac{1}{m+n}\right)\cdots\left(P-\frac{x-1}{m+n}\right)\times(q)\left(q-\frac{1}{m+n}\right)\cdots\left(q-\frac{r-x-1}{m+n}\right)\times\frac{1}{1\cdots\left(1-\frac{r-1}{m+n}\right)}$$

$$=\binom{r}{x}P^{x}q^{r-x}$$

This result implies that we can approximate the probabilities $\frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}$ by $\binom{r}{x}P^{x}q^{r-x}$ by setting p = $\frac{m}{m+n}$ provided m, n are large. This is true for all x = 0, 1, 2, …., r.

If m, n, are large, approximate the hypergeometric distribution by an appropriate binomial distribution. If the need arises, we may also go a step further in approximating the binomial distribution by the appropriate Poisson distribution.

## Summary of Study Session 10

In this study sessions, you have been able to:

1. Describe an hypergeometric experiment giving its probability function as
2. Show that the mean and variance of the hypergeometric distribution are mr/(m+n) and mnr(m+n-r)/((m+n)^2 (m+n-1) ) respectively
3. Approximate the hypergeometric distribution using the binomial distribution
4. Solve problems on hypergeometric distribution

## Self-Assessment Questions (SAQs) for Study Session

1. A shipment of 50 mechanical devices consists of 42 good ones and the rest defective. An inspector selects five devices at random without replacement.

(a) What is the probability that exactly three are good?

(b) What is the probability that at least 3 are good?

2. Simplify $\frac{m(m-1)\,r(r-1)}{(m+n)(m+n-1)} + \frac{rm}{m+n} - \left(\frac{rm}{m+n}\right)^2$

to obtain $\frac{rm}{m+n}\left(1-\frac{m}{m+n}\right)\left(\frac{m+n-r}{m+n-1}\right) = \frac{mnr(m+n-r)}{(m+n)^2(m+n-1)}$

Fourteen applicants apply for the post of a marketer in a company. 8 of them have university degrees. If 5 of these applicants are randomly selected for an interview, what is the probability that 3 of them have college degrees?

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Newnes. Elsevier. Linacre House, Jordan Hill, Oxford.

Odeyinka, J. A. and Oseni, B. A. (2008). *Basic Tools in Statistical Theory*. Highland Publishers.

# Study Sessions 11: Negative Binomial and Geometric Distributions

## Expected duration: 1 week or 2 contact hours

## Introduction

Negative Binomial and Geometric distributions are two families of discrete distributions that are very important in Statistics. The Geometric distribution is so named because the values of the Geometric density are the terms of a geometric series while the Negative binomial distribution is sometimes also referred to as the Pascal's distribution.

## Learning Outcomes for Study Session 11

At the end of this study sessions, students should be able to:

11.1 Describe the negative binomial and geometric distributions

11.2 Establish the relationship between the negative binomial and geometric distributions.

## 11.1 Negative Binomial Distribution

Consider a succession of Bernoulli trials, let $P(r)$ denote the probability that exactly $r + k$ $(k > 0)$, trials are needed to produce $k$ successes. This will so happen when the last trial, that is, $(r + k)$th trials is a success with probability p and the previous $(r + k - 1)$ trials must have $(k - 1)$ successes with probability

$r + k - 1_{C_{k-1}} P^{k-1} q^r$, where $q = 1 - p$

$$P(r) = prob\ of\ (k-1)\ successes\ in\ (x + k - 1) trials \times prob\ of\ (x + k)th\ success$$

$$= r + k - 1_{C_{k-1}} P^{k-1} q^r . p$$

$= r + k - 1_{C_{k-1}} P^k q^r \qquad r = 0, 1, 2 \ldots\ldots$ ……………eqn. (1)

$$= \frac{P^k(k+r-1)(k+r-2)\ldots.[k+r+1-(r+1)]}{r!} q^r$$

$$= \frac{P^k(k+r-1)(k+r-2)\ldots.(k+1)k}{r!} q^r$$

$$= P^k(-1)^r \frac{(-k)(-k-1)\ldots.(-k-r+1)}{r!} q^r$$

$$= P^k(-1)^r {-k}_{C_r} q^r$$

$= {-k}_{C_r} P^k(-q)^r$…………..eqn. (2)

Note that:

$$r+k-1_{C_{k-1}} P^k q^r, \qquad r = 0, 1, 2, \ldots \ldots$$

$$= r+k-1_{C_r} P^k q^r, \qquad r = 0, 1, 2, \ldots \ldots$$

$\sum_0^{\infty} P(r) = P^k \sum_{r=0}^{\infty} {-k}_{C_r} (-q)^r$

$$= P^k[1-q]^{-k}$$

$$= p^k p^{-k} = 1$$

Equations (1) and (2) for $k \geq 0$ are known as negative binomial distribution.

