STA 221: PROBABILITY DISTRIBUTION 1

University of Ilorin
Centre for Open & Distance Learning

CODL

Published by the Centre for Open and Distance Learning,
University of Ilorin, Nigeria

E-mail: codl@unilorin.edu.ng
Website: https://codl.unilorin.edu.ng

Foreword

Courseware remains the nerve centre of Open and Distance Learning. Whereas some institutions and tutors depend entirely on Open Educational Resources (OER), CODL at the University of Ilorin considers it necessary to develop its own materials. Rich as OERs are and widely as they are deployed for supporting online education, adding to them in content and quality by individuals and institutions guarantees progress. Doing it in-house as we have done at the University of Ilorin has brought the best out of the Course Development Team across Faculties in the University. Credit must be given to the team for prompt completion and delivery of assigned tasks in spite of their very busy schedules.

The development of the courseware is similar in many ways to the experience of a pregnant woman eagerly looking forward to the D-day when she will put to bed. It is customary that families waiting for the arrival of a new baby usually do so with high hopes. This is the apt description of the eagerness of the University of Ilorin in seeing that the centre for open and distance learning [CODL] takes off.

The Vice-Chancellor, Prof. Sulyman Age Abdulkareem, deserves every accolade for committing huge financial and material resources to the centre. This commitment, no doubt, boosted the efforts of the team. Careful attention to quality standards, ODL compliance and UNILORIN CODL House Style brought the best out from the course development team. Responses to quality assurance with respect to writing, subject matter content, language and instructional design by authors, reviewers, editors and designers, though painstaking, have yielded the course materials now made available primarily to CODL students as open resources.

Aiming at a parity of standards and esteem with regular university programmes is usually an expectation from students on open and distance education programmes. The reason being that stakeholders hold the view that graduates of face-to-face teaching and learning are superior to those exposed to online education. CODL has the dual-mode mandate. This implies a combination of face-to-face with open and distance education. It is in the light of this that our centre has developed its courseware to combine the strength of both modes to bring out the best from the students. CODL students, other categories of students of the University of Ilorin and similar institutions will find the courseware to be their most dependable companion for the acquisition of knowledge, skills and competences in their respective courses and programmes.

Activities, assessments, assignments, exercises, reports, discussions and projects amongst others at various points in the courseware are targeted at achieving the objectives of teaching and learning. The courseware is interactive and directly points the attention of students and users to key issues helpful to their particular learning. Students' understanding has been viewed as a necessary ingredient at every point. Each course has also been broken into modules and their component units in sequential order.

At this juncture, I must commend past directors of this great centre for their painstaking efforts at ensuring that it sees the light of the day. Prof. M. O. Yusuf, Prof. A. A. Fajonyomi and Prof. H. O. Owolabi shall always be remembered for doing their best during their respective tenures. May God continually be pleased with them, Amen.

Bashiru, A. Omipidan
Director, CODL

Course Guide

Probability Distribution I is 3 – credit first semester course. It is available to students offering Bachelor of Science, B. Sc., Statistics, Mathematics, Computer Science and Information Systems degrees.

The mathematical theory of probability gives us the basic tools for constructing and analyzing mathematical models for random phenomena. In studying a random phenomenon, we are dealing with an experiment of which the outcome is not predictable in advance. Experiments of this type that immediately come to mind are those arising in games of chance.

Sometimes the nature of the data suggests the form of the probability model that is assumed. For instance, suppose that an engineer wants to find out what proportion of computer chips, produced by a new method, will be defective, he might select a group of these chips with the resulting data being the number of defective chips in this group. Provided that the chips were "randomly" chosen, it is reasonable to suppose that each one of them is defective with probability p, where p is the unknown proportion of all the chips produced by the new method that are defective.

Course Goal

The goal of this course is to expose students to the knowledge of probability models in the presence of variations that we often encounter in our daily endeavors. When we incorporate the variations into our thinking and analyses, we can make informed judgments from our results that are not invalidated by the variation.

The course covers description of different sample spaces from random experiments, description of different types of events and combinations of events from the defined sample spaces. It also describes the Venn diagrams representation of random experiments, explains the axioms of probability and discusses the addition and multiplication rules of probability. Conditional probability, Prior and posterior probability of events are discussed and its applications through Bayes theorem are demonstrated. Discrete and continuous density functions. Cumulative distribution functions. Mean, Variance, and higher-moments. Chebyshev's inequality. Binomial, Poison, Uniform and Normal distributions are also covered in the course.

Course Development Team

- Content Authoring
  Dr. Alfred Abiodun

- Content Editor
  Mr. Mahmud Abdulwahab

- Instructional Design
  Olawale S. Koledafe

  Olorunfemi Ayodele Sunday

  Damilola Adesodun

  Eyibayo T. Owoeye

From the Vice Chancellor

Courseware development for instructional use by the Centre for Open and Distance Learning (CODL) has been achieved through the dedication of authors and the team involved in quality assurance based on the core values of the University of Ilorin. The availability, relevance and use of the courseware cannot be timelier than now that the whole world has to bring online education to the front burner. A necessary equipping for addressing some of the weaknesses of regular classroom teaching and learning has thus been achieved in this effort.

This basic course material is available in different electronic modes to ease access and use for the students. They are available on the University's website for download to students and others who have interest in learning from the contents. This is UNILORIN CODL's way of extending knowledge and promoting skills acquisition as open source to those who are interested. As expected, graduates of the University of Ilorin are equipped with requisite skills and competencies for excellence in life. That same expectation applies to all users of these learning materials.

Needless to say, that availability and delivery of the courseware to achieve expected CODL goals are of essence. Ultimate attention is paid to quality and excellence in these complementary processes of teaching and learning. Students are confident that they have the best available to them in every sense.

It is hoped that students will make the best use of these valuable course materials.

Professor S. A. Abdulkareem
Vice Chancellor

WORK PLAN

Learning Outcomes

To achieve the aim set out, the course has a set of objectives. Each unit has specific objectives which are included at the beginning of the unit. You may wish to refer to them during your study to check on your progress.
Below are the complete objectives of the course. By meeting these objectives, you should have achieved the aim of the course as a whole. Thus, in addition to the aim earlier stated, this course sets to achieve the under listed objectives. At the end of the course, learners should be able to:

STA 221
Probability Distribution 1

Week 01
- Understand the history of Statistics.

Week 02
- Understand the basic concepts and techniques used in Probability Distribution I.

Week 03
- ·Discuss the concepts of random experiments, sample spaces and events.
- ·Explain how different events and combination of events can be formed from a given random experiment.

- Explain various interpretations of probability and describe the various probability concepts on the specified random experiment.
- Describe random variables of the discrete and continuous types, derive the expectations, compute their associated probabilities and determine higher moments from them.
- ·Describe various probability distributions of the discrete and continuous forms, which includes Binomial, Poisson, Negative Binomial, Poisson, Geometric, Uniform and Normal distributions.
- Discuss the approximation methods of determining probabilities from these probability distributions under certain conditions.

Course Guide

Module 1

Unit 1: Random experiments and sample space
Unit 2: Combination of events
Unit 3: Combinational analysis
Unit 4: Probability measure
Unit 5: Sampling with and without replacement

Module 2

Unit 1: Conditional probability
Unit 2: Baye's theorem
Unit 3: Independent events

Module 3

Unit 1: Probability distribution for discrete random variables
Unit 2: Mathematical expectation for discrete random variables
Unit 3: Probability distribution for continuous random variables
Unit 4: Mathematical expectation for continuous random variables
Unit 5: Joint distributors of two discrete random variables
Unit 6: Joint distributors of two continuous random variables

Module 2

Unit 1: Higher moments of random variables
Unit 2: Binomial distribution
Unit 3: Poisson distribution
Unit 4: Negative binomial and geometric distributions
Unit 5: Uniform distribution
Unit 6: Normal distribution

Requirements for Success

The CODL Programme is designed for learners who are absent from the lecturer in time and space. Therefore, you should refer to your Student Handbook, available on the website and in hard copy form, to get information on the procedure of distance/e-learning. You can contact the CODL helpdesk which is available 24/7 for every of your enquiry.

Visit CODL virtual classroom on http://codllms.unilorin.edu.ng. Then, log in with your credentials and click on STA 221. Download and read through the unit of instruction for each week before the scheduled time of interaction with the course tutor/facilitator. You should also download and watch the relevant video and listen to the podcast so that you will understand and follow the course facilitator.

At the scheduled time, you are expected to log in to the classroom for interaction.

Self-assessment component of the courseware is available as exercises to help you learn and master the content you have gone through.
You are to answer the Tutor Marked Assignment (TMA) for each unit and submit for assessment.

Embedded Support Devices

Throughout your interaction with this course material, you will notice some set of icons used for easier navigation of this course materials. We advise that you familiarize yourself with each of these icons as they will help you in no small ways in achieving success and easy completion of this course. Find in the table below, the complete icon set and their meaning.

Introduction	**Learning Outcomes**	**Main Content**
Summary	**Tutor Marked Assignment**	**Self Assessment**

Web Resources	**Downloadable Resources**	**Discuss with Colleagues**
References	**Futher Reading**	**Self Exploration**

Grading and Assessment

TMA	CA	Exam	Total
20%	20%	60%	100%

Lectures

When discussing the concept of probability theory and statistics, a probability distribution is the mathematical function that gives the probabilities of occurrence of different possible outcomes for an experiment. It is a mathematical description of a random phenomenon in terms of its sample space and the probabilities of events.

The course is designed to introduce students to the basic concepts of probability and their applications and prepare them for future challenges in the use of probability distribution theories in their everyday activities.

Course Justification

Probability concept is a daily phenomenon that characterizes human activities and decisions. A good number of real life events do occur by chance and proper knowledge of various probability theories and applications would be a useful tool to determine the occurrence or non-occurrence of such as the situation arises. A thorough understanding of probability theories would enhance good understanding of several other statistical courses that have their own numerous life applications.

Source: pexels.com

MODULE 1

Unit 1: Random experiments and sample space
Unit 2: Combination of events
Unit 3: Combinational analysis
Unit 4: Probability measure
Unit 5: Sampling with and without replacement

UNIT 1 RANDOM EXPERIMENTS AND SAMPLE SPACE

Introduction

Welcome to the first unit of the first module of this Statistics Course. The mathematical theory of probability we will be studying gives us the basic tools for constructing and analyzing mathematical models for random phenomena. In studying a random phenomenon, we need to describe sample spaces and events for various random experiments.

Learning Outcomes

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

1. Define a random experiment that can result in different outcomes
2. Describe various Sample spaces for various random experiments.
3. Represent various sample spaces and events using tree diagrams.

Main Content

Random experiment

7mins

A random experiment, also known as statistical experiment or stochastic experiment is an experiment that can result in different outcomes, when it is repeated in the same manner every time.

For an experiment to be described as a random experiment, the following conditions must be satisfied

(i) All the possible outcomes of the experiment are known in advance

(ii) Any performance of the experiment results in an outcome that is not known in advance.

(iii) The experiment is repeated under identical conditions.

Outcome is the result obtained when a trial is made

Examples of a random experiment include:

(i) Tossing a fair coin can result in an outcomes head and tail

(ii) Tossing a fair die can result in outcomes corresponding to the number of sports (1, 2, 3, 4, 5, 6) on the faces of the die.

(iii) Picking one item from a batch of items to see whether it is defective or non- defective

From the example of tossing a coin, it is known in advance that the coin could result in a head or a tail, but it is not known in any particular toss whether a head or a tail will come up. Also, from example of tossing a die, it is known that all the possible outcomes should be numbers 1,2,3,4,5,6; but it is not known in advance whether any of these numbers will come up in . Condition **(ii)** implies that it is not known in advance whether it is head or tail that will come out in a particular toss.

Sample Space: The sample space of a random experiment is the set consisting of all possible outcomes of the experiment. It is commonly denoted S. The elements of sample space is usually enclosed in curly brackets.

Consider an experiment involving the tossing of a fair coin once, then the sample space can be given as

$$S = \{\text{Head, Tail}\} \text{ or simply}$$
$$S = \{H, T\}$$

Suppose that, you conduct an experiment that consists of tossing a fair die once, the sample space can be given as the set consisting of all number of spots on the faces of the die

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

If two fair coins are tossed together, then the sample space can be obtained as follows: Both coins can result in heads (HH), the first can result in a head while the second results in a tail (HT), the first can result in a tail while the second results in a head (TH) and both can result in tails (TT). The sample space then is

$$S = \{HH, HT, TH, TT\}$$

It's important for you to note: The same results can be achieved by tossing a coin twice. In that case, (HH) implies heads in first and second tosses, (HT) implies a head in first toss and a tail in second toss, (TH) implies a tail in first toss and a head in second toss while (TT) implies tails in both tosses.

If an experiment consists of tossing two dice together or a die is tossed twice. If we let i denote the number of spot(s) on the first die and j denotes the number on the second die, the sample space consists of (i,j) spots, represented as

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

Above sample space can also be represented as

		1st die					
		1	23	4	5	6	
	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
2nd die	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Now let us consider an experiment in which you select a molded plastic part, such as a connector, and measure its thickness. The possible values for thickness depend on the resolution of the measuring
instrument, and they also depend on upper and lower bounds for thickness. However, it
might be convenient to define the sample space as simply the positive real line

$$S = R^{+} = \{x | x > 0\}$$

Example

If it is known that all connectors will be between 1 and 15 millimeters thick, the sample space could be written as

$$S = R^{+} = \{x | 0 < x < 15\}$$

Sample Point: Each outcome of a sample space is called a sample point. For example, if you toss a coin twice , the sample points are elements (HH), (HT),(TH) or (TH).

In example of tossing two dice, the sample points are (1,1), (1,2), (1,3),.........(6,6). Often, sample points, elements and outcomes are used interchangeably.

A sample space can be said to be discrete if it consists of a finite or countable infinite set of outcomes. For example, the sample spaces of tossing coins and dice are examples of discrete sample space.

A sample space is continuous if it contains an interval (either finite or infinite) of real numbers. The example on the thickness of plastic part is continuous sample space. Often, discrete sample space can be defined from continuous sample space. For example, if we classify the thickness as low if it lies between 1 and 5 millimeters, medium if it lies between 6 and 10 millimeters and high if it is above 10 millimeters. the sample space might be taken to be the set of three outcomes:

$$S = \{\text{low, medium, high}\}$$

Event: Event is any subset of the sample space. It can be denoted by letters A, B,C, D, E etc.

Examples

(i) In example of tossing a coin above, we can define event A ={H}, as an event that a head turns up in the toss.
(ii) In the example involving tossing a coin twice, we can define event E = {HH, HT} as the event that a head occurs in the first toss
(iii) For the toss of two dice, event C ={(1,6), (2,5), (3,4),(4,3), (5,2), (6,1)}is the event that the sum of the two numbers that occur on the two dice equals 7.

Since event is a set also, it is enclosed in a curly bracket
Several events can be defined on a particular sample space. For example, consider an experiment consisting of tossing a coin and rolling a die together once. The sample space consists of the various combinations of head (H) or tail (T) with numbers 1,2,3,4,5,6 and can be given as
S_1 = {H1, H2, H3, H4, H5, H6, T1, T2, T3, T4, T5, T6}

The following different events can be defined on S1
E_1:An event that the sample point contains a head, and this is given as
E_1= {H1, H2, H3, H4, H5, H6}
E_2 : An event that the sample point contains an even number
E_2 ={H2, H4, H6, T2, T4, T6}
E_3: An event that the sample point contains a tail and a number greater than 2
E_3 ={T3,T4, T5, T6}

Other forms of events can also be described, which include
Null event: An event that does not contain any element or sample point and it denoted ϕ or { }

Elementary Event: An event that contains only one element

Compound Event : Contains more than one element. In the toss of a coin twice with sample space

S={HH,HT,TH,TT}

The even given by A={HT} is an elementary event. An example of compound event is

B={HH,HT}

Containment: It is said that a set A is contained in another set B if every element of the set A also belongs to the set B. This relation between two events is expressed symbolically by the expression $A \subset B$, which reads A is a subset of B. Equivalently, if $A \subset B$, we may say that B contains A and may write $B \supset A$.

Consider an example of rolling, suppose that A is the event that an even number is obtained defined by A={2, 4, 6} and C is the event that a number greater than 1 is obtained given as C ={2, 3, 4, 5, 6}, it follows that $A \subset C$

Activity 1

Suppose that three items are selected at random from a manufacturing process. Each item inspected is classified defective, D or non-defective, N. List the elements of the sample space.