Mean and Variance of the Negative Binomial Distribution

Mean

Recall that the moment generating function (MGF) of a random variable *X*,

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

using the moment generating function approach, therefore, from equation (1), the MGF of *R* is

$$M(t) = E(e^{tr}) = \sum_{r=0}^{\infty} e^{tr} \binom{r+k-1}{r} P^k q^r,$$

But $(1-x)^{-n} = \sum_{j=0}^{\infty} \binom{-n}{j} (-x)^j = \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j$ for -1<x<1

Therefore

$M(t)$=$\sum_{r=0}^{\infty} e^{tr} \binom{-k}{r} P^k (-q)^r,$

=$\sum_{r=0}^{\infty} \binom{-k}{r} P^k (-qe^t)^r$

$$= P^k (1-qe^t)^{-k}$$

Now, $M'(t) = k\, qe^t P^k (1-qe^t)^{-k-1}$

$E(R) = M'(t)|t = 0 = \frac{kq}{p}$

$E(R)^2 = M''(t)$

$$= k\, qe^t P^k (1-qe^t)^{-k-1} + (k+1) qe^t P^k (1-qe^t)^{-k-2} kqe^t$$

Complete the solution using $V(R) = E(R)^2 - (E(R))^2$ (see Post –test 4)

## 11.2 Geometric distribution

If in equation (1), we put $k = 1$, we have

$$r + k - 1_{C_{k-1}} P^k q^r$$

$$= r_{C_0} P \quad q^r$$

$= q^r p, r = 0,1,2, \ldots..$

and $q = 1 - p$, we have geometric distribution.

The following describes the Geometric distribution.

Consider a sequence of Bernoulli trials with probability $p$ of success. This sequence is observed until the first success occurs. Let R denote the number of failures before this first success. For instance, if the sequence starts with F representing failure and S success, with F, F, F, S ……. then R=3, i.e. this distribution describes the event of first success after $n^{th}$ independent trials with probability ρ, $0 < \rho < 1$ .

Moreover, the probability of such a sequence is $P[R=3] = (q)(q)(q)(p) = q^3 p = (1-p)^3 \ p.$

Generally, the p.d.f, f(r) = $P[R = r]$ of R is given by

$$f(r) = (1-p)^r p, \qquad r = 0, 1, 2 \ \ldots....$$

$$f(r) = q^r p, \qquad r = 0,1,2, \ldots \ldots .... \qquad (1)$$

Some authors define the geometric distribution by assuming 1 (instead of 0) is the smallest mass point. The p.d.f then has the form

$$f(r) = \begin{cases} p(1-p)^{r-1} \\ 0, \ \ elsewhere \end{cases}, r = 1, 2, 3 \ldots\ldots\ldots\ldots \qquad (2)$$

Mean and Variance of a geometric distribution

Consider equation (2)