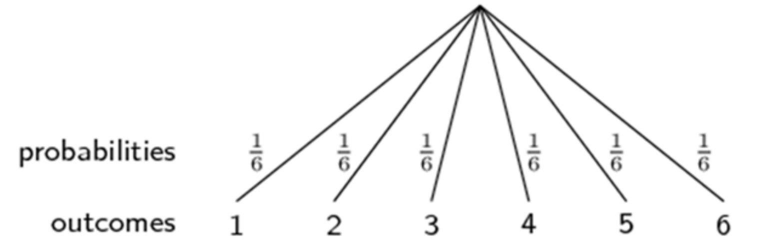
Solution: S ={DDD, DDN, DND, DNN, NDD, NDN, NND, NNN }

Tree diagram

A tree diagram is simply a way of representing a sequence of events. It is useful for organizing and visualizing the different possible outcomes of a sequence of events. For each possible outcome of the first event, we draw a line where we write down the probability of that outcome and the state of the world if that outcome happened. Then, for each possible outcome of the second event we do the same thing.

Example

The tree diagram below represents the possible outcomes tossing a 6-sided die.

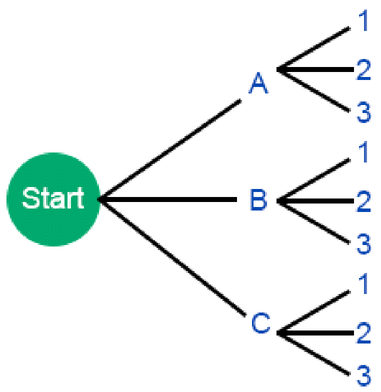


Activity 2

Suppose that you want to play a game, there are three doors A, B, C and 3 windows 1,2,3 after each door to get to the next point. How many choices will be there for a person to get to the next point? Draw a tree diagram to visualize the different possibilities.

Solution:
Each of the doors A, B, C can be linked to each of the windows and we have all the possible combinations as follows
S = {A1, A2, A3, B1, B2, B3, C1, C2, C3}

The tree diagram will give the branching of doors first, and then for each door there will be branching of windows. This can be represented as below



You can flip the coin 3 times in one turn. State the events that one will get exactly 3 points in one turn? Draw a tree diagram to visualize the different possibilities.

The possibilities are A3, B3, C3. Thus the event can be listed in the usual form in a set E as

E ={A3, B3, C3}

Summary

Thus far in this unit, you have gotten to understand that a random experiment, also known as statistical experiment or stochastic experiment which was defined as an experiment that can result in different outcomes when it is repeated in the same manner. Sample spaces for various random were describe and represented using tree diagrams. I also defined random experiment, sample space and event.

Self Assessment Questions

1. State the conditions under which an experiment can be termed a random experiment.
2. Provide a reasonable description of the sample space for each of the random experiments below
   (a) Each of three machined parts is classified as either above or below the target specification for the part.
   (b) Each of four transmitted bits is classified as either in error or not in error.
   (c) In the final inspection of electronic power supplies, three types of non-conformities might occur: functional, minor, or cosmetic.
3. Write the sample space S for the following random experiments.
   (a) We toss a coin until we see two consecutive tails. We record the total number of coin tosses.

   (b) A bag contains 4 balls: one is red, one is blue, one is white, and one is green. We choose two distinct balls and record their color in order.
   (c) A customer arrives at a bank and waits in the line. We observe T, which is the total time (in hours) that the customer waits in the line. The bank has a strict policy that no customer waits more than 20 minutes under any circumstances.
4. A survey of some women were carried out on where the delivery of their first child took place. The possible responses are government hospital (G), private hospital (P) or home (H). A child delivered could either be male (M) or a female (F). Specify the sample space for the survey.
5. A bag contains 3 black and 5 white balls, two ball are drawn at random. Draw a tree diagram for the two draws.

Tutor Marked Assignment

1. There are 4 Kings from a deck of card. One coin is tossed and one card is chosen from those 4 Kings. Obtain the sample space and represent it in a tree diagram
2. (a) Construct a sample space that describes all three-child families according to the genders of the children with respect to birth order by denoting boy with "b" and girl with "g"
   (b) Use tree diagram to represent the sample space in (a)
3. List the elements of the sample space in picking 2 marbles, one at a time, from a bag that contains many blue and red marbles.

References

1. Hogg, R.V. and Tanis, E.A. (1977). Probability and statistical Inference.
2. Walpole, R.E., Myers, R.H., Myers, S.L.. and Keying, Ye .(2004). Probability & Statistics for Engineers & Scientists. ISBN 81-7808-613-1.

Further Reading

https://www.siyavula.com/read/maths/grade-11/probability/10-probability-04
https://www.pdfdrive.com/mathematical-statistics-with-applications-7th-edition-e28265865.html

UNIT 2 COMBINATION OF EVENTS

Introduction

In this unit, we will be focusing on the combination of events having known what an event is from the previous unit. A new event can be described from the combination of two or more events. Because events are also subsets, we can use basic set operations such as unions, intersections, and complements to form other events of interest.

Learning Outcomes

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

1. Describing various forms of combination of events
2. Represent events by means of venn diagrams
3. State the rules guiding set operation under the combination of events

Main Content

Combination of Events

Some of the basic set operations in terms of events are summarized below.

Union: The union of two events is the event that consists of all outcomes that are contained in either of the two events. We can denote the union of events A and B as $A \cup B$. The concept of union extends to more than two sets. For example, The *union of k sets* $A_1, \ldots, A_k$ is defined to be the set that contains all outcomes that belong to at least one of these k sets. The notation for this union is either of the following: $A_1 \cup A_2 \cup \ldots \cup A_k$

Example 1

Let A = {a, b, c} and B = {b, c, d, e},
then the union of events A and B is given as
$A \cup B$ = {a, b, c, d, e}

Example 2

Suppose that P is an event that a person randomly selected from a tobacco company smokes cigarette, and Q denotes that event that the selected employee drinks alcohol, then the event $Q \cup P$ is the set of all employees who either smokes or drinks or do both.

Intersection: The intersection of two events is the event that consists of all outcomes that are contained in both of the two events. We denote the intersection of A and B as $A \cap B$

Example 1

Suppose that A person randomly selected at a restaurant while dining is a taxpayer, and B be event that the person is over 60 years of age, then $A \cap B$ is the set of all taxpayers who are above 60 years of age.

Example 2

Let M = {a, e, i, o, u} and N ={b, c, d, e},
then $M \cap N = \{e\}$

Complement: The complement of an event in a sample space is the set of outcomes in the sample space that are not in the event. We denote the complement of event E as E'

Example 1

Let E be the event that a red card is selected from a bag containing playing cards and let S be the entire deck consisting of red and black cards. Then E' is the event that the card selected from the bag is not red but black card

Example 2

Consider the sample space
S = {book, catalyst, cigarette, precipitate, engineer, rivet}
Let A = {catalyst, rivet, book, cigarette}. Then
$A' = \{\text{precipitate}, \text{engineer}\}$

Activity 1

In an experiment involving tossing a fair coin three times, suppose that the sample space consists of eight sample points given as
S = {HHH, HHT, HTH, HTT, THH, THT, TTT}
If we define the following events on S

A: A head appears on the first coin ={HHH,HHT,HTH}
B: The point contains two heads={HHT,HTH,THH}

Obtain the following events

(a) Union of A and B
(b) Intersection of A and B
(c) Complement of A

Solution:
(a) The union of A and B is obtained by listing all the elements of A and B without repeating any of them, given as

$$A \cup B = \{HHH, HHT, HTH, THH\}$$

(b) The intersection of events A and B is obtained by listing the element/elements that appears/appear in both events and is given as

$$A \cap B = \{HHT, HTH\}$$

(c) The complement of event A is obtained by listing the element/elements that appears/appear in S but not in A, and this is given by $A' = \{HTT, THH, THT, TTT\}$

Mutually exclusive events

Two events A and B are said to be mutually exclusive if the event that correspond to their intersection is an empty set. That is $A \cap B = \phi$
For example, consider a sample space consisting of integers 1 through 10. Suppose we define the following events given by

S = {1, 2, 3, 4, 5, 6, 7, 8, 9,10}

A: The set containing even numbers ={2, 4, 6, 8, 10}
B: The set containing numbers between 2 and 6 inclusive ={2, 4, 6}
C: The set containing odd numbers {1, 3, 5, 7, 9}

AB = {2, 4, 6}
AC = ϕ : Mutually exclusive
BC = ϕ: Mutually exclusive

Activity 2

I will like you to consider the following sample space of an experiment $S = \{x|0<x<\infty\}$, and let's define the following events

$E_1 = \{x|1 \le x<10\}$
$E_2 = \{x|3<x<118\}$

Give the elements of the following events
(i) $E_1 \cup E_2$
(ii) $E_1 \cap E_2$
(iii) E'_1

Solution:
(i) $E_1 \cup E_2 = \{x|1<x<118\}$
(ii) $E1 \cap E_2 = \{x|3<x<10\}$
(iii) $E'_1 = \{x|x \ge 10\}$

Venn Diagrams

Venn diagrams are often used to describe relationships between sets. We can use Venn diagrams to represent a sample space and events in a sample space.

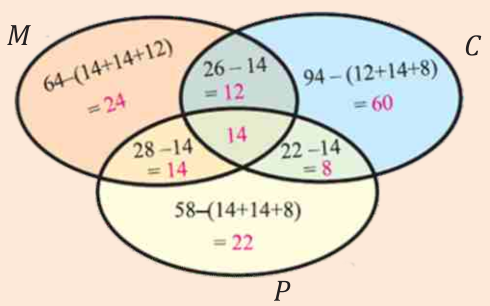
(a) (b) (c)

The Venn diagram above represents the sample space of the random experiment as the points in the rectangle S, where events A and B are the subsets of points in the indicated regions; (b) illustrates two events with no common outcomes in which $A \cap B = \phi$, and the shaded region in (c) represents the intersection between events A and B.

Activity 3

In a survey of university students, 64 had taken mathematics course, 94 had taken chemistry course, 58 had taken physics course, 28 had taken mathematics and physics, 26 had taken mathematics and chemistry, 22 had taken chemistry and physics course, and 14 had taken all the three courses. Find how many had taken one course only.

Solution:
Let M, C and P represent the courses Mathematics, Chemistry and Physics respectively.
With respective numbers
n(M) = 64, n(C) = 94, n(P) =58



From the Venn diagram above, we have
No. of students who had taken only math = 24
No. of students who had taken only chemistry = 60
No. of students who had taken only physics = 22
Total no. of students who had taken only one course
= 24 + 60 + 22 = 106

Operation of Set

Let A, B, C be events (sets) defined on the sample space S, then the following rules hold.

Commutative rule
(i) $A \cup B = B \cup A$, $A \cap B = B \cap A$

Associative rule
(ii) $A \cup (B \cup C) = (A \cup B) \cup C$, $A \cap (B \cap C) = (A \cap B) \cap C$

Distributive rule
(iii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B)$

The following rules also hold
(iv) $A \cup A = A$ and $A \cap A = A$
(v) $A \cup S = S$ and $A \cap S = A$
(vi) $A \cap A' = \phi$ and $A \cup A' = A + A' = S$

De Morgan's rule
(vii) $(A')' = A$ $(A \cup B)' = A' \cap B'$
(viii) $(A \cap B)' = A' \cup B'$

Activity 4

Suppose that a die is tossed once with the sample space given by S = {1, 2, 3, 4, 5, 6}. Define the following events
A = {1, 3, 5}
B = {2, 4, 6}
C = {1, 2, 3.4}
Use these to demonstrate the rules of probability listed below
(i) $A' \cup B') = A(\cap B)'$
(ii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Solution:
(i) From the sets, we have that

$A' = \{2, 4, 6\}$
$B' = \{1, 3, 5\}$
$A' \cup B' = \{1, 2, 3, 4, 5, 6\}$ ***
$A \cap B = \phi$
$(A \cap B)' = \{1, 2, 3, 4, 5, 6\}$ +++

From *** and +++,
(ii) $B \cup C = \{1, 2, 3, 4, 6\}$
$A \cap B = \phi$
$A \cap C = \{1, 3\}$
From these sets
$A \cap (B \cup C) = \{1, 3\}$ and $(A \cap B) \cup (A \cap C) = \{1, 3\}$

Summary

In this unit, I gave the definition of random experiment as the experiment whose outcomes in any performance in not known in advance even though all the possible outcomes must be known in advance. Sample space was also defined and various examples were given for its illustrations. Also, I defined different types of events for various sample spaces. Combinations of events, including union, intersection and complement were also defined and some numerical examples were used for the illustrations.

Self Assessment Questions

1. Consider an experiment with sample space given as

$$S = \{AA, AD, DA, DD\}$$

Suppose we define the following events

$E_1=\{AA, AD, DA\}$

$E_2 = \{DD\}$

$E_3 = \{AD, DA, DD\}$

Obtain

(a) $E_1 \cup E_2$

(b) $E_1 \cap E_2$

(c) E'_1

2. Let S={1, 2,..., 10}, and we define the following events

A = {1, 3, 5}

B = {1, 4, 6}

C = {2, 5,7}

Determine

(a) $S \cup C$ (b) $A \cup B$ (c) $A' \cap C$ (d) $A' \cup (B \cap C)$

3. Last July, there were 20 windy days and 15 rainy days, yet 5 days were neither windy nor rainy.

Represent the information on a venn diagram

4. Suppose that a die is tossed once with the sample space given as

S = {a, b, c, d, e, f}

Define the following events

A = {a, c, e}

B = {b, d, f}

C = {a, b, c, d}

(a) Use these events to demonstrate the rules of probability listed below

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

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

(b) What principles of set operation guide your answers to questions (i) and (ii) in (a)?

Tutor Marked Assignment

1. In a town 85% of the people speak Tamil, 40% speak English and 20% speak Hausa. Also 32% speak Tamil and English, 13% speak Tamil and Hausa and 10% speak English and Hausa, find the percentage of people who can speak all the three languages.
2. Let A, B, C be arbitrary events. Find expressions for the events that of A, B, C
   (a) none occur
   (b) only A occurs
   (c) at least one occurs
   (d) B and C occur but A does not occur
3. Consider a random experiment involving the toss of a fair die once. Suppose we define the
   Following events:
   A = { 2,3,5}
   B = {2, 4, 6}
   C = { 1, 3, 5}
   Obtain
   (a) $A \cup B$
   (b) $A \cap C$

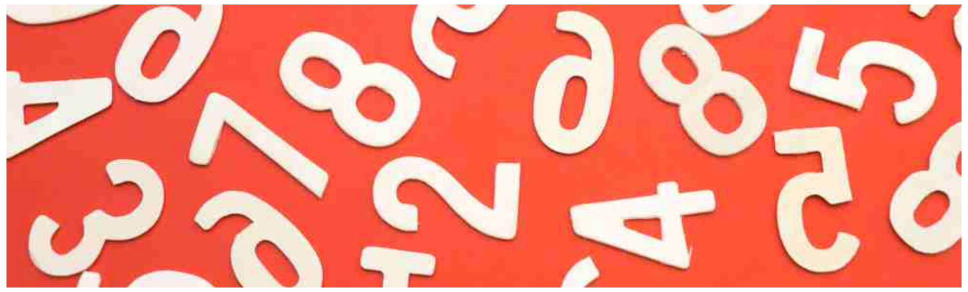
References

1. Hogg, R.V. Tanis, E.A. (1977). Probability and statistical Inference.
2. Walpole, R.E. Myers, R.H. Myers, S.L. & Keying, Ye (2004). Probability & Statistics for Engineers & Scientists. ISBN 81-7808-613-1.

Further Reading

https://www.onlinemath4all.com/venn-diagram-word-problems-with-3-circles.html

https://www.pdfdrive.com/mathematical-statistics-with-applications-7th-edition-e28265865.html



UNIT 3 COMBINATORIAL ANALYSIS

Introduction

It is important to note at this point that the combinatorial analysis stems from the fundamental counting principle which states that if you wish to find the number of outcomes for a given situation, simply multiply the number of outcomes for each individual event. For example, if a child has two different choices of drink and three different choices of meat. Then the total number of choices he/she can make is 2 times 3 which gives total of 6 number of outcomes possible. The basic counting principle may be expanded to any number of events. For example, suppose for example that a customer wishes to install Trimline telephone and can choose from n_1 =10 decorator colours which we shall assume are available in any of n_2=3 optional cord lengths with n_3 =2 types of dialing namely, rotary. or touch-tone. These three classifications result in $n_1 \times n_2 \times n_3 = 10 \times 3 \times 2 = 60$ different ways for a customer to order one of these phones. So in details we will be looking at combinatorial analysis.

Learning Outcomes

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

1. Describe the concept of counting principle
2. Apply the fundamental counting principle to determine the number of outcomes
3. Describe ordered and unordered arrangements of objects and illustrate with numerical examples.

Main Content

Basic Principles of Counting

5mins

Let's assume you have two experiments to perform. If experiment 1 can result in any one of n_1 outcomes and if for each outcome of experiment 1, there are n_2 possible outcomes of experiment 2, then together there are $N = n_1 \times n_2$ possible outcomes of the two experiments.

Proof of the Basic Principles

The basic principle may be proved by enumerating all possible outcomes of the two experiments as follows:

(1, 1), (1, 2),…, (1, n2)
(2, 1), (2, 2),…, (2, n2)
.
.
.
$(n_1, 1), (n_1, 2), \ldots, (n_1, n_2)$

We say that outcome is (i, j) if experiment 1 results in its ith outcome and experiment 2 results in the jth of its possible outcomes. Thus the set of possible outcomes consists of m rows, each row containing m elements, and this proves the results.