Mean

$$E(R) = \sum_{\boldsymbol{r=1}}^{\infty} r \ p(1-p)^{r-1}$$

let $1 - p = q$

$$\Rightarrow E(R) = \sum_{r=1}^{\infty} r \; p(q)^{r-1}$$

$$= \sum_{r=1}^{\infty} p \frac{d}{dq} \; (q)^r$$

$$= p \frac{d}{dq} \sum_{r=1}^{\infty} (q)^r$$

$$= p \frac{d}{dq} (q + q^2 + q^3 + \cdots)$$

But $(q + q^2 + q^3 + \cdots) = q(1 + q + q^2 + q^3 + \cdots)$

$$= q \left( \frac{1}{1-q} \right)$$

Therefore, $E(R) = \; p \frac{d}{dq} \left( \frac{q}{1-q} \right)$

$$= p \left( \frac{(1-q)(1) - q(-1)}{(1-q)^2} \right)$$

$$= p \left( \frac{(1 - q + q)}{(1-q)^2} \right)$$

$$= p \left( \frac{(1)}{(p)^2} \right)$$

$$E(R) = \frac{1}{p}$$

Variance

$$E(R^2) = \sum_{r=1}^{\infty} r^2 \; p(1-p)^{r-1}$$

$$= \sum_{r=1}^{\infty} r^2 \; p(q)^{r-1}$$

$$= p\sum_{r=1}^{\infty}\frac{d}{dq}(rq^r)$$

$$= p\frac{d}{dq}\sum_{r=1}^{\infty}(rq^r)$$

$$= p\frac{d}{dq}(q + 2q^2 + 3q^3 + \cdots)$$

$$= p\frac{d}{dq}q(1 + 2q + 3q^2 + 4q^3 + \cdots)$$

Recall that $1 + 2x + 3x^2 + 4x^3 + \cdots = \frac{1}{(1-x)^2}$

Therefore, we have $p\frac{d}{dq}q\left(\frac{1}{(1-q)^2}\right)$

$$= p\left[\frac{(1-q)^2(1) + 2q(1-q)}{(1-q)^4}\right]$$

$$= p\left[\frac{[(1-q)][(1-q+2q)]}{(1-q)^4}\right]$$

$$= p\left[\frac{1+q}{(1-q)^3}\right]$$

$$= p\left[\frac{1+q}{(1-q)^3}\right]$$

$$= \left[\frac{1+q}{(p)^2}\right]$$

$E(R^2) = \left[\frac{2-p}{(p)^2}\right]$ since $q = 1 - p$

Therefore, $V(R) = E(R)^2 - \left(E(R)\right)^2$

$$= \left[\frac{2-p}{(p)^2}\right] - \left(\frac{1}{p}\right)^2$$

$$= \frac{1-p}{p^2}$$

Thus, the mean and variance of this form geometric distribution are $\frac{1}{p}$ and $\frac{1-p}{p^2}$ respectively.

Example 1: A fair die is cast on successive independent trials until second six is observed. What is the probability of observing exactly 10 non-sixes before the second six is cast.

Solution: This is a negative binomial distribution problem. So,

$$\binom{r+k-1}{k-1} p^k (1-p)^r \qquad r = 0, 1, 2 \ldots \ldots$$

Therefore, we have $\binom{10+2-1}{1}\left(\frac{1}{6}\right)^2 \left(5/6\right)^{10} = 0.049$

Example 2:

Team A plays team B in a seven game with series. That is the series is over when either of the teams wins four games. For each game, $\rho(A \text{ wins}) = 0.6$ and the games are assumed to be independent. What is the probability that the series will end in exactly six games.

Solution:

The game will end is either A or B wins the game series.

$$\begin{aligned}
\rho(\text{game ends}) &= \rho(\text{A wins series in 6 games}) + \rho(\text{B wins series in 6 games}) \\
&= \binom{5}{3}(0.6)^4(0.4)^2 + \binom{5}{3}(0.4)^6(0.4)^2 \\
&= 0.207 + 0.092 \\
&= 0.299
\end{aligned}$$

Note: that

$$\begin{aligned}
\rho(\text{A wins series in 6 games}) &= \rho[\text{A looses 2 games before 4 wins}] \\
&= \rho(\text{Y} = 2) \\
&= \binom{5}{3}(0.6)^4(0.4)^2 \\
&= 0.207
\end{aligned}$$

Example 3:

In a sequence of independent rolls of a fair die;

i. What is the probability that the first four is observed in the sixth trial?

Solution: This is geometric distribution problem

P(R = 5) = $\left(\frac{5}{6}\right)^5\left(\frac{1}{6}\right)$= 0.067 where R denotes the number of non-fours before the occurrence of the first four.

What is the probability that at least six trials are required to observe a four?

Solution: P$[R \geq 5] = 1 - P[R \leq 4]$

$$= 1 - [P[R=0] + P[R=1] + P[R=2]] + P[R=3 + P[R=4]]$$

$$P[R=0] = \left(\frac{5}{6}\right)^0\left(\frac{1}{6}\right) = \frac{1}{6}$$

$$P[R=1] = \left(\frac{5}{6}\right)^1\left(\frac{1}{6}\right) = \frac{5}{36}$$

$$P[R=2] = \left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right) = \frac{25}{216}$$

$$P[R=3] = \left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right) = \frac{125}{1296}$$

$$P[R=4] = \left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right) = \frac{625}{7776}$$

$$= 1 - \frac{4651}{7776}$$

Complete the solution

## Summary of Study Session 11

In this study sessions, we have studied

1. The negative binomial and the geometric distribution.
2. The probability distribution of negative binomial distribution is given by 〖=r+k-1〗 _(C_(k-1) ) P^k q^r
3. The geometric distribution is a special case of the negative binomial distribution which holds when k=1 and is q^r p, r=0,1,2,…..

## Self-Assessment Questions(SAQs) for study Session 11

The probability of a successful missile launch is 0.9. Test launches are conducted until three successful launches are achieved. What is the probability of each of the following:

Exactly six launches will be required.