Generally, if there are n_1 ways to carry out the first stage of the experiment; for each of these n_1 ways, there are n_2 ways to carry out the second stage; for each of these n_2 ways, there are n_3 ways to carry out the third stage, and so forth. Then the total number of ways in which the entire task can be accomplished is given by the product

$N = n_1 \times n_2 \times \ldots n_k.$

Permutation (Unordered arrangement)

This is the number arrangements of n distinct objects in a row taking all at a time. It is the number of different arrangement of objects in a row when the order of arrangements is not relevant. It is given as

n! = n(n - 1) (n - 2)…3.2.1. For example, in the arrangement of 3 letters a,b,c, there are 6! possible permutations, which are abc, acb, bac, bca, cab, cba.

This results follows the basic principle as follows. Since the first object in the permutation can be any of the 3, the second object in the permutation can then be chosen from any of the remaining 2, and the third object in the permutation is then chosen from the remaining one. Thus, there are 3x2x1 = 6 possible permutations.

Permutation taking r < n at a time

To develop a formula for the number of permutation of anything taken r at a time, we consider the special of 4 objects arranged in sets of 3. In this case, there are 4 choices for the first letter and after the first letter has been chose, there are choices left for the second letter. Thus for the first two letters, there are (4)(3) =12 possible choices. Now, for the third letter, there remains two letters choices, so there are (4)(3)(2) =24 choices of 3 letters out of a total of 4, this shows that, in general, if we denote the number of permutations of n objects taking r at a time by ^nP_r, we need to take the product of the r numbers n, n-1,n-2,…, or

$$^nP_r = \frac{r \text{ factors}}{n(n-1)(n-2)\ldots(n-r+1)}$$

If we multiply and divide the right-hand side of above by (n-r)(n-r-1)…(2)(1),
the numerator becomes n!, thus we can write

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

Example

The number of permutations of 3 objects from 4 using the formula can be obtained as follows:

$$^4P_3 = \frac{4!}{(4-3)!}$$

$$(\frac{4!}{(4-3)!} = 24$$

Permutation of similar objects

The permutation of n objects in which n_1 are alike, n_2 are alike, n_k are alike is

$$\frac{n!}{n_1!n_{2!}\cdots n_k!}, \text{ where } \sum_{j=1}^{k} n_j = n$$

For example, suppose there are 12 ball of which 3 are white, 4 are red and 5 are black. The number of possible arrangements in a row if only their colour are distinguished can be obtained as follows

Number of white = 3
Number of red = 4
Number of black = 5

The required number of arrangement

$$\frac{12!}{3!4!5!} = \frac{479001600}{6(24)(120)}$$

$$=27720$$

Activity 1

In a class of 6 men and 4 women, an examination was given and the students were ranked according to their performance. Suppose that no two students obtained the same score

(a) How many different rankings are possible?
(b) How many different rankings are possible if the men are ranked just among themselves and the women among themselves?

Solution:

(a) If men and women are not distinguished, then the possible number of rankings are

$(6+4)! = 10! = 10\,(9)(8)\ldots 3.2.1$
$= 3{,}628{,}800$ rankings

(b) Men can be ranked among themselves in 6! ways and the women can be ranked among themselves in 4! ways. Thus the total number of rankings is (6!)(4!) = 17, 280

Activity 2

List all the possible permutations of the numbers 11223 that are possible

Solution: There are

$$\frac{5!}{2!4!!} = 30 \text{ possible permutations of the numbers}$$

The numbers are

11223	11232	11322	12123	12132
12213	12231	12312	12321	13122
13212	13221	21123	21132	21213
21231	21312	21321	22113	22131
22311	23112	23121	23211	31122
31212	31221	32112	32121	32211

Combination (Ordered arrangement)

This is the number of different groups of r objects that could be selected from a total of n objects, when the order of selection is relevant.

To develop a formula for the number of combinations of n objects taken r at a time, we note that for every set of r letters there is just one combination if all r letters are to be used,

As each group of r objects is counted r! times, it follows that the number of different groups of r objects that could be formed from a set of n objects is

$$\binom{n}{r} = \frac{n(n-1)(n-2)\ldots(n-r+1)!}{r!} = \frac{{}^nP_r}{r!}$$

Thus the number of selection of r objects from n when ordering is relevant is

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

Activity 3

From a group of 5 men and 7 women, how many different groups of 5 can be selected consisting of 2 men and 3 women?

Solution:

There are $\binom{5}{2} = \frac{5!}{(5-2)!2!} = 10$ possible groups of 2 men from 5 and

$\binom{7}{3} = \frac{7!}{(7-3)!3!} = 35$ possible groups of 3 women from 7

Thus from the basic counting rule, there are $\binom{5}{2} \times \binom{7}{3} = 350$ possible groups of 5 that can be selected

Activity 4

Find the value (s),o of n for which

(a) ${}^{n+1}P_3 = {}^nP_4$

(b) $\binom{n+1}{3} = \frac{7}{3}\binom{n}{2}$

Solution:

(a) ${}^{n+1}P_3 = {}^nP_4$ implies that

$$\frac{(n+1)!}{(n+1-3)!} = \frac{n!}{(n-4)!}$$

$$\frac{(n+1)(n)(n-1)(n-2)!}{(n-2)!} = \frac{n(n-1)(n-2)(n-3)(n-4)!}{(n-4)!}$$

$n+1 = (n-2)(n-3)$
$n+1 = n^2 - 5n + 6$
$n^2 - 6n - 5 = 0$
$n(n-5) - 1(n-5) = 0$
Thus n =1 or n =5.

(b) $\binom{n+1}{3} = \frac{7}{3}\binom{n}{2}$ implies that

$$\frac{(n+1)!}{3!(n+1-3)!} = \frac{7}{3}\left(\frac{n!}{(n-2)!}\right)$$

$$\frac{(n+1)!}{(3).2!(n-2)!} = \frac{7\,n!}{(3)2!(n-2)!}$$

$(n+1)! = 7n\,!$

$(n+1)n! = 7n\,!$

$n+1 = 7$

$\therefore\ n = 6$

Activity 5

A box contains 4 red, 3 white and 2 blue balls. Three balls are drawn at random. How many ways can balls of different colours be selected?

Solution:

1 red ball can be selected in ${}^{4}C_{1}$ ways.
1 white ball can be selected in ${}^{3}C_{1}$ ways.
1 blue ball can be selected in ${}^{2}C_{1}$ ways.
Total number of ways
$= {}^{4}C_{1} \times {}^{3}C_{1} \times {}^{2}C_{1}$
$= 4 \times 3 \times 2 = 24$

Summary

This unit I gave the basic principle of counting. The definition of permutation as unordered arrangement of distinct objects was given and various forms of permutations I illustrated with examples. Also, combination, which is a selection of say r objects from n distinct objects when ordering is important was given with illustrative examples.

Self Assessment Questions

1. Suppose you are buying a new car where there are 2 body styles, 5 colours available and 3 models. How many choices are possible?
2. A circular table has 6 chairs, out of this 6, five are identical. In how many ways can the six people be arranged on these chairs?
3. In how many ways can letters of the word PAWPAW be arranged?
4. What is the value of $5_{P_2} \times 3_{P_2} - 3!$

Tutor Marked Assignment

1. In how many different ways can 5 girls and 5 boys form a circle such that the boys and the girls alternate?
2. Consider an operation that can be completed in 3 steps. If step 1 can be completed in 2 ways, step 2 in 3 ways and step 3 in 4 ways, how many ways can be the operation be completed?
3. In how many ways can 10 people be seated on a bench if 4 seats are available?

References

1. Bluman, A.G. Elementary Statistics. A step by step approach. ISBN:0-07-254907-6.
2. Walpole, R.E. Myers, R.H. My qtf, S.L. and Keying Ye. (2004). Probability & Statistics for Engineers & Scientists. ISBN 81-7808-613-1.

Further Reading

https://www.careerbless.com/aptitude/qa/permutations_combinations.php

UNIT 4 PROBABILITY MEASURE

Introduction

The likelihood of the occurrence of an event resulting from a statistical experiment is evaluated by means of a set of real numbers called weights or probabilities, ranging from 0 to 1. To every point in the sample space, a probability is assigned such that the sum of all probabilities is 1. If we have reason to believe that a certain sample point is likely to occur when the experiment is conducted, then the probability assigned should be close to 1. On the other hand, a probability close to zero is assigned to a sample point that is not likely to occur. In an experiment such as tossing a coin, or a die, all the sample points have the same chance of occurring and are assigned equal probabilities. For points outside the sample space, that is for simple events that cannot probably occur, a probability of zero is assigned.

Learning Outcomes

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

1. Describe various interpretations of probability.
2. Describe the concept of probability of an event and the rules guiding probability measure.
3. Calculate some probabilities based on the various types of events.

Main Content

Interpretations of Probability

8mins

The true meaning of probability is still a highly controversial subject and is involved in many current philosophical discussions pertaining to the foundations of statistics. However, in its simplest form, probability can be defined as a mathematical term which shows the degree of belief which we have about something. There are three approaches through which probability can be interpreted. Each of these interpretations can be very useful in applying probability theory to practical problems.

Classical Interpretation

This interpretation is based on the concept of equally likely outcomes. For example, when a coin is tossed, there are two possible outcomes, head or tail. If the outcomes are assumed equally likely to occur, then the probability of head occurring equals thee probability of tail occurring which equals 1/2. Suppose a die is tossed once, the possible outcomes are 1, ...,6. If these outcomes are equally likely, then each of them has probability 1/6 of occurring. Generally, if the outcome of some process must be one of n different outcomes, and if these n outcomes are equally likely, then the probability of each outcome is 1/n. The difficulty of classical approach is that no systematic method is given for assigning probabilities to outcomes that are not assumed to be equally likely.

Frequentist Interpretation

This interpretation describes the relative frequency with which the outcome will be obtained if the process were repeated a large number of times under similar conditions. For example, the probability of obtaining a head when a coin is tossed is considered to be 1/2 because the relative frequency of heads should be approximately 1/2 when the coin is tossed a large number of times under similar conditions. The limitation of this approach are:

(1) The conditions of large number of times is vague since there is no definite indication of actual number that is considered large.

(2) The identical condition under which the experiments is to be performed is not precisely described.

Subjective Interpretation

This approach also known as Bayesian interpretation centers on personal belief of the occurrence of events. This personal judgement is based on the person's information about the process. Another person who may have different belief or different information may assign a different probability to the same outcome. For example, tossing a coin once, a person without information about the coin or the way in which it is tossed might regard a head and a tail to be equally likely outcomes and then assign a subjective probability of 1/2 of obtaining a head. However, the person who tossed the coin might feel that a head is much more likely to occur than a tail and thus assign a probability of 3/4 of head occurring.

The difficulty of subjective interpretation is that a person's judgement of the relative likelihoods of an infinite number of events be completely consistent and free from conditions does not seem to be humanly attainable. Also, The approach provides no objective basis for two or more scientist working together to reach a common evaluation of common interest areas.

Probability of an Event

Let S be a sample space of a random experiment and A be any event defined on S. The probability of event A denoted Pr(A) is defined as

$$\Pr(A) = \frac{\text{Measure of A}}{\text{Measure of S}}$$

For experiment with set of equally likely outcomes, the probability of event a is given by

$$\Pr(A) = \frac{\text{Number of outcomes in A}}{\text{Number of outcomes in S}}$$

Activity 1

In the toss of a fair die twice, find the probability that

(a) The same numbers appear on the two dice
(b) Different numbers appear on the two dice
(c) The number that appears on the first die at least 4

Solution:

Let the sample space be given by
{(1,1),(1,2),...,(6,5),(6,6)}
The number of outcomes of the sample space n(S)=36)
(a) Let A: the same number appear on the two dice, i.e {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}
n(A) =6
Thus Pr(A) = 6/36 = 1/6
(b) Let B: Different numbers appear on the two dice,
Pr(B) = 1 - Pr(A)
1 - 1/6 = 5/6
Let C: the number that appears on the first die is at least i.e {(4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}, n(C) = 18
Pr(C) = 18/36

Probability Measure

In a given random experiment, it is necessary to assign to each event A in the sample space *S* a number that indicates the probability that A will occur. In order to satisfy the mathematical definition of probability, the number Pr(A) that is assigned must satisfy three specific axioms. These axioms ensure that the number Pr(A) will have certain properties that we intuitively expect a probability to have under each of the various interpretation of probability.

Definition: Let S be a sample space of a random experiment and A be any event defined on *S*. A set function Pr(A) defined on S is called a "**Probability measure**" if it satisfies the following conditions, known as "**Axioms of Probability**".

Axiom 1: $Pr(A) \geq 0$ for all A in S
Axiom 2: $Pr(S) = 1$
Axiom 3: For any two events A_1 and A_2 such that $A_1 \cap A_2 = \phi$,
$Pr(A_1 \cup A_2) = A_1 + A_2$, where
$A_1 \cap A_2 = \phi$, implies that events A_1 and A_2 are mutually exclusive or disjoint.
Axiom 3 can be generalized as follows, for any countable collection of mutually exclusive events $A_1, A_2, \ldots$ in S, where $Pr(A_i \cap A_j) = \phi$ for all $i \neq j$,
$Pr(A_1 \cup A_2 \cup \ldots) = Pr(A_1 + A_2 + \ldots)$

Operations of sets

Theorem: If A is any event in S, then $Pr(A') = 1 - Pr(A)$ and
Proof: From rule (vi) above, $A' = S - A$,
Then $Pr(A') = Pr(S-A)$
$= Pr(S) - Pr(A)$
$= 1 - Pr(A)$ (Since the probability of the sample space S)
= 1 from Axiom 2 above
Also, $Pr(A) = 1 - Pr(A')$

Addition rule of probability: Let A and B be any two events in S, then
$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$
Proof: The event (AB) can be represented as a union of disjoint sets, namely,)

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

By Axiom 3 of probability,

$$\Pr(A \cup B) = \Pr(A) + \Pr(A' \cap B) \qquad \text{(aa)}$$

Which is a union of disjoint sets. Thus
However,

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

Thus

$$\Pr(B) = \Pr(A \cap B) + \Pr(A' \cap B) \qquad \text{(bb)}$$

from which

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

The result in (bb) is substituted into (aa) to obtain

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

Above can be extended to more than two events
For example, for three events A, B and C,
$Pr(A \cup B \cup C) = Pr(A) + Pr(B) + Pr(C) - Pr(A \cap B) - Pr(A \cap C) - Pr(B \cap C) + Pr(A \cap B \cap C)$

Activity 2

A total of 28 percent of American males smoke cigarettes,7 percent smoke cigars, and 5 percent smoke both cigars and cigarettes. What percentage of males smoke neither cigars nor cigarettes?

Solution:
Let E be the event that a randomly chosen male is a cigarette smoker and let B
be the event that he is a cigar smoker and $E \cap B$ be event that he is a smoker of both cigars and cigarettes so that $Pr(E) = 0.28$, $Pr(B) = 0.07$ and $Pr(E \cap B) = 0.05$
Let $(E \cup B)'$ be the event that he smokes neither cigars nor cigarettes smoker with probability $\Pr(E \cup B)'$. If the probability of being either a cigar or cigarette smoker is $Pr(E \cup B)$, then knowing that
$\Pr(E \cup B) + \Pr(E \cup B)' = 1$, from which $\Pr(E \cup B)' = 1 - \Pr(E \cup B)$
But,

$$\Pr(E \cup B) = \Pr(E) + \Pr(B) - \Pr(E \cap B)$$
$$= 0.28 + 0.07 - 0.05 = 0.3$$

So that $\Pr(E \cup B)' = 1 - 0.3$
$= 0.7$
Therefore 70 percent of American males smoke neither cigarettes nor cigars.

Activity 3

A die is loaded in such a way that an even number is twice as likely to occur as a n odd number. If A is the event that a number less than 4 than 4 occurs on a single toss of the die, find Pr(A)
Solution: The sample space is $S = \{1,2,3,4,5,6\}$
We assign a probability of y to each even number and 2y to every even number as follows

Number	Probability
1	y
2	2y
3	y
4	2y
5	y
6	2y

Since the sum of the probabilities must be 1, we have
$y + 2y + y + 2y + y + 2y = 9y = 1$
so that $y = 1/9$ the probability of each even number
let $A = \{1, 2, 3\}$, then
$Pr(A) = 1/9 + 2/9 + 1/9 = 4/9$

Activity 4

Suppose that the history of a number of wafers in a semiconductor manufacturing process are presented in the table below. The wafers were further classified as either in the "center'' or at the "edge'' of the sputtering tool that was used in manufacturing, and by the degree of contamination. The table shows the proportion of wafers in each category.