Fewer than six launches will be required

At least four launches will be required.

Using the information in example (3) above, what is the probability that at most five trials are required to observe a four.

A man pays ₦1.00 a throw to win a ₦3.00 gift. His probability of wining on each throw is 0.1.

(i) What is the probability that two throws will be required to win a gift?

(ii) What is the probability that more than three throws will be required to win a gift?

(iii) What is the expected number of throw needed to win a gift?

Show that the variance of the negative binomial distribution is $\frac{kq}{p^2}$ .

Show that the mean and variance of the geometric distribution of the form in equation (1) are $\frac{q}{p}$ and $\frac{q}{p^2}$

## References

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics*. McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Newnes. Elsevier. Linacre House, Jordan Hill, Oxford.

Odeyinka, J. A. and Oseni, B. A. (2008). *Basic Tools in Statistical Theory*. Highland Publishers.

Shittu, O. I. (2011). *Notes on Probability II.*

# Study Sessions 12: Multinomial Distribution

## Expected duration: 1 week or 2 contact hours

## Introduction

Considering the previous study sessions on binomial distribution, the conclusion that you are aware that each trial of a binomial experiment can result in two and only two possible outcomes is not a misfire. In the multinomial experiment, however, each trial can have two or more possible outcomes. So, a binomial experiment is a special case of a multinomial experiment.

## Learning Outcomes for Study Session 12

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

12.1 Describe and solve problems using multinomial experiment

## 12.1 Multinomial experiment

A multinomial experiment is a statistical experiment that has the following properties:

The experiment consists of $n$ repeated trials

Each trial has a discrete number of possible outcomes

The probability that a particular outcome will occur is constant for any given trial

The trials are independent

A multinomial distribution is the probability distribution of outcomes from a multinomial experiment.

Definition: Suppose a multinomial experiment consists of $n$ trials, and each trial can result in any of $k$ possible outcomes $E_1, E_2, E_3, \ldots\ldots, E_k$. Suppose, also, that each possible outcome can occur with probabilities $p_1, p_2, p_3, \ldots\ldots, p_k$. Then, the probability $p$ that $E_1$ occurs $n_1$ times, $E_2$ occurs $n_2$ times,......., and $E_k$ occurs $n_k$ times is

$$p = \left[\frac{n!}{(n_1!n_2!\ldots\ldots n_k!)}\right]\left[{p_1}^{n_1}\,{p_2}^{n_2}\ldots\ldots\,{p_k}^{n_k}\right] \text{ where } n = n_1 + n_2 + n_3 + \cdots + n_k$$

Example 1:

A bowl consists of 2 red marbles, 3 green marbles and 5 blue marbles. 4 marbles are randomly selected from the bowl with replacement. What is the probability of selecting 2 green marbles and 2 blue marbles?

Solution:

The experiment consists of 4 trials, so $n = 4$.

The 4 trials produce 0 red marbles, 2 green marbles and 2 blue marbles; so

$n_{red} = 0, n_{green} = 2, n_{blue} = 2$

On any particular trial, the probability of drawing a red, green or blue marble is 0.2, 0.3 and 0.5 respectively.

Using the multinomial formula, we have

$$p = \left[\frac{n!}{(n_1!\, n_2! \ldots\ldots n_k!)}\right] [{p_1}^{n_1}\, {p_2}^{n_2} \ldots\ldots {p_k}^{n_k}]$$

$$\left[\frac{4!}{(0!\, 2!\, 2!)}\right] [(0.2)^0 (0.3)^2 (0.5)^2]$$

Therefore $p = 0.135$.

Example 2:

Suppose a card is drawn randomly from an ordinary deck of playing cards and then put back in the deck. This exercise is repeated five times. What is the probability of drawing 1 spade, 1 heart, 1 diamond and 2 clubs?

Solution:

The experiment consists of 5 trials, n=5

The 5 trials produce 1 spade, 1 heart, 1 diamond and 2 clubs; so $n_1 = 1, n_2 = 1, n_3 = 1, n_4 = 2$

On any particular trial, the probability of drawing a spade, heart, diamond or club is 0.25, 0.25, 0.25 and 0.25 respectively. Thus, $p_1 = 0.25,\ p_2 = 0.25,\ p_3 = 0.25, p_4 = 0.25$

Using the multinomial formula, we have

$$p = \left[\frac{n!}{(n_1!\, n_2! \ldots\ldots n_k!)}\right][p_1{}^{n_1}\, p_2{}^{n_2} \ldots\ldots\, p_k{}^{n_k}]$$

$$\left[\frac{5!}{(1!\,1!\,1!\,2!)}\right][(0.25)^1(0.25)^1(0.25)^1(0.25)^2]$$

$$p = 0.05859$$

## Summary of Study Session 12

In this study session, you have learnt the following:

How to describe a multinomial experiment stating its properties.

Suppose a multinomial experiment consists of n trials, and each trial can result in any of k possible outcomes E_1,E_2,E_3,…..,E_k. Suppose, also, that each possible outcome can occur with the probabilities p_1,p_2,p_3,…..,p_k. Then, the probability p that E_1 occurs n_1 times, E_2 occurs n_2 times,......., and E_k occurs n_k times is

p=[n!/((n_1 !n_2 !…..n_k !) )][〖p_1〗^(n_1 ) 〖〖 p〗_2〗^(n_2 )…..〖〖 p〗_k〗^(n_k ) ] where n=n_1+n_2+n_3+⋯+n_k

## Self-Assessment Question (SAQs) for Study Session 12

Suppose that a fair die is rolled 9 times. Find the probability that 1 appears 3 times, 2 and 3 twice each, 4 and 5 once each.

In a city on a particular night, television channels 4, 3 and 1 have the following audiences: channel 4 has 25 percent of the viewing audience, channel 3 has 20 percent of the viewing audience and channel 1 has 50 percent of the viewing audience. Find the probability that among ten television viewers randomly chosen in that city on that particular night, 4 will be watching channel 4, 3 will be watching channel 3 and 1 will be watching channel 1.

## References

Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Roussas, G. G. (1973). *A First Course in Mathematical Statistics*. Addison-Wesley Publishing Company.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

Odeyinka, J. A. and Oseni, B. A. (2008). *Basic Tools in Statistical Theory.* Highland Publishers.

# Study Sessions 13: Poisson Process

## Expected duration: 1 week or 2 contact hours

## Introduction

From our previous discussion on Poisson distribution, Poisson random variables represent the number of occurrences of some outcomes not in a given number of trials but in a given period of time, region, length or space. Also, Poisson distribution provides a realistic model for many random phenomena, e.g. the number of telephone calls per hour coming into the switchboard of a large company, etc.

## Learning Outcome for Study Session 13

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

13.1 Define a counting process
13.2 Define a Poisson process
13.3 Solve problems on Poisson process

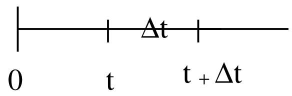
## 13.1 The Counting Process

Definition 1: A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if N(t) represents the total number of "events" that have occurred up to time t.

Example 1: Let N(t) denote the number of goals that a given footballer has scored by time *t*, then $\{N(t), t \geq 0\}$ is a counting process.

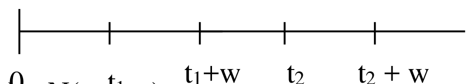
Definition 2:

Independent Increments: A counting process is said to possess independent increment if the number of events which occur in disjoint interval are independent.




Definition 3:

Stationary Increments: A counting process is said to possess stationary increments if the number of events in the interval ($t_1$ + w, $t_2$ + w), i.e. $\left[N(t_2+w) - N(t_1+w)\right]$ has the same distribution as the number of events in the interval ($t_1$, $t_2$) $\left[i.e\ N(t_2) - N(t_1)\right]$ for all $t_1 < t_2$ and w > 0.




$N(t_2)$- $N(t_1)$, $N(t_2+w)$ – $N(t_1 + w)$

### 13.1.1 Properties of a Counting Process N(t)

$N(t) \geq 0$

N(t) is integer-valued.

if there exists a time $w \leq t$, then $N(w) \leq N(t)$.

for w< t, N(t) – N(w) equals the number of events that have occurred in the interval (w, t)

## 13.2 Poisson Process

A Poisson process is a stochastic process in which events occur continuously and independent of one another. Examples of situations that are well modelled as Poisson processes include the radioactive decay of atoms, telephone calls arriving at a switchboard, page view requests to a website and rainfall.

The Poisson process is a collection {N(t): t ≥ 0} of random variables where N(t) is the number of events that have occurred up to time t (starting from time 0). The number of events between time $a$ and time $b$ is given as $N(b) - N(a)$ and has a Poisson distribution.

The Poisson process is a continuous time process. It discrete-time counterpart is the Bernoulli process.