Number of Contamination Particles	**Centre**	**Edge**	**Total**
0	0.30	0.10	0.40
1	0.15	0.05	0.20
2	0.10	0.05	0.15
3	0.06	0.04	0.10
4	0.04	0.01	0.05
≥5	0.07	0.03	0.10
	0.72	0.28	1.00

What is the probability that a wafer was either at the edge or that it contains four or more particles?

Solution:
We shall define the following events.
A: the event that a wafer contains four or more particles, and
B : the event that a wafer is at the edge
From above Pr(A) = 0.05 + 0.10 = 0.15, Pr(B) = 0.28 and
$\Pr(A \cap B) = 0.04$
Then the required probability is
$\Pr(A \cup B) = 0.15 + 0.28 - 0.04 = 0.39$

Summary

In this unit, we focused on three approaches to interpreting probabilities of events were highlighted, which include classical, frequentist and Subjective approaches. Some numerical examples of probability of events were given. Various rules guiding probability were described and examples of probabilities based on these rules were given.

Self Assessment Questions

1. Briefly explain the following approaches to probability interpretation
   (a) Classical approach
   (b) Frequentist approach
   (c) Bayesian approach
2. A box contains 200 identical machine parts of which 100 are produced by machine A, 60 by machine B and 40 by machine C. If a part is randomly selected from the box, what is the probability that it is produced by
   (a) machine A or B.
   (b) neither machine A or B.
3. What is the probability of getting a total of 7 or 11 when a pair of dice are tossed?
4. A random experiment can result in one of the outcomes {a, b, c, d} with probabilities 0.1, 0.3, 0.5 and 0.1 respectively. Let A denote the event {a, b}, B the event {b, c, d}, and C the event {d}.
   Find the probabilities
   (a) $\Pr(A')$.
   (b) $\Pr(A \cap B)$.
   (c) $(A \cup B)$.
5. Suppose that a fair die is tossed twice, what is the probability that the number of spots that appear on the first die is greater than that of second die.

Tutor Marked Assignment

1. If $\Pr(A) = 0.3$, $\Pr(B) = 0.2$ and $\Pr(A \cap B) = 0.1$, determine the following probabilities
   (a) $\Pr(A \cup B)$ (b) $\Pr(A')$ (c) $\Pr(A' \cap B)$ (d) $\Pr(A \cap B')$
   (e) $\Pr((A \cup B)')$ (f) $\Pr(A' \cup B)$
2. Consider an experiment involving experimental stimulus administered to an animal which may either respond (R) or not respond (N). If the stimulus is administered to three animals, what is the probability that at least 2 of them will respond?
3. Three candidates A, B and C are contesting an election to the Senate, candidates A and B have equal probabilities of winning while candidate C has two third the probability of A of winning. Find the probability that candidate C will lose the election

References

1.Hogg, R.V. and Tanis, E.A. (1977). Probability and statistical Inference.
2. Ross, S.M. (2004). Introduction to Probability and Statistics for Engineers and Scientists.

Further Reading

Montgomery, D. C. Runger, George C. Applied Statistics and Probability for Engineers.
John Wiley & Sons. Inc.
https://www.pdfdrive.com/probability-and-statistics-textbook-e20201392.html

UNIT 5 SAMPLING WITH AND WITHOUT REPLACEMENT

Introduction

We suppose that a set contains n objects, and we wish to draw r < n objects from the set. The order in which the objects are drawn may or may not be important. In addition, it is possible that a drawn object is replaced or not replaced before the next one is drawn. So basically, in this unit we will be focusing on sampling with and without replacement.

Learning Outcomes

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

1. Describe sampling with and without replacement
2. Describe hypergeometric distribution
3. Describe the mean and variance of hypergeometric distribution
4. Solve some numerical problems

Main Content

Sampling with and without replacement

3mins

Definition: Sampling with replacement occurs when an object is selected and then replaced before the next one is selected. By multiplication rule, the number of possible selections of r objects from n distinct objects when sampling is done without replacement is n^r

Sampling with replacement can be exemplified by the toss of a fair dice twice with the number of distinct outcomes $6^2 = 36$

These can be listed as follows

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Activity 1

Three-digit numbers are to be selected from numbers 1, 2, 3
(a) List all the possible numbers when sampling is done with replacement
(b) If a three-digit number is selected at random, what is the probability that they are the same number?

Solution:
There 3^3 = 27 possible numbers, and these are as listed below

(111), (112),(113),(121), (122) ,(123),(131), (132),(133)

(211), (212),(213),(221), (222), (223),(231), (232) ,(233)

(311) ,(312), (313),(321), (322),(323),(331) ,(332) ,(333)

(a) Let B = event that the three digits are the same i.e {(111), (222), (333)}
Pr(B) = 3/27 =1/3

Definition: Sampling without replacement occurs when the object is not replaced after it has been selected. The number of possible of selection of r objects from n objects when sampling is without replacement is

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

Illustrative example 1: Consider a lot consisting of 100 fuses which is inspected by the following procedure. Suppose that 20 of the fuses are bad and 80 are good. Five fuses are chosen at random and tested, if all five "blow" at correct amperage, the lot is accepted. If X is a random variable equal to the number of bad fuses in the sample of five, the probability of accepting the lot is

$$\Pr(X=0) = \frac{\binom{20}{0}\binom{80}{5}}{\binom{100}{5}} = 0.32$$

In a more general representation, suppose that X is one of the numbers 0, 1, 2, 3 , 4, 5; then

$$\Pr(X=x) = \frac{\binom{20}{x}\binom{80}{5-x}}{\binom{100}{5}}, \text{ x = 0, 1, 2, 3, 4, 5}$$

Hypergeometric Distribution

When sampling is done without replacement, the binomial distribution does not give exact probabilities, since the trials are not independent, but in hypergeometric distribution, the trials are not independent. The smaller the number of the population, the less accurate the binomial probability will be.

Illustrative Example
Suppose that a day's production of 850 manufactured parts contains 50 parts that do not conform to customer requirements. Two parts are selected at random, without replacement from the batch. Let the random variable **X** equal the number of nonconforming parts in the sample. We can define the events as follows: **A** and **B** denote the events that the first and second parts are nonconforming respectively. The knowledge that the first part is nonconforming suggests that it is less likely that the second part selected is also nonconforming.

Hypergeometric distribution can be conceptualized as follows. Consider a collection of $N = N_1+N_2$ similar objects in which N_1 of them belong to one of two dichotomous classes and N_2 of them belong to the second class. Suppose that a sample of n objects is selected from these N objects at random without replacement. We can find the probability that x of these n objects (where the integer x satisfies $x \le r$, $x \le N_1$ and $r - x \le N_2$) belongs to the first class and n – x belongs to the second class. In doing this, we can select x from the first class in $\binom{N_1}{x}$ ways and n-x from the second class in $\binom{N_2}{n-x}$ ways. By the multiplication principle, the product $\binom{N_1}{x}\binom{N_2}{n-x}$ equals the number of ways the joint operation can be performed. If the n objects can be selected from $N = N_1+N_2$ in $\binom{N}{n}$ ways, then the probability of selecting exactly x objects from $\binom{N_1}{x}$ and n-x from $\binom{N_2}{n-x}$ is given by

$$\Pr(X=x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$$

The Mean and Variance of Hypergeometric Random

3mins

From a sample of n objects drawn at random, one at a time from a collection of N objects of which N_1 are of one kind and $N_2 = N - N_1$ are of another kind. We think of the first kind as "success", and coded 1; the other kind as "failure "and coded 0. Let $X_1, \ldots, X_n$ denote the sequence of coded outcomes, that is X_i is 0 or 1 according to whether the i^{th} draw results in success or failure The total number (S_n) of successes in n trials is just the sum of the X's. $S_n = X_1+X_2+\ldots+X_n$ as it is in the case of independent Bernoulli trials. It is also known that X's, are identically distributed, meaning that the probability of "p" of a 1 is the same at each trial.

$$p = \Pr(X=1) = \frac{N_1}{N}$$

The Mean and Variance of the hypergeometric random variable can be determined from the Bernoulli trials that comprise the experiment.

Thus $E(S_n) = E(X_1+X_2+\ldots+X_n) = p + p + \ldots + p$

$$= np = n.\frac{N_1}{N}$$

From the above, it is noted that the mean of hypergeometric random variable is similar to that of binomial random variable, which implies that whether the object the expected number of type "1" among the objects drawn is the same. One thus obtain the mean of X as

The variance of X is also given as $\text{Var}(X) = np(1-p)\left(\frac{N-n}{N-1}\right)$

This differ from the binomial variable by the term $\left(\frac{N-n}{N-1}\right)$

which is referred to as the finite population correction factor. Sampling with replacement is equivalent to sampling from an infinite set because the proportion of success remains constant for every trial in the experiment. As mentioned previously, if sampling were done with replacement, **X** would be a binomial random variable and its variance would be **np(1-p)**. Consequently, the finite population correction represents the correction to the binomial variance that results because the sampling is without replacement from the finite set of size N.

Activity 2

Suppose that X has a hypergeometric distribution with N =20, N_1 =4 and n = 4, find the following probabilities

(a) Pr(X =1)

(b) Pr(X =4)

(c) Pr(X ≤ 2)

(d) Determine the mean and variance of X

Solution:

Let the pdf be given by $\Pr(X=x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$, $x = 0, 1, \ldots, 4$

(a) $$\Pr(X=1) = \frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}}$$

$$= \frac{4(16)(15)(14)/6}{20(19)(18)(17)/24} = 0.4623$$

(b) $$\Pr(X=4) = \frac{\binom{4}{4}\binom{16}{0}}{\binom{20}{4}}$$

$$= \frac{1}{20(19)(18)(17)/24} = 0.00021$$

(c) $$\Pr(X \le 1) = \frac{\binom{4}{0}\binom{16}{4}}{\binom{20}{4}} + \frac{\binom{4}{1}\binom{16}{3}}{\binom{20}{4}} + \frac{\binom{4}{2}\binom{16}{2}}{\binom{20}{4}}$$

$$= \frac{(16)(15)(14)(13)/24 + 4(16)(15)(14)/6 + 6(16)(15)/2}{20(19)(18)(17)/24}$$

$= 0.9866$

(d) The mean of X

Obtain $p = N_1/N =$

$= 4/20 = 0.2$

Thus $E(X) = np$

$= 4(0.2) = 0.8$

$$\text{Variance of X} = 4(0.8)(0.8)\left(\frac{20-4}{20-1}\right)$$

$$4(0.8)(0.8)\left(\frac{16}{19}\right) = 0.539$$

Activity 3

A lot of 10 compressor tanks contains 3 non-conforming and 7 conforming articles. Suppose that 4 articles are drawn from the lot without replacement. The probability function for the number of non-conforming articles in the sample of 4 is

$$f(x) = \frac{\binom{3}{x}\binom{7}{4-x}}{\binom{10}{4}}, \; x = 0, 1, \ldots, 3$$

If X is the number of non-conforming articles in the sample, then

(a) Verify that f(x) is indeed a pdf

(b) Compute the mean of X

Solution:

We first obtain Pr(X = x) for x = 0,1,2,3 as follows

$$f(0) = \frac{\binom{3}{0}\binom{7}{4}}{\binom{10}{4}} = \frac{35}{210}$$

$$f(1) = \frac{\binom{3}{1}\binom{7}{3}}{\binom{10}{4}} = \frac{105}{210}$$

$$f(2) = \frac{\binom{3}{2}\binom{7}{2}}{\binom{10}{4}} = \frac{63}{210}$$

$$f(3) = \frac{\binom{3}{3}\binom{7}{1}}{\binom{10}{4}} = \frac{7}{210}$$

(a) f(x) > 0 for every x and

$$f(0)+f(1)+f(2)+f(3) = \frac{35}{210} + \frac{105}{210} + \frac{63}{210} + \frac{7}{210} = 1$$

Thus f(x) is pdf

$$\text{(b) Mean} = 0 \times \frac{35}{210} + 1 \times \frac{105}{210} + 2 \times \frac{63}{210} + 3 \times \frac{7}{210}$$

$$\frac{105}{210} + \frac{126}{210} + \frac{21}{210} + \frac{252}{210} = 1.2$$

Alternative method: Given the following information, N =10, N_1=3, N_2=7 n=4

Obtain p =N_1/N=

= 3/10,

so that mean = np = 4(3/10) = 1.2

Summary

In this unit, we had done sampling with and without replacement and its defination. Hypergeometric distribution was defined from the background of sampling without replacement and contrasted with binomial distribution which was defined from sampling with replacement background. The mean and variance of hypergeometric random variable were briefly described. Numerical problems relating to the distribution were solved.

Self Assessment Questions

1. Differentiate between sampling with and without replacement
2. Suppose that 4 balls are drawn at random, one at a time without replacement from 10 balls where 3 are black and 7 are white. What is the probability that the third ball drawn is black?
3. Suppose that seven balls are selected at random without replacement from a box containing five red and ten blue balls. If $\overline{X}$ denotes the proportion of red balls in the sample, what are the mean and variance of $\overline{X}$?
4. Suppose that two numbers are selected from numbers 1 through 10 without replacement
   (a) List the elements of the sample space
   (b) What is the probability that the sum of the two numbers is even
   (c) What is the probability that the first number is a prime number
5. Suppose that X has a hypergeometric distribution with N =25, N_1 = 5 and n = 4, find the following probabilities

(a) Pr(X =1)

(b) Pr(X =4)

(c) $Pr(X \leq 2)$
(d) Determine the mean and variance of X

Tutor Marked Assignment

1. A lot contains 12 manufactured items of which 3 are good and 9 are bad. Three items are selected at random without replacement to see whether there are any bad ones among them. If at least one item is bad, the lot is rejected. Find the probability that the lot will be rejected.
2. A bag contains 4 white, 5 red and 7 black balls. If three balls are drawn at random without replacement, what is the probability that one is white, one is red and one is black?
3. Consider an experiment involving selecting two numbers from integers 1, 2, 3, 4, 5, 6 without replacement. Let i and j denote the first and the second selections respectively, find $Pr(3 \leq i+j < 7)$

References

1. Lindgren, B. W. Statistical Theory. Third Edition. Macmillan Publishing Co. Inc. New York
2. Montgomery, D.C. and Runger, G.C. (2003). Applied Statistics and Probability for Engineers.Third Edition. *John Wiley & Sons, Inc.*

Further Reading

1. Gupta, S. C. Fundalmentals of Statistics. Himalaya Publishing House

Source: pexels.com

MODULE 2

Unit 1: Conditional probability
Unit 2: Baye's theorem
Unit 3: Independent events

UNIT 1 CONDITIONAL PROBABILITY

Introduction

Suppose you are to perform an experiment for which the sample space of all possible outcomes is S and also that probabilities have been specified for all the events in S. We can then study the way in which the probability of an event, say A changes after it has been learnt that some other event say B has occurred. This unit we will deal with this subject matter of probability.

Learning Outcomes

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

1. Define conditional probability of events and illustrate with examples
2. Demonstrate conditional probabilities with numerical examples
3. Describe the multiplication rule of probability
4. Demonstrate multiplication rule of probability with numerical examples.

Main Content

Conditional Probability

7mins

We shall start the discussion of conditional probability with an illustrative example. Suppose that in a manufacturing process, 10% of the parts contain surface flaws and 25% of the parts with surface flaws are defective parts. However, only 5% of parts without surface flaws are defective parts. Here, the probability of a defective part depends on our knowledge of the presence or absence of a surface flaw. This example illustrates that probabilities need to be re-evaluated as additional information becomes available. The notation and details are further illustrated for this example.

Let **D** denote the event that a part is defective and let **F** denote the event that a part has a surface flaw. Then, we denote the probability of **D** given, or assuming, that a part has a surface flaw as Pr(D|F). This notation is read as the conditional probability of **D** given **F**, and it is interpreted as the probability that a part is defective, given that the part has a surface flaw.

Definition
Let S be a sample space of a random experiment and B be an event defined on S such that Pr(B)>0. Suppose also that A is any arbitrary event in S, then the conditional probability of event A occurring given that B has occurred is given as

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

The conditional probability given above is not defined if Pr(B) = 0
This definition can be understood in a special case in which all outcomes of a random experiment
are equally likely. If there are n total outcomes,

Pr(B) = (number of outcomes in B)/n

Also, $\Pr(A \cap B)$=(number of outcomes in $A \cap B$)/n

Consequently,

$$\frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\text{numberof outcomes in } (A \cap B)}{\text{numer of outcomes in B}}$$

The conditional probability $\Pr(A|B)$ has a simple meaning in terms of the frequency interpretation given in unit 1 above. According to the interpretation, if an experimental process is repeated a large number of times, then the proportion of repetitions in which the event B will occur is approximately Pr(B) and the proportion of repetitions in which both the events A and B will occur is approximately $\Pr(A \cap B)$

Therefore, among those repetitions in which the events B occurs, the proportion of repetitions in which event A will also occur is approximately equal to

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

This definition can be understood in a special case in which all outcomes of a random experiment
are equally likely. If there are n total outcomes,

Activity 1

Suppose that a coin is tossed twice. If we assume that all the four points in the sample space are equally likely, what is the conditional probability that both tosses result in heads given that the first toss results in a head?