Definition 4: The Poisson process is a continuous-time counting process $\{N(t), t \geq 0\}$ that possesses the following properties:

1. N (0) = 0

2. The number of occurrences counted in disjoint interval are independent from each other (that is independent increments)

3. The probability distribution of the occurrences counted in any time interval only depends on the length of the interval (i.e. stationary increments)

4. No counted occurrences are simultaneous.

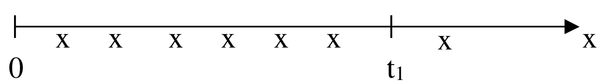
The consequences of this definition include:

The probability distribution of N (t) is a Poisson distribution.

The probability distribution of the waiting time until the next occurrence is an exponential distribution.

Assumptions :

Consider this sketch




Let × denote the occurrence of a happening.

Seven happenings occurred between time 0 and time $t_1$.

Assume that there exists a positive quantity λ which satisfies the following:

1. The probability that exactly one happening will occur in a small time interval of length Δt is approximately equal to λΔt i.e.

P(one happening in interval of length Δt) = λΔt + 0 (Δt)

Note: 0(Δt) [some function of smaller order than Δt] denotes an unspecified function which satisfies $\lim_{\Delta t \to 0} \frac{0(\Delta t)}{\Delta t} = 0$

where λ is the mean rate at which happening occurs per unit of the [mean rate of occurrence]

2. The probability of more than one happening in a small time interval of length Δt is negligible when compared to the probability of just one happening in the same time interval, i.e. P[2 or more happening in intervals of length Δt]= 0(Δt).

3. The numbers of happening in non-overlapping time intervals are independent

as Δt → 0, if the assumptions are valid, then the distribution of N(t) is Poisson.

Therefore, a counting process {N(t), t ≥ 0} is said to be a Poisson process having rate λ, λ > 0, if:

N(0) = 0, i.e. the process begins at time 0

The numbers of events that occur in disjoint time intervals are independent (i.e. independent increment).

The distribution of the number of events that occur in a given interval depends only on the length of that interval and not on its location (stationary increment).

P {N(Δt) = 1}= λΔt + 0 (Δt) or P[N (t + Δt) –N(t) = 1]= λΔt + 0Δt

P[N(Δt) ≥ 2] = 0 (Δt) or P[N(t + Δt) – N (t) ≥ 2] = 0 (Δt)

If conditions (i) – (v) hold, then for t ≥ 0.

$$P_n(t) = P\,[N(t) = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

Proof

Method I:

Divide the interval (o, t) into, say n time subintervals, each of length Δt = $t/n$. The probability that $k$ happenings occur in the interval (o, t) is approximately equal to the probability that exactly one happening has occurred in each of $k$ of the $n$ subintervals that the interval $(o, t)$ is divided into.

Now, the probability of a happening or "success", in a given subinterval is λΔt. Each subinterval provides us with a Bernoulli trial; either the subinterval has a happening or it does not. Also, on view of the assumptions made, these Bernoulli trials are independent, repeated Bernoulli trials; hence, the probability of exactly k "successes" in the n trials is given by

$$\binom{n}{k} (\lambda\Delta t)^k \ (1 - \lambda\Delta t)^{n-k}$$

$$= \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

which is an approximation to the desired probability that k happenings will occur in time interval (0, t).

To get an exact expression,

$$\binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

$$= \frac{n!}{(n-k)!k!} \frac{(\lambda t)^k}{n^k} \left(1 - \frac{\lambda t}{n}\right)^{-k} \left(1 - \frac{\lambda t}{n}\right)^{n}$$

$$= \frac{n(n-1)(n-2).......(n-k+1)(n-k)!}{(n-k)!k!} \frac{(\lambda t)^k}{n^k} \left(1 - \frac{\lambda t}{n}\right)^{-k} \left(1 - \frac{\lambda t}{n}\right)^{n}$$

$$= \frac{n(n-1)(n-2)........(n-k+1)}{n^k} \frac{(\lambda t)^k}{k!} \left(1 - \frac{\lambda t}{n}\right)^{-k} \left(1 - \frac{\lambda t}{n}\right)^{n}$$

but $\lim_{n \to \infty} \left(1 - \frac{\lambda t}{n}\right)^{n} \equiv e^{-\lambda t}$

$$\lim_{n \to \infty} \left(1 - \frac{\lambda t}{n}\right)^{-k} \equiv 1$$

$$\lim_{n \to \infty} \frac{n(n-1)(n-2)........(n-k+1)}{n^k} \equiv 1$$

So that we have $\frac{(\lambda t)^k e^{-\lambda t}}{k!}$ $\quad$ *for* $k = 0, 1, 2$ …..