Solution: The sample space of the experiment is given as
S = {HH, HT, TH, TT} and the number of elements in the sample space is
n(S) = 4

Let A be the event that both tosses result in heads so that A={HH}, and Let B be event that first toss results in a head so that B={HH, HT},

From the definition of conditional probability, we also need event $A \cap B$, which is the intersection between A and B and this is
$A \cap B$ ={HH}
Let n($A \cap B$) denote number of elements in $A \cap B$ and n(B) denote the number of elements in B
Then the required probability is $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$

$$= \frac{n(A \cap B)}{n(S)} \bigg/ \frac{n(B)}{n(S)}$$

$$= \frac{1/4}{2/4} = \frac{1}{2}$$

Condition probability can be calculated from information (outcomes of an experiment) that are displayed in tabular form. For example, consider the table below which classifies information (outcomes) according to two factors A and B. Factor A has 3 levels (A_1, A_2, A_3) and factor B has 2 levels (B_1, B_2).

Factor A	**Factor B**		**Total**
	B_1	B_2	
A_1	$n(A_1 \cap B_1)$	$n(A_1 \cap B_2)$	$n(A_1)$
A_2	$n(A_2 \cap B_1)$	$n(A_2 \cap B_2)$	$n(A_2)$
A_3	$n(A_3 \cap B_1)$	$n(A_3 \cap B_2)$	$n(A_3)$
Total	$n(B_1)$	$n(B_2)$	N

From the table n is the total number of observations which is equivalent to the sample space n(A1∩B1) is the number of outcomes classified jointly by factor levels A1 and B1. Others are similarly defined. n(A1) is the marginal total of outcomes in factor level A1. Others are similarly defined.

Activity 2

A student is to be selected at random from a group of freshers using their new matriculation numbers. Suppose that the students' population is made up according to their faculty and gender as follows

Faculty	Gender		Total
	Male	**Female**	
Agric	89	23	112
Arts	201	74	275
Social Sciences	52	41	93
Education	193	85	278
Engineering	175	4	179
Health Sciences	45	20	65
Physical and Life Sciences	198	53	251
Total	953	300	1253

(a) What is the probability that a randomly selected student is a male and from faculty of Education?

(b) If a student from the faculty of Agric is known to have been selected at random, what is the probability that he is a Female?

Solution:

The sample space here is the total number of students which is 1253

(a) Let M: event that the student is a male and n(A) denotes the number of elements in A

E: event that the student is from education and n(B) is the number of elements

The required probability is $Pr(M \cap E) = \frac{193}{1253}$

$= 0.154$

(b) Let A: the student is from faculty of Agric

F: the student is a Female

$F \cap A$: the student is a female and from the faculty of Agric and the number of elements is $n(F \cap A) = 23$, $n(A) = 112$

Required to find $Pr(F|A) = \frac{n(F \cap A)}{n(S)} \Big/ \frac{n(A)}{n(S)}$

$= \frac{23/1253}{112/1253} = \frac{23}{112}$

$= 0.205$

Multiplication Rule

Recall the definition of the conditional probability of event A_1 given that event A_2 has occurred given

$$Pr(A_1|A_2) = \frac{Pr(A_1 \le A_2)}{Pr(A_2)}$$

We can write the definition of conditional probability above to provide a general expression for the probability of the intersection of two events. This formula is referred to as a **multiplication rule** for probabilities. And we have

$$Pr(A_1 \cap A_2) = Pr(A_1|A_2)Pr(A_2)$$

and

$$Pr(A_1 \cap A_2) = Pr(A_2|A_1)Pr(A_1)$$

The last expression is obtained by interchanging A_1 and A_2

For example, the formula

$$Pr(A \cap B) = Pr(A|B)Pr(B)$$

Follows immediately from the definition of conditional probability and allows one to compute $Pr(A \cap B)$ from the knowledge of $Pr(A|B)$ and Pr(B).

Activity 3

Suppose that the probability that an automobile battery subject to high engine compartment temperature suffers low charging current is 0.30. The probability that a battery is subject to high engine compartment temperature is 0.45. What is the probability that a battery is subject to low charging current and high engine compartment?

Solution:

Let A_1 denote the event that a battery suffers low charging current, and let A_2 denote the event that a battery is subject to high engine compartment temperature. The probability that a battery is subject to low charging current and high engine compartment temperature is

$$Pr(A_1 \cap A_2) = Pr(A_1|A_2)Pr(A_2)$$

$$0.30 \times 0.45 = 0.135$$

Activity 4

One bag contains 4 white and 3 black balls, and a second bag contains 3 white and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. What is the probability that a ball now drawn from the second bag is black?

Solution:

Let B1, B2 and W1 represent, respectively the drawing of a black ball from bag 1, a black ball from bag 2 and a white ball from bag. Interest is on the union of the mutually exclusive events $B_1 \cap B_2$ and $W_1 \cap B_2$

Now $\Pr(B_1 \cap B_2)$ or $\Pr(W_1 \cap B_2) = \Pr(B_1 \cap B_2) + \Pr(W_1 \cap B_2)$

$$= \Pr(B_1)\Pr(B_2 | \,)) + \Pr(W_1 \cap B_2)$$

$$= \Pr(B_1)\Pr(B_2 | B_1) + \Pr(W_1)\Pr(B_2 | W_1)$$

$$= \left(\frac{3}{7}\right)\left(\frac{6}{9}\right) + \left(\frac{4}{7}\right)\left(\frac{5}{9}\right) = \left(\frac{38}{63}\right)$$

Summary

This unit we have covered the conditional probability of two or several events. Independence of two and several events was also highlighted and some numerical examples were given for illustration. The law of total probability was stated.

Self Assessment Questions

1. Define the conditional probability of event A given B
2. One batch of items contains 5 defective and 4 good items, and a second bag contains 4 defective and 6 good items. One item is drawn from the first batch and placed unseen in the second batch. What is the probability that an item now drawn from the second batch is good?
3. Disks of polycarbonate plastic from a supplier are analyzed for scratch and shock resistance. The results from 100 disks are summarized as follows:

		shock resistance	
		high	*low*
scratch	*high*	70	9
resistance	*low*	16	5

Let A denote the event that a disk has high shock resistance, and let B denote the event that a disk has high scratch resistance. Determine the following probabilities

(a) Pr(A)
(b) Pr(B)
(c) Pr(A/B)

4. Define the multiplication rule of probability
   Suppose that the probability that uncle Jones parks in a no-parking zone in the school and gets a parking ticket is 0.06, and the probability that Jones can find a legal parking space and has to pack in the no-parking zone is 0.20. If Jones arrives at school and has to pack in a no-parking zone, what is the probability that he will get a parking ticket?
5. Suppose that the probability that a student who offered further maths passes physics is 0.45, the probability that a student offer physics 0.25. What is the probability that a student offers further maths and passes physics?

Tutor Marked Assignment

1. Suppose that the probability that a regularly schedule fight departed on time is P(D)=0.83; the probability that it arrived on time is Pr(A) = 0.75; and the probability that it departed and arrived on time is $\Pr(D \cap A) = 0.57$. Find the probability that a plane
   (a) departed on time given that it arrived on time
   (b) arrived on time given that it departed on time
2. Consider the following events defined on an experiment involving tossing two coins together
   A_1 : at least one head turns up
   A_2 : exactly two heads turn up
   Find $\Pr(A_2 | A_1)$
3. The following are the numbers of registered members in two parties (APC and PDP) by gender (Male and Female) in a ward: (APC, Male)=40, (PDP, Male)=10, (APC, Female)=20 and (PDP, Female)=30. If a party member is selected at random, what is the probability of selecting an APC member given that he is a male?
4. Suppose that Pr(A/B)=0.4 and Pr(B)=0.5, find $\Pr(A \cap B)$
5. Suppose that two dice are rolled. If we define the following events.
   E_1: 2 appears on at least one die
   E_2: sum of numbers appearing 6
   Find $\Pr(E_1 | E_2)$
6. A market survey was conducted in four cities to find out the preference for Milo over Bournvita. The responses are shown below

	City 1	City 2	City 3	City 4	Total
Yes	45	55	60	50	210
No	35	45	35	45	160
No opinion	5	5	5	5	100
Total	85	105	100	100	20

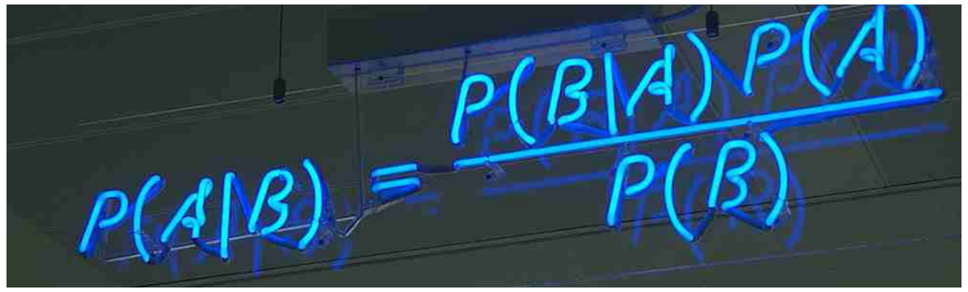
   (a) What is the probability that a consumer prefers Milo given that he is from City 4?
   (b) Given that a consumer prefers Milo, what is the probability that he is from city 2?

References

1. Lindgren, B. W. Statistical Theory Macmillan Publishing Co. Inc. New York
2. Hogg, R.V. and Craig, A. T. Introduction to Mathematical Statistics. Macmillan Publishing Co. Inc. New York

Further Reading

Gupta, S. C. Fundamentals of Statistics. Himalaya Publishing House.
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=2ahUKEwjg4qSL9tjnAhUFyYUKHS2NAPIQFjADegQIChAB&url=https%3A%2F%2Focw.mit.edu%2Fcourses%2Fmathematics%2F18-05-introduction-to-probability-and-statistics-spring-2014%2Freadings%2FMIT18_05S14_Reading3.pdf&usg=AOvVaw0u_kkShuxvMr5EyxK4lHNL



UNIT 2 BAYE'S THEOREM

Introduction

Often times, we do not have complete information about some events, but we might know one conditional probability but we would like to calculate a different one. Also, conditional probabilities are not only interesting as new probabilities given some additional information, they may also be used as tools in the computation of unconditional probabilities. So in this unit we will focus on stating and proving Baye's theorem, we will define multiplication rule of probability, and also use it to solve some numerical problems.

Learning Outcomes

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

1. Describe the law of total probability
2. State and prove Bayes's theorem
3. Calculate some numerical problems based on Baye's theorem

Main Content

Law of Total probability

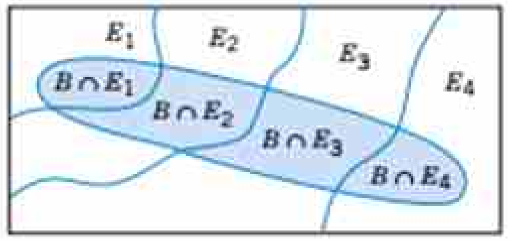
7mins

I will like you to understand the concept of total probability by first considering the partitioning a sample space into two mutually exclusive events A and A'. Suppose that B is any other event such that Pr(B)>0 which intersects and . Since and are disjoint, then we can write B as the union of the part of B in A and the part of B in . That is,
$B = (B \cap A) \cup (B \cap A')$. Thus the probability of B can be written as the sum of the probabilities of mutually exclusive parts ($A \cap B$) and ($A' \cap B$)
That is

$$\Pr(B) = \Pr(B \cap A) + \Pr(B \cap A')$$
$$= \Pr(B|A)\Pr(A) + \Pr(B|A')\Pr(A')$$

Partitioning of a sample space into more than two events

Definition: Let $E_1, E_2, \ldots, E_k$, k = 2, 3,..., be events which are mutually exclusive such that $E_i \cap E_j = \phi$ for $i \neq j$ and collectively exhaustive such that $S = E_1 + E_2 + \cdots + E_k$. Under these conditions, $E_1, E_2, \ldots, E_k$ are called partitions of sample space S. The figure below is an example of partition of S into four events.



$B=(B\cap E_1)\cup(B\cap E_2)\cup(B\cap E_3)\cup(B\cap E_4)$

Patition of sample space into four mutually exclusive events

Proof of Baye's Theorem

Let B be any other event such that Pr(B)>0 and

$$B = B\cap(E_1+E_2+E_3+E_4)$$
$$= (B\cap E_1)+(B\cap E_2)+(B\cap E_3)+(B\cap E_4)$$

From the law of total probability,

$\Pr(B) = \Pr(B\cap E_1)+\Pr(B\cap E_2)+\Pr(B\cap E_3)+\Pr(B\cap E_4)$ (I)

From the definition of conditional probability,

$\Pr(B\cap E_j) = \Pr(B|E_j)\Pr(E_j)$, j=1, 2, 3,4 (II)

so that (I) can be written as

$\Pr(B) = \Pr(B|E_1)\Pr(E_1)+\Pr(B|E_2)\Pr(E_2)+\cdots+\Pr(B|E_4)\Pr(E_4)$ (III)

The expression on the left of (II) can also be written as

$\Pr(B\cap E_j) = \Pr(E_j|B)\Pr(B)$ (IV)

Equating (II) and (IV) we have

$\Pr(E_j|B)\Pr(B) = \Pr(B|E_j)\Pr(E_j)$, from which

$$\Pr(E_j|B)=\frac{\Pr(B|E_j)\Pr(E_j)}{\Pr(B)}$$
$$=\frac{\Pr(B|E_j)\Pr(E_j)}{\Pr(B|E_1)\Pr(E_1)+\Pr(B|E_2)\Pr(E_2)+\ldots\Pr(B|E_4)\Pr(E_4)}$$

For example,

$$\Pr(E_1|B)=\frac{\Pr(B|E_1)\Pr(E_1)}{\Pr(B|E_1)\Pr(E_1)+\Pr(B|E_2)\Pr(E_2)+\ldots\Pr(B|E_4)\Pr(E_4)}$$

From the above, the probabilities $\Pr(E_1)$, $\Pr(E_2)$,..., $\Pr(E_4)$ which are already given or known before conducting an experiment are termed a priori or prior probabilities. The conditional probabilities $\Pr(E_1|B)$, $\Pr(E_2|B)$,..., $\Pr(E_4|B)$, which are computed after conducting the experiment are called the posterior probabilities.

Activity 1

Suppose that 10% of the drivers in a city are incompetent. Suppose also that a diagnostic test is available which is 90% effective as follows. If a driver is incompetent, the probability that the test will so indicate is 0.9, and if a driver is competent, the probability that the test will also so indicate is 0.9. Given that the test is effective, what is the probability that the driver is in fact incompetent?

Solution:

We define the following events from the question

E_1: event that a driver is incompetent

E_2: event that the driver is competent.

B: event that the diagnostic test is effective.

The following probabilities are given

$\Pr(E_1) = 0.1$

$\Pr(E_2) = 0.9$,

$\Pr(B|E_1) = \Pr(B|E_2)=0.9$

We are required to find $\Pr(E_1|B)$

This is given as

$$\Pr(E_1|B)=\frac{\Pr(B|E_1)\Pr(E_1)}{\Pr(B|E_1)\Pr(E_1)+\Pr(B|E_2)\Pr(E_2)}$$
$$=\frac{0.9\times0.1}{(0.9\times0.1)+(0.9\times0.9)} = 0.1$$

Activity 2

The No-work-no-pay move by the Federal Government has generated a hot debate among the general public, leading to the following probabilities. The probability that a randomly selected person is a university student is 1/20; the probability that a university student actually supports the move is 1/10; a non-university student supports the move with probability 7/10. If a randomly selected person supports the move, what is the probability that he is a university student?

Solution:

Suppose we define the following events

U: The selected person is a university student

S: The selected person supports the move

We have the following probabilities

$\Pr(U) = 1/20$, $\Pr(S|U)=1/10$, $\Pr(S'|U)=7/10$

Required to find $\Pr(U|S)$

$$\Pr(U|S)=\frac{\Pr(U\cap S)}{\Pr(S)}$$
$$=\frac{\Pr(S|U)\Pr(U)}{\Pr(S|U)\Pr(U)+\Pr(S'|U)\Pr(U)}$$

$$= \frac{1/10 \times 1/20}{(1/10 \times 1/20) + (7/10 \times 1/20)}$$
$$= 0.125$$

Activity 3

Each of the three identical jewelry boxes has 2 drawers. In each drawer of the first box, there is a gold watch. In each drawer of the second box there is a silver watch. In one drawer of the third box, there is a gold box while in the other drawer, there is a silver watch. If we select a drawer at random, open one of the drawers and find it to contain a silver watch, what is the probability that the other drawer has the gold watch?

Solution:
We shall define the following events
Ai: i^{th} box is selected; (i =1, 2, 3)
B: The open drawer of the selected box contains a silver watch
We have the following
$Pr(A_1) = Pr(A_2) = Pr(A_3) = 1/3$
$Pr(B \mid A_1) = 0$; $Pr(B \mid A_2) = 1$; $Pr(B \mid A_3) = ½$
We are given that one the drawers of the selected box contains a silver watch and we want that the second drawer of the box contains a gold watch. This is possible if only if the third box is selected because it is only box that contains the silver and gold watches in its drawers. We are thus required to find the probability $Pr(A_3|B)$
By Baye's theorem this is given as

$$Pr(A_1|B) = \frac{Pr(B|A_1)Pr(A_1)}{Pr(B|A_1)Pr(A_1) + Pr(B|A_2)Pr(A_2) + Pr(B|A_3)Pr(A_3)}$$

$$Pr(A_1|B) = \frac{\frac{1}{2} \times \frac{1}{3}}{0 \times \frac{1}{3} + 1 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3}} = \frac{1}{3}$$

Activity 1

Two sets of candidates are competing for the position on the Board of Director of a Company. The probabilities that the first and the second sets will win are 0.6 and 0.4 respectively. If the first set wins, the probability of introducing a new product is 0.8 and the corresponding probability of the second set wins is 0.3. What is the probability that the product will be introduced?

Solution:
Let A_1, A_2, denote the events that the first and second sets of candidates respectively win the position and let B be the event of introducing a new product.
We are given
$Pr(A_1) = 0.6$; $Pr(A_2) = 0.4$; $Pr(B \mid A_1) = 0.8$; $Pr(B \mid A_2) = 0.3$
The event B can be described in the following mutually exclusive ways:
(I) 1^{st} set wins and the new product is introduced i.e $Pr(A_1 \cap B)$

(II) 2^{nd} set wins and the new product is introduced i.e $Pr(A_1 \cap B)$
Thus by addition theorem of probability we have

$$Pr(B) = Pr(A_1 \cap B) + Pr(A_2 \cap B)$$
$$= Pr(B \mid A_1)Pr(A_1) + Pr(B \mid A_2)Pr(A_2)$$
$$= 0.8 \times 0.6 + 0.3 \times 0.4 = 0.48 + 0.12 = 0.6$$

Summary

The concept of multiplication rule of conditional probability has been given and illustrative example was Provided to you in this unit. The law of total probability was stated when sample space is partitioned into non-overlapping events. Baye's theorem was described as an extension of conditional probability. The formula for solving problems relating to Baye's theorem was derived, and real live examples were provided for its illustration.

Self Assessment Questions

1. Define the law of total probability
2. State and prove Baye's theorem
3. In a rubber factory, machines A, B, C manufacture respectively 20%, 30% and 50% of the total of its outputs. Of them, 5, 4 and 2% respectively are defective bolts. Suppose that a bolt is drawn at random from the product and it is found to be defective. What is the probability that it was manufactured by machine B?
4. A company has two plants that manufacture generators. Plant I manufactures 70% of the generators and plant II manufacture 30% of the product. 80% of the generators manufactured by plant I meets the specified quality while 90% of those manufactured by plant II meets the specified quality. A generator is picked at random and found to meet the specified standard, what is the probability that it is manufactured by plant II
5. A bag I contain 4 white and 6 black balls while another Bag II contains 4 white and 3 black balls. One ball is drawn at random from one of the bags, and it is found to be black. Find the probability that it was drawn from Bag I.

Tutor Marked Assignment

1. Three automatic machines produce a similar automobile parts. Machine 1 produces 40% of the total, machine 2 produces 25% while machine 3 produces the rest. On the average, 10% of parts produced by machine 1 are defective while for machines 2 and 3, the percentage of defective parts are 5% and 1% respectively. What proportion of parts produced by the three machines is defective?
   If a part is selected at random from the combined output and found to be defective, what is the probability that it is produced by machine 1?
2. In a yearly examination being conducted before admission into a post- secondary institution, a candidate has a chance of 3 out of 5 of passing the examination if question 1 is made compulsory. However, if question 1 is made optional, the candidate has a chance of 3 out of 4 of passing the examination. Record shows that question 1 has been made compulsory 3 times in the past 10 examinations conducted by the institution
   (a) What is the probability that the candidate passes the examination?
   (b) If it is known that a candidate has passed the examination, what is the probability that question 1 is made compulsory?
3. In the production of certain electronic devices, 2% of the device produced is defective. A quick test detects a defective device with probability 95%; however, with probability 10% it gives a false alarm for an intact device.
   If the test gives an alarm, what is the probability that the device just tested is indeed defective?
4. An insurance company insured 2000 scooter drivers, 4000 car drivers and 6000 truck drivers. The probability of an accident involving a scooter driver, car driver and a truck driver is 0.01, 0.03 and 0.15 respectively. If one of the insured person meets with an accident, what is the probability that he is a scooter driver?

References

1. Montgomery, D.C. and Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. John Wiley & Sons, Inc.
2. Soong, T. T Fundamentals of Probability and statistics for Engineers.(2004). John Wiley & Sons Inc.

Further Reading

Gupta, S. C. Fundamentals of Statistics. Himalaya Publishing House.
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=2ahUKEwjg4qSL9tjnAhUFyYUKHS2NAPIQFjADegQIChAB&url=https%3A%2F%2Focw.mit.edu%2Fcourses%2Fmathematics%2F18-05-introduction-to-probability-and-statistics-spring-2014%2Freadings%2FMIT18_05S14_Reading3.pdf&usg=AOvVaw0u_kkShuxvMr5EyxK4lHNL

UNIT 3 INDEPENDENT EVENTS

Introduction

Having understood conditional probability from the previous unit, we can now proceed to independent events in this unit. Suppose that two events occur independently of each other in the sense that the occurrence or non-occurrence of either of them has no influence on the occurrence or non-occurrence of each other, then it is natural to assume that the joint probability of both of them occurring is equal to the product of their individual probabilities.

Learning Outcomes

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

1. Define independence of two or more events
2. Demonstrate several events that could be independent
3. Compute probabilities of events that are independent

Main Content

Independence of two events

3mins

Consider the conditional probability of event A_1 given that event A_2 has occurred given by

$$\Pr(A_1|A_2) = \frac{\Pr(A_1 \cap A_2)}{\Pr(A_2)}$$

If the occurrence of event A_1 does not depend on the occurrence event A_2 (that is the probability that A_2 occurs does not change the probability that A_1 occurs), then event A_1 will be said to be independent of event A_2.

Definition: Two events A and B are said to be statistically independent or simply independent if the probability corresponding to their intersection can be written as the product of their marginal probability $(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2)$

Example

Suppose that A is the event that a head is obtained when a fair coin is tossed, and B is the event that either number 1 or 2 is obtained when a balanced die is rolled. Then event A will occur with a relative

frequency of 1/2 when the coin is tossed repeatedly and the event B will occur with a relative frequency of 1/3 when the die is rolled repeatedly. Hence the relative frequency with which both A and B occur simultaneously is will be (1/2)(1/3).
Thus we have that Pr(A) = 1/2 and Pr(B) =1/3, so that

$$\Pr(A \cap B) = \Pr(A)\Pr(B) = (1/2)(1/3) = 1/6$$

Activity 1

Suppose that two machines 1 and 2 in a factory are repeated independently of each other. Let A be the event that machine 1 will become inoperative during a given period, and let B be the event that machine 2 will become inoperative in the same period. If Pr(A) = 1/3 and Pr(B) = 1/4. Determine the probability that at least one of the machines will become inoperative during the given period

Solution:
Given that
Event A: Machine 1 will become inoperative during a given period and Pr(A) =1/3
Event B: Machine 2 will become inoperative during the same period, and Pr(B) =1/4
Since the two machines are repeated independently, we have that the probability that both will become inoperative during the period is $\Pr(A \cap B) = \Pr(A)\Pr(B) = (1/3)(1/4) = 1/12$
Therefore the probability that at least one of the machines will become inoperative during the given period is

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A)\Pr(B) = (1/3) + (1/4) - 1/12 = 1/2$$

Activity 2

Consider an experiment of rolling a balanced die. Let **A** be the event that an even number is obtained, and let **B** be the event that one of the numbers 1, 2, 3, or 4 is obtained. Show whether or not events **A** and B are independent.
Solution:
Let the sample space S = {1, 2, 3,4, 5, 6}
Let event A ={2,4,6} and event B = {1, 2, 3, 4}
Pr(A) = 1/2 and Pr(B) = 2/3.
The intersection of A and B is A ∩ B = {2,4}
Then Pr(A ∩ B) = 1/3 and Pr(A)Pr(B) = (1/2)(2/3) =1/3
Thus it follows that A and B are independent
It can also be stated that if events A and B are independent, then the following events are also independent
(a) A and B' (b) B and A′

Proof

(a) $A \text{ and } B' = A \cap B' = A \cap (S - B)$

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

$= A - A \cap B$

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

$= \Pr(A) - [\Pr(A)\Pr(B)]$

$= \Pr(A)[1 - \Pr(B)]$

$= \Pr(A)\Pr(B')$

(b) Similar to (a),

$A' \text{ and } B = A' \cap B$

$= (S - A) \cap B$

$= B - A \cap B$

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

$= \Pr(B) - [\Pr(A)\Pr(B)]$

$= \Pr(B)[1 - \Pr(A)]$

$= \Pr(B)\Pr(A')$

(c) $A' \cap B' = (S - A) \cap (S - B)$

$= (S - B - A + A \cap B)$

$\Pr(A' \cap B') = \Pr(S - B - A + A \cap B)$

$= \Pr(S) - \Pr(B) - \Pr(A) + \Pr(A \cap B)$

$= [\Pr(1 - \Pr(B)] - \Pr(A)[1 - \Pr(B)]$

$= [\Pr(1 - \Pr(A)][1 - \Pr(B)]$

$= [\Pr(1 - \Pr(A)][1 - \Pr(B)]$

$= \Pr(A')\Pr(B')$

Independence of three events

2mins

The notion of independence can be extended to three events. Thus an appropriate definition of the independence of three events A, B, and C would have to go further than merely assuming that all of the ${}^3C_2$ pairs of events are independent. We are thus led to the following definition.
Events A_1, A_2, and A_3, independent if and only if

(i) $\Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2)$

(ii) $\Pr(A_1 \cap A_3) = \Pr(A_1)\Pr(A_3)$

(iii) $\Pr(A_1 \cap A_2) = \Pr(A_1)\Pr(A_2)$

(iv) $\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1)\Pr(A_2)\Pr(A_3)$

The above four relations are thus necessary for the characterization of independence of A1, A2, A3
Conditions (i) –(iii) are termed pairwise independence.

Pairwise independence or independence in pairs means that each combination of two of the three events is a set of independent events. Independence of three events should be treated with caution. This shall be illustrated as follows.

(I) Pairwise independence does not imply condition

Example 1

Let the sample space of a random experiment be given by
S ={1, 2, 3, 4}, such that each element of the sample space has equal probability
Pr(1) = Pr(2) = Pr(3) =Pr(4) = ¼
Suppose we define the following events
$E_1=\{1,2\}$, $E_2=\{1,3\}$, $E_3=\{1,4\}$, the

$$(E_1\cap E_2)=(E_1\cap E_3)=(E_2\cap E_3)=\{1\},$$

and $(E_1\cap E_2\cap E_3)=\{1\}$

$$\Pr(A_1\cap A_3)=\Pr(A_1)\Pr(A_3)$$
$$\Pr(A_2\cap A_3)=\Pr(A_2)\Pr(A_3)$$

Thus $\Pr(E_1\cap E_2)=\Pr(E_1\cap E_3)=\Pr(E_2\cap E_3)=\Pr(E_1\cap E_2\cap E_3)=1/4$

Now

$$\Pr(E_1\cap E_2)=\Pr(E_1)\Pr(E_2)=(1/2)(1/2)\;=1/4\,,$$
$$\Pr(E_1\cap E_3)=\Pr(E_1)\Pr(E_{32})=(1/2)(1/2)\;=1/4\,,$$
$$\Pr(E_2\cap E_3)=\Pr(E_2)\Pr(E_3)=(1/2)(1/2)\;=1/4\,,$$

but

$$\Pr(E_1\cap E_2\cap E_3)=1/4\ ?\ \Pr(E_1)\Pr(E_2)\Pr(E_3)=(1/2)\,(1/2)\,(1/2)\;=1/8$$

(II) The fact that event **A** is independent of event **B** and also independent of event **C** does not necessarily imply that A is independent of **BC.**

Example 2

Suppose that two fair dice are thrown. Let **A** denote the event that the sum of the dice is 7. Let **B** denote the event that the first die equals 4 and let **C** be the event that the second die equals 3. Now it can be shown that **A** is independent of **B** and that **A** is also independent of **C**; but clearly **A** is not independent of **BC**

Proof

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

$(A\cap B)=\{4,3\}$, $(A\cap C)=\{4,3\}$, $(B\cap C)=\{4,3\}$, $(A\cap B\cap C)=\{4,3\}$

$$\Pr(A|B)=\frac{\Pr(A\cap B)}{\Pr(B)}=\frac{1}{36}\Big/\frac{6}{36}$$
$$=\frac{1}{6}=\Pr(A)$$

Thus **A** is independent of **B**
Also

$$\Pr(A|C)=\frac{\Pr(A\cap C)}{\Pr(C)}=\frac{1}{36}\Big/\frac{6}{36}$$
$$=\frac{1}{6}=\Pr(A)$$

Also **A** is independent of **C**
Now $\Pr(A\mid BC)=\Pr(A\mid B\cap C)$

$$\Pr(A|B\cap C)=\frac{\Pr(A\cap B\cap C)}{\Pr(B\cap C)}$$
$$=\frac{1}{36}\Big/\frac{1}{36}=1\neq\Pr(A)$$

Thus **A** is not independent of **BC**

(III) If A, B, C are independent, then A will be independent of any event formed from F and G. for example, A is independent of $(B\cup C)$

Proof:

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

By addition rule of probability

$$=\Pr(A)\Pr(B)+\Pr(A)\Pr(C)-\Pr(A)\Pr(B)\Pr(C)$$
$$=\Pr(A)[\Pr(B)+\Pr(C)-\Pr(B)\Pr(C)]$$
$$=\Pr(A)\Pr(B\cup C)$$

Activity 3

Consider events A, B, C, D defined on sample space S. Let Pr(A) = 0.8, Pr(B) = 0.68, Pr(C) = 0.9, Pr(D)= 0.75. If A, B, C, D are independent, then find
(I) Pr(A ≤ " Ψ (II) A ∩ " ᶊ (III) $\Pr(A\cup B\cup C\cup D)$
(IV) (A ∪ B′ ∪ C′ $ Ψ

Solution:

(i) $\Pr(A \cap B) = \Pr(A) \times \Pr(B)$

$= 0.8 \times 0.68 = 0.544$

(ii) $\Pr(A \cap B') = \Pr(A)\Pr(B')$

$= 0.8 \times (1 - 0.68) = 0.256$

(iii) $\Pr(A \cup B \cup C \cup D) = 1 - \Pr(A \cup B \cup C \cup D)'$

$= 1 - \Pr(A' \cap B' \cap C' \cap D')$ From De Morgan's rule

$= 1 - [\Pr(A')\Pr(B')\Pr(C')\Pr(D')]$

$= 1 - [(1-0.8) \times (1-0.68) \times (1-0.9) \times (1-0.75)]$

$= 1 - 0.0016$

$= 0.9984$

(iv) To be solved by students

Summary

In this unit, I gave the definition of independence of two or more events. Proofs were demonstrated that various events formed from independent events are also independent. Numerical problems relating to independent events were solved.

Self Assessment Questions

1. Define independence of event A given that event B had occurred
2. Show that if events A and B are independent, then B and A′are also independent
3. If Pr(A) = 0.3, Pr(B) = 0.2, Pr(C) = 0.1, and A, B, C are independent events, find the probability of occurrence of at least one of the events A, B and C.
4. Consider two events E_1 and E_2 . If the occurence or non-occurence of E_1 does not affect the probability of occurence of E_2, which of the following expressions is correct?

A. $\Pr(E_1 / E_2) = \Pr(E_2)$ B. $\Pr(E_2 / E_1) = \Pr(E_2)$

C. $\Pr(E_2 / E_1) = \Pr(E_1)$ D. $\Pr(E_1 / E_2) = \Pr(E_2 / E_1)$

Tutor Marked Assignment

1. An electronic device is made up of three components A, B and C. The probability that component A will fail in some fixed period of time is 0.01, component B will fail with probability 0.1 and component C will fail with probability 0.02. If the three components work independently of one another, find the probability that the device will work satisfactorily.

2. Two coins are tossed and the following events are defined
   A denotes the event that at most one head on the two tosses
   B denotes the event that one head and one tail in both tosses
   Are A and B independent?