Method II:

Let t be a point in time after time 0

So, the time interval [0, t] has length t

The time interval [t, t + $\Delta t$] has length $\Delta t$.

Let $P_n(s) = P[Z_{(s)} = n]$

$= P$[exactly n happenings in an interval of length s]

Then, $P_o(t + \Delta t)$ = p[no happenings in interval (0, t + $\Delta t$)]

= p [no happening in (0, t) and no happening in (t, t + $\Delta t$)]

= p [no happening in (0, t)]. P[no happening in (t, t + $\Delta t$)]

$= P_o(t)\, P_o(\Delta t)$ (using the independence assumption)

but P[no happening in (t, t + $\Delta t$)] = 1 – P[one or more happenings in (t, t + $\Delta t$)]

= 1 – P[one happening in (t, t + $\Delta t$)] – P[more than one happening in (t, t + $\Delta t$)]

$= 1 - [\lambda \Delta t + 0(\Delta t)] - 0\,(\Delta t)$

$= 1 - \lambda \Delta t - 0(\Delta t) - 0(\Delta t)$

$\therefore P_o(t + \Delta t) = P_o(t)\, P_o(\Delta t)$

$= P_o(t)\,[1 - \lambda \Delta t - 0(\Delta t) - 0(\Delta t)]$

$P_o(t + \Delta t) = P_o(t) - P_o(t)\, \lambda \Delta t - 0(\Delta t)\, P_o(t) - 0(\Delta t) P_o(t)$

This implies that

$P_o(t + \Delta t) - P_o(t) = - P_o(t)\, \lambda \Delta t - P_o(t)\,[0(\Delta t) + 0(\Delta t)]$

Divide both side by $\Delta t$

$$\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = \frac{-P_0(t)\lambda\Delta t}{\Delta t} - \frac{P_0(t)[0\Delta t)+0(\Delta t)]}{\Delta t}$$

$$\frac{P_0(t+\Delta t)-P_0(t)}{\Delta t} = -\lambda P_0(t) - P_o(t)\left[\frac{0(\Delta t) + 0(\Delta t)}{\Delta t}\right]$$

Recall : $\frac{dP_n(t)}{dt} = \lim_{\Delta t \to 0} \frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = P_n^1(t)$ (differentiation from first principle)

If we take limit as Δt →0, we have

$$\lim_{\Delta t \to 0} \frac{P_o(t+\Delta t) - P_o(t)}{\Delta t} = -\lambda P_o(t)$$

$$= \frac{dP_o(t)}{dt}$$

$$= P_o^1(t)$$

i.e. $P_o^1(t) = -\lambda P_o(t)$

the solution of this differential equation $P_o^1(t) = -\lambda P_o(t)$

$\Rightarrow P_o(t) = Ce^{-\lambda t}$

Let $P_o(t) = P_o(0) = 1$

this implies that C = 1

so, $P_o(t) = e^{-\lambda t}$

Similarly, $P_1(t+\Delta t) = P_1(t)P_o(\Delta t) + P_o(t)P_1(\Delta t)$

$$\Rightarrow P_1(t+\Delta t) = P_1(t)[1 - \lambda\Delta t - 0(\Delta t)] + P_o(t)[\lambda\Delta t + 0(\Delta t)]$$

$$= P_1(t) - \lambda P_1(t)\Delta t - P_1(t)0(\Delta t) + P_o(t)\lambda\Delta t + P_o(t)0(\Delta t)$$

divide through by $\Delta t$

$$\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t} = \frac{-\lambda P_1(t)\Delta t}{\Delta t} - \frac{P_1(t)0(\Delta t)}{\Delta t} + \frac{P_o(t)\lambda\Delta t}{\Delta t} + \frac{P_o(t)0(\Delta t)}{\Delta t}$$

$$\lim_{\Delta t \to 0} \frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = -\lambda P_1(t) + \lambda P_o(t)$$

$$\therefore\ P_1^1(t) = -\lambda P_1(t) + \lambda P_o(t)$$

Using $P_1(0) = 0$,

$$P_1(t) = \lambda t e^{-\lambda t}$$

Continuing in a similar way, we have

$$P_n^1(t) = -\lambda P_n(t) + \lambda P_{n-1}(t)$$

This system of differential equations is satisfied by

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$ using mathematical induction.

Thus, $N(t) \sim$ Poisson$(\lambda t)$ and $\mu = E[N(t)] = \lambda t$

Example 1 :

Suppose that the average number of telephone calls arriving at the switchboard of a small corporation is 30 calls per hour.