3. Let A and B be independent events defined on the same sample space. Suppose they have the following probabilities: Pr(A) = x and Pr(B) = y. Find the probabilities of the following events
   (a) Neither event A nor event B occurs
   (b) Event A occurs but event B did not occur

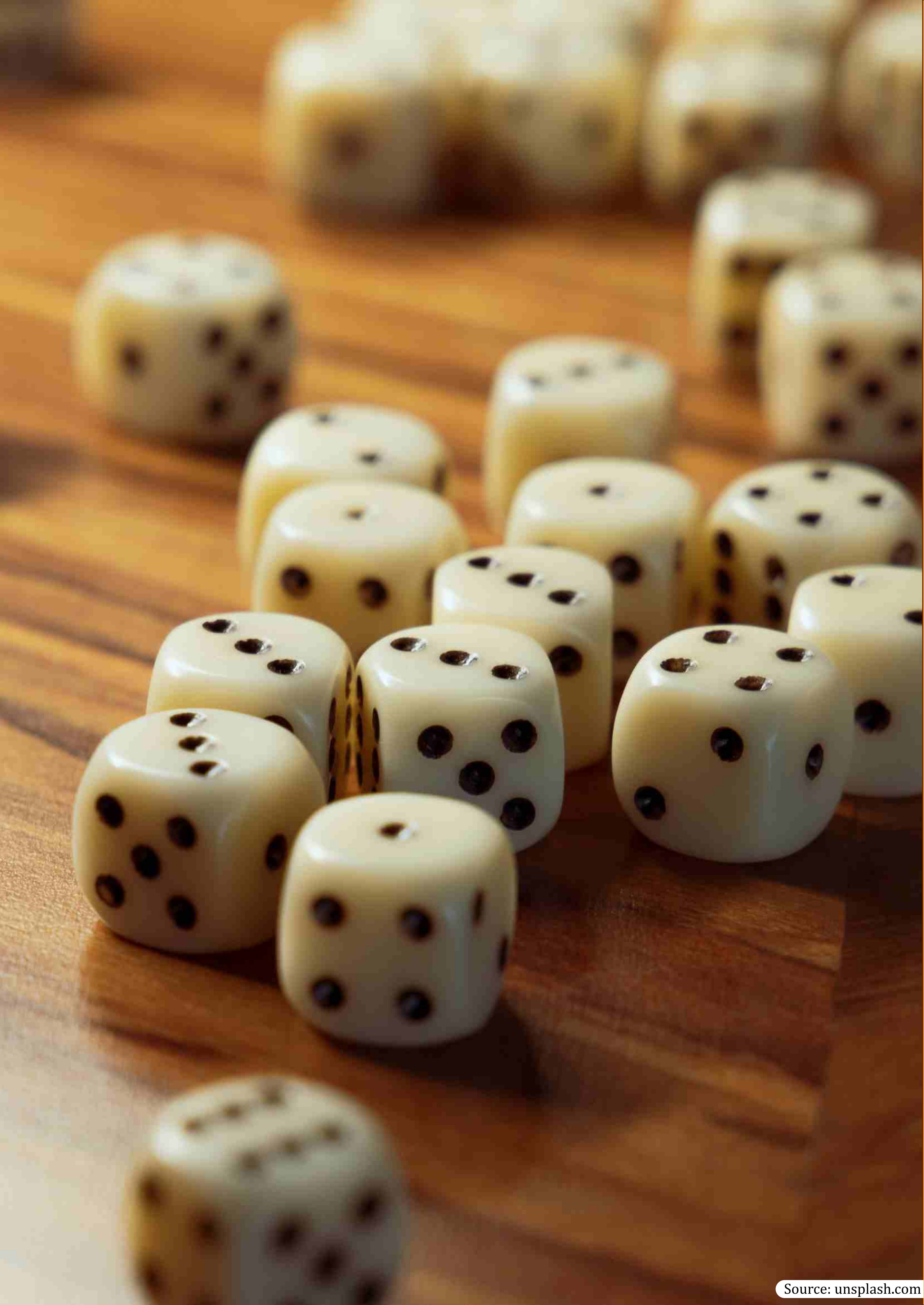
References

1. Dekking, F.M., Kraaikamp, C. Lopuhaa, H.P. Meester. L.E. A Modern Introduction to Probability and Statistics. Springer
2. Soong, T.T. Fundamentals of Probability and statistics for Engineers.(2004). John Wiley & Sons Inc.

Further Reading

https://www.pdfdrive.com/probability-and-statistics-textbook-e20201392.html

https://www.pdfdrive.com/mathematical-statistics-with-applications-7th-edition-e28265865.html


Source: unsplash.com

MODULE 3

Unit 1: Probability distribution for discrete random variables
Unit 2: Mathematical expectation for discrete random variables
Unit 3: Probability distribution for continuous random variables
Unit 4: Mathematical expectation for continuous random variables
Unit 5: Joint distributors of two discrete random variables
Unit 6: Joint distributors of two continuous random variables

UNIT 1 PROBABILITY DISTRIBUTION FOR DISCRETE RANDOM VARIABLES

Introduction

Most often, outcomes from a random experiment are summarized by a simple number. In some cases, however, descriptions of outcomes by a simple number may not be adequate. Because the particular outcome of the experiment is not known in advance, therefore it is useful to associate a number with each outcome in the sample space. The function that associates a number with the outcome of a random experiment is referred to as a random variable. So we will be looking at defining discrete random variables and describing cumulative distribution function of discrete random variables.

Learning Outcomes

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

1. Define the discrete random variables as well as the different formulations of probability functions
2. Describe the cumulative distribution function of discrete random variables
3. Compute probabilities from discrete random variables.

Main Content

Random variables and associated probability functions 5mins

Definition: Let S be a sample space associated with a random experiment E, a function that assigns a real number X(s) to each element s of the sample space S is called a random variable and it is denoted X. The testing of a number of electronic components is an example of random experiment. It is often important to allocate a numerical description to the outcome. For example, the sample space when three electronic components are tested may be written

$$S = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$$

Where N denotes “non-defective” and D denotes “defective”. One is naturally concerned with the number of defectives that occur. Thus each point in the sample space will be assigned a numerical value. Such a numerical value is a random variable.

Discrete random variables

X will be referred to as a discrete random variable if the range space R_X contains finite or countably infinite numbers of points

In the example of electronic components given above,

S = {NNN, NND, NDN, DNN, NDD, DND, DDN, DDD}

If we are interested in the number of defectives, then we can define a random variable X on S such that for

$s \in S$, X(s) is the number of defectives in the test as follows.

DDD: $X(DDD) \longrightarrow$ 3
DDN: $X(DDN) \longrightarrow$ 2
DND: $X(DND) \longrightarrow$ 2
NDD: $X(NDD) \longrightarrow$ 2
DNN: $X(DNN) \longrightarrow$ 1
NDN: $X(NDN) \longrightarrow$ 1
NND: $X(NND) \longrightarrow$ 1
NNN: $X(NNN) \longrightarrow$ 0

Clearly, X is a mapping of S into R_X, where R_X is a set of real numbers called the range space, and represented as

$R_X = \{0, 1, 2, 3\}$

Example

Consider an experiment involving the toss of a fair die twice (or tossing two dice together once). Let the sample space S consist of the 36 ordered pairs (i,j),

i = (1, 2, 3, 4, 5, 6); j = (1, 2, 3, 4, 5,6) where i is the outcome on the first die and j the outcome on the second die. Then

S = {(1,1), (1,2), (1,3),. . . ,(5,6), (6,6)}

Suppose we define the random variable X such that for $(i,j) \in S$

X(i,j) =i+j,

then we can map each element of the sample space to a set of real numbers as follows

(1,1): $X(1,1) \longrightarrow$ 2
(1,2),(2,1): $X(1,2), X(2,1) \longrightarrow$ 3
(1,3), (2,2), (3,1): $X(1,3), X(2,2), X(3,1) \longrightarrow$ 4
: : :
: : :
: : :
(5,6),(6,5): $X(5,6), X(6,5) \longrightarrow$ 11
(6,6): X(6,6) 12

Then $R_x = \{2,3,4\ldots,11,12\}$

Probability density (mass) function

The probability density (mass) function of a discrete random variable X (denoted f(x)) is a real valued function defined as f(x) = Pr(X = x). Thus if X is a discrete random variable which can take on values x1, x2,..., $\in R_x$, then f(x) will be called a probability density (mass) function if

(i) $p(x) = 0$

(ii) $\sum_{R_x} p(x) = 1$

(iii) $\Pr(X = x) = f(x)$

Activity 1

From example of electronic components given above

(a) Obtain the probability mass function of X
(b) Demonstrate that f(x) is indeed a probability mass function (pmf)
(c) Find (i) Pr(X<2) (ii) Pr(X>0)

Solution:

(a) The probability mass function is obtained as follows

X	0	1	2	3
f(x)	1/8	3/8	3/8	1/8

(b) From the table

$f(0)=\Pr(X=0)=1/8$
$f(1)=\Pr(X=1)=3/8$
$f(2)=\Pr(X=2)=3/8$
$f(3)=\Pr(X=3)=1/8$

It is observed that f(x) > 0 for all x and f(0) + f(1) + f(2) + f(3) = 1
Thus f(x) is indeed a pdf.
(c) Based on results in (a) above,

(i) $\Pr(X<2) = \Pr(X=0)+\Pr(X=1)$
$= 1/8+3/8 = 1/2$

(ii) $\Pr(X>0) = 1-\Pr(X=0)$
$= 1-1/8$
$= 7/8$

Activity 1

A shipment of 8 similar microcomputers to a retail outlet contains 3 that are defective. If school makes random purchase of 2 of these components, find the probability distribution for the number of defectives.

Solution:
Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can be any of the numbers 0, 1 and 2.

Now, $f(0) = \Pr(X = 0) = \dfrac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}} = \dfrac{10}{28}$

$$f(1) = \Pr(X = 1) = \frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}} = \frac{15}{28}$$

$$f(0) = \Pr(X = 0) = \frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}} = \frac{3}{28}$$

Thus the probability distribution of X is

x	0	1	2
f(x)	10/28	15/28	3/28

Activity 3

If f(x) = k(2x +1) , x = 1, 2,..., 6, is a probability density function.

(a) Determine the value of constant k?

(b) Find P(X >2).

Solution:
(a) Since f(x) is a pdf, then

$\sum_{R_x} f(x) = 1$, which implies that

$$\sum_{x=1}^{6} k(2x+1) = 1$$

k[(2×1+1)+ (2×2+1)+(2×3+1)+ (2×4+1)+ (2×5+1)+ (2×6+1)=1

48k =1

k=1/48

Thus f(x) = 1/48(2x +1) , x = 1, 2,..., 6

(b) Pr(X>2) = 1-Pr(X≤2)
= 1-1/48[3+5]
=1-8/48
= 5/6

Cumulative distribution function (cdf) of discrete random variables

Definition: The cumulative distribution function (cdf) for a discrete random variable X, denoted F(x) is defined as

$$F(x) = \Pr(X \le x) = \sum_{u \le x} f(u)$$

where f(u) is the discrete probability density function of a random variable X.

If X is a discrete random variable which takes values x1,x2,...xk then the cdf is given as

$F(x)=0$, $-\infty < x < x_1$
$= f(x_1)$, $x_1 \le x < x_2$
$= f(x_1)+f(x_2)$, $x_2 \le x < x_3$
:
:
$= f(x_1)+f(x_2)+\ldots+ f(x_k)$, $x_k \le x < \infty$

Example

Suppose that a fair coin is tossed twice so that the sample space is S = {HH, HT, TH, TT}. Let X represent the number of heads that comes up, then

X(HH) = 2 for HH
X(HT) = X(TH) = 1 for HT and TH
X(TT) = 0 for TT
R_x={0,1,2}
Then Pr(X=0)=1/4, Pr(X=1) = ½ , Pr(X=2) =1/4
The probability distribution function is thus given in the table below

X	0	1	2
Pr(X=x)	¼	½	¼

The cumulative distribution function is as follows

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

Activity 1

Let the probability mass function of a discrete random variable be given by

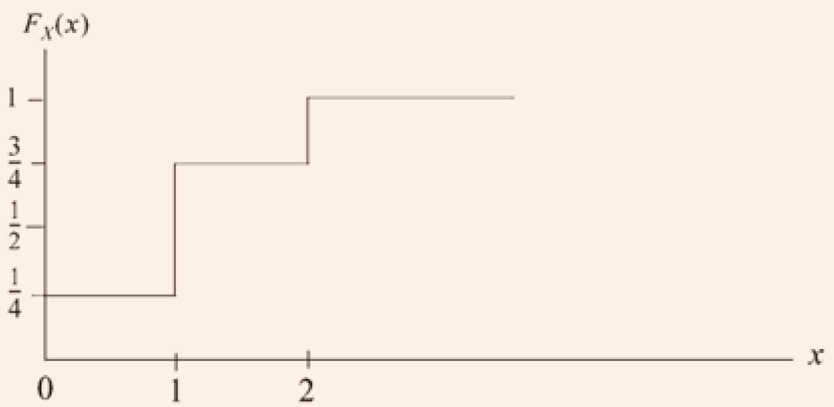
$$f(x)=\begin{cases}1/4 & x=0\\ 1/2 & x=1\\ 1/4 & x=2\end{cases}$$

(a) Obtain the cdf of X
(b) Plot the graph of the cdf in (a)

(a)

$$F_X(x)\begin{cases}0 & x<0\\ 1/4 & 0\le x<1\\ 3/4 & 1\le x<2\\ 1 & x\ge 2\end{cases}$$

(b)



We can obtain probability density function (pdf) from cumulative distribution function (cdf).

Example

From the cdf given below

$$F(x)=\begin{cases}0, & -\infty<x<3\\ 1/3, & 3\le x<4\\ 1/2, & 4\le x<5\\ 2/3, & 5\le x<6\\ 1, & 6\le x<\infty\end{cases}$$

The pdf is obtained as

$$f(x)=\begin{cases}1/3, & x=3\\ 1/2-1/3=1/6, & x=4\\ 2/3-1/2=1/6, & x=5\\ 1-2/3=1/3, & x=6\end{cases}$$

We can compute the following probabilities

(i) Pr(3<X≤5)

(ii) Pr(X≤5)

(i) Pr(3<X5) = f(4)+f(5)
= 1/6+1/6 =1/3

(ii) Pr(X5) =1-f(6)
= 1-1/3 =2/3

Summary

In this unit, Discrete random variables were defined and different formulations of probability functions on discrete random variables were given. The cumulative density function of discrete random variables were also defined and some probabilities based on discrete.

Self Assessment Questions

1. Define discrete random variable
2. Consider an experiment involving tossing a fair coins three times. If we define a random variable X as the number of heads that appears
   (a) Obtain the probability density function f(x) of X
   (b) Demonstrate that f(x) is indeed a probability density function (pdf)
   (c) Find (i) Pr(X<2) (ii) Pr(X>0)
3. Let the pdf of a random variable Y be given by

$$f(y)=ky^2, 1,2,3,4,5$$

Find
   (a) the value of k
   (b) Pr(2≤Y≤4)

4. Suppose that the pdf of variable X is given by f(x)=(2x + 1)/25, x=0, 1, 2, 3, 4, find Pr(X≤2) 5. Suppose that $f(x)=(3/4)(1/4)^x$, x=1,2,3,4,5; find Pr(X > 2)

Tutor Marked Assignment

1. The sample space of a random experiment is given as S= {a, b, c, d, e, f}, and each outcome is equally likely. Suppose that a random variable is defined as follows

Outcome:	a	b	c	d	e	f
x:	0	0	1.5	1.5	2	3

   (a) Determine the probability mass function of X
   (b) Find Pr(X<2)

2. If the pdf of a random variable is given as f(x) =Cx, for x=1,...,6.
   (a) Determine the value of C
   (b) Find Pr(X < 2)

3. Suppose that the probability mass function of the number of defectives R in a batch of items is

$$\Pr(R=r)=\begin{cases}\binom{4}{r}(0.2)^{r}(0.80)^{4-r}, & r=0,1,\ldots 4\\ 0 & \text{otherwise}\end{cases}$$

(a) Find the cdf of R

(b) What is the probability that fewer than two items will be defective?

References

1. Montgomery, D.C. & Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. *John Wiley & Sons, Inc.*
2. Roussas, G. G. A Course in Mathematical Statistics. Academic Press, New York
3. T.T. Soong. Fundamentals of Probability and statistics for Engineers. (2004). John Wiley & Sons Inc.

Further Reading

https://www.pdfdrive.com/probability-and-statistics-textbook-e20201392.html

https://www.pdfdrive.com/mathematical-statistics-with-applications-7th-edition-e28265865.html

UNIT 2 MATHEMATICAL EXPECTATION FOR DISCRETE RANDOM VARIABLES

Introduction

An extremely important concept in summarizing important characteristics of distributions of probability is mathematical expectation or simply called Expected value. So in this unit, we will be considering mathematical expectation for discrete random variables.

Learning Outcomes

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

1. Define mathematical expectation for discrete random variables
2. State and prove the properties of discrete random variables
3. Use the concept of mathematical expectation to determine the mean and variance of discrete random variables

Main Content

Principle of mathematical expectation

3mins

We shall introduce the unit by the following example. An engineer who needs a little extra money devises a game of chance in which some of his friends might wish to participate. The game he proposed was to let his friends cast an unbiased die and then receive a payment according to the following schedule.