What is the probability that no calls will arrive in a 3-minute period?

What is the probability that more than 5 calls will arrive in a 5-minute interval?

Solution: Assume that the number of calls arriving during any time period has a Poisson distribution.

Assume that time is measured in minutes

30 calls per hour ≡ 0.5 calls per minute.

$\therefore$ mean rate of occurrence $\lambda$ = 0.5 per minute.

$\therefore$ (i) P [no call in a 3 minute period] = $\dfrac{e^{-\lambda t}(\lambda t)^0}{0!}$

$$= e^{-(0.5)(3)}(0.5x3)^o = e^{-1.5} = 0.223$$

(ii) P[more than 5 calls in a 5-minute interval] = $\sum_{n=6}^{\infty} \dfrac{e^{-\lambda t}(\lambda t)^n}{n!}$

$$= \sum_{n=6}^{\infty} \frac{e^{-(0.5)(5)}(2.5)^n}{n!}$$

= 1 – P[n ≤ 5]

= 1 – [P(0) + P(1) + P(2) + P(3) + P(4) + P(5)]

= 0.042

since P(0) = 0.223

$$P(1) = \frac{e^{(0.5)(3)}(0.5x3)^1}{1!} = e^{-1.5}(1.5) =$$

$$P(2) = \frac{e^{-(0.5)(3)}(0.5x3)^2}{2!} = \frac{e^{-1.5}(1.5)^2}{2} =$$

$$P(3) = \frac{e^{-(0.5)(3)}(0.5x3)^3}{3!} = \frac{e^{-1.5}(1.5)^3}{6} =$$

$$P(4) = \frac{e^{-(0.5)(3)}(0.5x3)^4}{4!} = \frac{e^{-1.5}(1.5)^4}{24} =$$

$$P(5) = \frac{e^{-(0.5)(3)}(0.5x3)^5}{5!} = \frac{e^{-1.5}(1.5)^5}{120} =$$

Complete the solution

Example 2:

People enter a gambling casino at a rate of 1 for every 2 minutes.

(a) What is the probability that no one enters between 12:00 and 12:05?

(b) What is the probability that at least 4 people enter the casino during that time?

Solution:

(a) $\lambda = 1$ per 2 minutes

$\lambda = 2.5$ per 5 minutes $[\lambda t = \frac{1}{2}(5)]$

so $P[N(5) = 0] = \dfrac{e^{-\lambda t}(\lambda t)^o}{0!}$

$= e^{-2.5} =$

(b) $P[N(5) \geq 4] = 1 - P[n(5) < 4]$

$= 1 - [P(N(5)=0] + P[N(5)=1] + P[N(5)=2] + P[N(5)=3\ ]$

$$= 1 - \left[\frac{e^{-2.5}(2.5)^o}{0!} + \frac{e^{-2.5}(2.5)}{1!} + \frac{e^{-2.5}(2.5)^2}{2!} + \frac{e^{-2.5}(2.5)^3}{3!}\right]$$

$$= 1 - e^{-2.5}\left[1 + \tfrac{5}{2} + \frac{25}{4x2} + \frac{125}{8x6}\right] =$$

Complete the solution

## Summary of Study Session 13

In this study sessions, we have:

i. Defined a counting process
ii. Defined a Poisson process stating its properties
iii. Shown that $P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$

where mean, $\mu = E[N(t)] = \lambda t$

## Self-Assessment Question(SAQs) for Study Session 13

1. Telephone calls are being placed through a certain exchange at random times on the average of 4 per minute. Assuming a Poisson law, determine the probability that in a 15-second interval, there are 3 or more calls.
2. Suppose that flaws on plywood occur at random with an average of one flaw per 50 square feet. What is the probability that a 4-foot x 8-foot sheet will:

i) have no flaws

ii) at most one flaw?

3. Weak spots occur in a certain manufactured tape on the average of 1 per 1000ft. Assuming a Poisson distribution of the number of weak spots on a given length of tape, what is the probability that:

(a) a 2400 –ft roll will have at most 2 defects?
(b) a 1200-ft roll will have no defect?

## References


Mood, A. M., Graybill, F. A. and Boes, D. C. (1974). *Introduction to theory of Statistics.* McGraw-Hill Inc.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Chapman& Hall/CRC.

Attenborough, M. (2003). *Mathematics for Electrical Engineering and Computing*. Newnes. Elsevier. Linacre House, Jordan Hill, Oxford.

Olubusoye, O. E. (2010). *Study sessions notes on STA 311*.