A: If {1,2,3} occurs, he receives N1000, B if {4,5} occurs, he receives N5000, and if C = {6} occurs, he receives N35000. Since the die is unbiased, the probabilities of the respective events are 3/6/ 2/6 and 1/6. The problem facing the Engineer is the determination of the amount that will be charged for the opportunity of playing the game. He reasons that if the game is played a large number of times, about 3/6 of the trials will require a payment of N1000, about 2/6 will require a payment of 2/6 while the payment of N35000 will require about 1/6 of the trials. Thus, the approximate average payment is

$$1000(3/6)+5000(2/6)+35000(1/6) = 500+1666.667+ 5833.333$$
$$=8000$$

Thus he expects to pay N8000 on the average.

It should be noted that he never pays exactly N8000. The payment is either N1000, N5000 or N35000. However, the weighted average of 1000, 5000 and 35000 in which the weights are the respective probabilities 3/6, 2/6 and 1/6 equal N8000. Such a weighted average is called "mathematical expectation" of payment. Thus if the young man decides to charge N10000 per play, he would make N2000 per play"on the average". Since the most that a player would lose at the charge of N10000 per play is N9000, he might find that several players are attracted by the possible gain of N25000.
A more mathematical way of formulating the preceding example would be to let X be the random variable defined by the outcome of the cast of the die.

Thus, the pdf of X is

$$f(x)=\frac{1}{6},\ x=1,2,3,4,5,6$$

The payment can then be given in terms of o the observed x by the function

$$h(x)=\begin{cases}1000, & x=1,\ 2,\ 3\\ 5000, & x=4,5\\ 35000, & x=6\end{cases}$$

So that the mathematical expectation of the payment is given as

$$\sum_{x=1}^{6} xf(x)=(1000)\left(\frac{1}{6}\right)+(1000)\left(\frac{1}{6}\right)+(1000)\left(\frac{1}{6}\right)+(5000)\left(\frac{1}{6}\right)$$

$$+(5000)\left(\frac{1}{6}\right)+(35000)\left(\frac{1}{6}\right)$$

$$=(1000)\left(\frac{3}{6}\right)+(5000)\left(\frac{2}{6}\right)+(35000)\left(\frac{1}{6}\right)$$

$$=\frac{48000}{6}=8000$$

Definition: If f(x) is the pdf of a random variable X of the discrete type with the range space R_X, then the mathematical expectation or expected value of X is given by

$$E(X)=\sum_{Rx} xf(x)$$

The discrete formula given above says to take a weighted sum of the values x of X, where the weights are the probabilities f(x)

Properties of Mathematical expectation

When it exists, mathematical expectation satisfies the following properties.

(i) If a is a constant, then E(a)= a

(ii) If a is a constant and h(x) is a function then

$$E[ah(X)] = aE[h(X)]$$

(iii) If a_1 and a_2 are constants and h_1 and h_2 are functions, then $E[a_1(h_1(X)+a_2(h_2(X)] = a_1E[(h_1(X)+a_2E[(h_2(X)]$

Proofs

(i) We can express E(a) as

$$E(a)=\sum_{Rx} af(x)$$

$$= a\sum_{Rx} f(x)$$

$= a$, since $\sum_{Rx} f(x) = 1$ from the condition (ii) of a pdf

(ii) Knowing that $E[ah(X)]=\sum_{R_x} ah(x)f(x)$

$$=a\sum_{R_x} h(x)f(x)$$

$$=aE[h(X)]$$

(iii) $E[a_1(h_1(X)+a_2(h_2(X)]=\sum_{Rx}[a_1(h_1(x)f(x)+a_2(h_2(x)f(x)]$

$$=\sum_{Rx}[a_1(h_1(x)+a_2(h_2(x)]f(x)$$

$$=\sum_{Rx}[a_1h_1(x)f(x)]+\sum_{Rx}[a_2h_2(x)f(x)]$$

By applying proof (ii), it follows that

$$E[a_1(h_1(X)+a_2(h_2(X)] = a_1E[(h_1(X)+a_2E[(h_2(X)$$

Note: Property (iii) can be extended to more than two terms, hence mathematical expectation can be called a *linear or distributive operator.*

The mean and variance of discrete random variables 3mins

Mean of X: Consider a random variable X, we can think of the mathematical expectation defined above as a weighted mean of the random variable, given by

$$\mu = E(X) = \sum_{R_X} xf(x)$$

Definition: The expectation of any function of discrete X, denoted h(x) can be given as

$$E(h(X)) = \sum_{R_X} h(x)f(x)$$

Suppose $h(x) = x^2$, then

$$E(X^2) = \sum_{R_X} x^2f(x)$$

Variance of X

Suppose that the function of X, h(x) = $(x-\mu)^2$, then for the discrete X the variance is given as

$$\sigma_x^2 = E(X-\mu)^2 = \sum_{R_X}(x-\mu)^2 f(x)$$

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

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

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

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

Activity 1

Suppose that the probability density function of a random variable X is as given in the table below

X	-2	1	3
f(x)	1/3	1/6	½

Find (a) Mean and variance of X

(b) E(2x+5)

Solution: (a) Mean=E(X)= -2(1/3)+1(1/6)+3(1/2)

= -2/3+1/6+3/2

= 1

Variance = $E(X^2)$-$(E(X))^2$ =4(1/3)+1(1/6)+9(1/2)-1^2

= 6-1

= 5

(b) E(2x+5) = (2 (-2) + 5) × 1/3 + (2(1) + 5) × 1/6 + (2(3) + 5)×1/2

= 1/3 + 7/6 + 11/2 = 7

Activity 2

Suppose that a game is played with a single die which is assumed fair. In this game, a player wins N90 if a 2 turns up, loses N30 if a 6 turns up while he neither wins nor loses if any other face turns up. Find the expected amount won/lost on the toss and its variance.

Solution:

Let X be a random variable denoting the amount won/lost on a toss with probability f(x) of winning or losing. For the equally likely experiment, f(x) =1/6 for all i = 1, 2,..., 6.

We can then express the scenario in terms of the observed X as

$$X = \begin{cases} 90, & f(x) = 1/6 \\ -30, & f(x) = 1/6 \\ 0, & f(x) = 4/6 \end{cases}$$

Thus, the expected amount won/lost on a toss is given by

E(X) = 90×1/6+(-30×1/6)+0×4/6

= N10

This is the amount won

To obtain the variance, we need $E(X^2)$ given by

$E(X^2)$ = $(90^2×1/6)+(-30^2×1/6)]$ = 1200

Variance = - 1200 -10^2

= 1100

Activity 3

Let X have the pdf given as

$$f(x) = \begin{cases} \frac{x}{10}, & x = 1,2,3,4 \\ 0, & \text{otherwise} \end{cases}$$

Obtain the mean of X and its variance

We can obtain the mean and the variance as follows

$$\text{Mean} = E(X) = \sum_{x=1}^{4} x\left(\frac{x}{10}\right) = (1)\left(\frac{1}{10}\right) + (2)\left(\frac{2}{10}\right) + (3)\left(\frac{3}{10}\right) + (4)\left(\frac{4}{10}\right) = 3$$

$$E(X^2) = \sum_{x=1}^{4} x^2\left(\frac{x}{10}\right) = (1^2)\left(\frac{1}{10}\right) + (2^2)\left(\frac{2}{10}\right) + (3^2)\left(\frac{3}{10}\right) + (4^2)\left(\frac{4}{10}\right) = 10$$

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

$$= 10 - (3)^2 = 1$$

Activity 1

Let X be a random variable denoting the square of the score of number shown when a die is tossed once.

Find the mean and the standard deviation of X

Solution:

Let the possible values of the experiment be given as

S = {1, 4, 9, 16, 25, 36}

Each outcome has probability of 1/6

The mean of X, E(X) = 1/6(1+4+9+16+25+36)

= 91/6 = 15.1667

Computing E(X2) = 1/6($1^2+4^2+9^2+16^2+25^2+36^2$)

= 1/6(1+16+81+256+625+1296)

= 2275/6 =379.1667

Standard deviation of X, $\sigma_X = \sqrt{379.1667 - (15.1667)^2}$ = 12.2

Summary

We describe the expectation for discrete random variable and gave some real live illustrative examples were given. Some simple properties of mathematical expectation were given and proved. We applied the concept of mathematical expectation to calculate the means and variances of random variables of discrete form.

Self Assessment Questions

1. Define the mathematical expectation of a discrete random variable X
2. State and prove three properties of mathematical expectation of a discrete random variable
3. Suppose that X has the pdf given as

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

   Find the mean and the variance of X
4. Two unbiased dice are cast. A payment equals to the sum of the spots on the top sides is given the caster. Compute the expected value of the payment. 5. Suppose that the pdf of variable X is given by f(x) = (2x + 1)/25, x = 0, 1, 2, 3, 4, find the mean of X

Tutor Marked Assignment

1. The discrete random variable Y has probability distribution

$$f(y) = \begin{cases} \frac{y}{36}, & \text{y} = 1,2,3,\ldots,8 \\ 0, & \text{otherwise} \end{cases}$$

   Find E(Y), Var(Y)
2. Suppose that the pdf of a random variable X is given be

$$f(x) = \begin{cases} \frac{x^2}{55}, & \text{x} = 1,2,3,4,5 \\ 0, & \text{otherwise} \end{cases}$$

   Find E(X) and Var(X)
3. The distribution of number of messages X sent per hour by MTN with their respective probabilities, denoted (x,f(x)) are as follows: (10, 0.08), (11, 0.15) and (12, 0.30). Find the average number of messages sent per hour nearest to the whole hour.

References

1. Dekking, F.M. Kraaikamp, C. Lopuhaa, H.P. Meester. L.E.A. Modern Introduction to Probability and Statistics.
2. Montgomery, D.C. & Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. *John Wiley & Sons, Inc.*

Further Reading

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=14&ved=2ahUKEwj1qJf8tNLnAhXjnVwKHd_jCR84ChAWMAN6BAgDEAE&url=http%3A%2F%2Fwww.columbia.edu%2F~kr2248%2F4109%2Fchapter4.pdf&usg=AOvVaw3Z36G692z7Wg4UKJ5WpDJN

UNIT 3 PROBABILITY DISTRIBUTION FOR CONTINUOUS RANDOM VARIABLES

Introduction

You are welcome to another course unit. There are often situations when we have random variables whose set of possible values may not be countable. For example, let us consider a random variable whose values are the heights of all people over 18 years of age. Between any two values, say 163.5 and 164.5 centimeters. There are an infinite numbers of heights, one of which is 164 centimeters. The probability of selecting a person at random who is exactly 164 cm is zero.

Learning Outcomes

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

1. Define the continuous random variables as well as the different formulations of probability functions.
2. Describe the cumulative distribution function of continuous random variables
3. Compute probabilities from continuous random variables.

Main Content

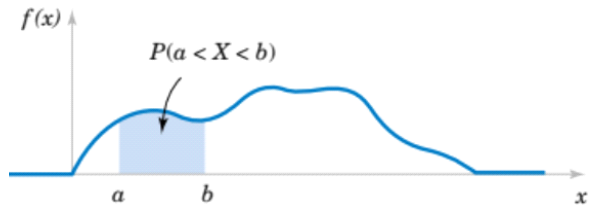
Probability density function of continuous random variables 3mins

SAQ 1

SAQ 2

SAQ 5

A probability density function f(x) can be used to describe the probability distribution of a continuous random variable X. Let us consider this example, consider the density of a system shown in the figure below. If an interval is likely to contain a value for X, then its probability corresponds to value for f(x). The probability that X is between a and b (the shaded region) is determined as the integral of f(x) from a to b. This integral is the area under the density function over this interval, and it can be loosely interpreted as the sum of all values of X over the interval.



Probability determined from the area under f(x)

Definition: Let us compute X to be a continuous random variable, then the probability density function (pdf) of X with range space R_x that is an interval or union of intervals, is a non-negative integrable function defined for all real values of $x \in (-\infty, \infty)$ and satisfying.

(i) f(x)≥0 for all

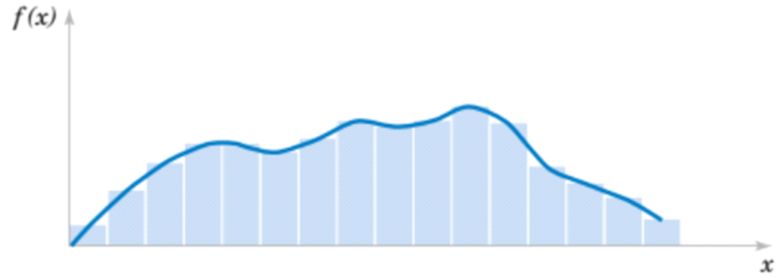
(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$ i.e the total area over the entire range space equals 1

(iii) For any interval (a, b), Pr(a ≤ X ≤ b) = $\int_a^b f(x)dx$

Note: When X is a continuous random variable, then the probability that X takes on any one particular value is generally zero. X is only defined for a continuous case when it lies between intervals of positive real values.

Thus for continuous random variable X,
Pr(a < X < b) = Pr(a < X ≤ b) = Pr(a ≤ X < b) = Pr(a ≤ X ≤ b) for all real x. A density function can be approximated by a histogram as seen in the figure below.

As observed, for each interval of the histogram, the area of the bar equals the relative frequency (proportion) of the measurements in the interval. This is an estimate of the probability that a measurement falls in the interval. Similarly, the area under f(x) over any interval equals the true probability that a measurement falls in the interval. One important point to note is that f(x) is used to calculate an area that represents the probability that X assumes a value in [a, b]. For example, suppose that X denotes baby's weight at birth, the probability that X results in [0.8kg, 2.0 kg] is the integral of the probability density function of X over this interval.



Histogram approximating a probability density

Activity 1

Suppose that X is a continuous random variable having the pdf given as

$$f(x) = \begin{cases} K(4x - 2x^2), & 0 < x < 2 \\ 0 & otherwise \end{cases}$$

Find (a) the value of K (b) Pr(X > 1)

Solution:
(a) Since f(x) is a pdf, we have

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

$$K\int_0^2 (4x - 2x^2)dx = 1$$

$$\frac{8K}{3} = 1 \text{ and } K = \frac{3}{8},$$

so that the pdf becomes

$$f(x) = \begin{cases} \frac{3}{8}(4x - 2x^2), & 0 < x < 2 \\ 0 & otherwise \end{cases}$$

(b) Pr(X>1) = $\frac{3}{8}\int_1^2 (4x - 2x^2)dx = \frac{1}{2}$

Activity 2

Given that a random variable Y has the pdf given as

$$f(y) = \begin{cases} e^{-y}, & y > 0 \\ 0 & otherwise \end{cases}$$

(a) Show that f(y) is indeed a pdf
(b) Compute Pr(1<Y<3)

Solution:
(a) Following the conditions for a pdf, we need to show here that f(y) ≥ 0 for all y, and $\int_{-\infty}^{\infty} f(y)dy = 1$

When y = 0, f(0) = $e^{-0} = 1$, when y = 8 , f(8) = $e^{-\infty} = 0$.

Thus f(y) = 0 for all y. also $\int_0^{\infty} e^{-y}dy = [-e^{-y}]_0^{\infty} = 1$

Therefore, f(y) is indeed a pdf.

(b) $\Pr(1 < Y < 3 = \int_1^3 e^{-y}dy = [-e^{-y}]_1^3 = e^{-1} - e^{-3}$ = 0.318

SAQ 3

SAQ 4

Distribution function of continuous random variables

The distribution function of a random variable X of the continuous type denoted F(x) is, defined in terms of the pdf of X is given by

$$F(x) = \Pr(X \le x) = \int_{-\infty}^{u} f(u)du$$

From the fundamental theorem of calculus, we have for x values for which the derivative given by

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

Since there are no steps or jumps in a distribution function of F(x) of the continuous type, it must be true that
Pr(X = b) = 0

For all values of b. this agrees with the fact that the integral $\int_b^b f(x)dx$ is taken to be zero, then it follows also that
Pr(a < X< b) = Pr(a < X ≤ b) = Pr(a ≤ X < b) = F(b) - F(a)

Activity 3

Suppose that Y is a continuous random variable with pdf
g(y) = 2y, 0 < y <1.
Find the distribution function F(y) of Y

Solution:

$$F(y)=\Pr(Y\le y)=\begin{cases}0, & y<0\\ \int_0^y 2u\,du=y^2, & 0\le y<1\\ 1, & y\ge 1\end{cases}$$

We can compute probabilities from a given distribution function
For example, use the last example above to compute the following probabilities

(a) $\Pr(\frac{1}{2}<Y\le\frac{3}{4})$ (b) $\Pr(\frac{1}{4}\le Y\le 1)$

Solution:

(a) $\Pr(\frac{1}{2}<Y\le\frac{3}{4})=F\left(\frac{3}{4}\right)-F\left(\frac{1}{2}\right)$

$$=\left(\frac{3}{4}\right)^2-\left(\frac{1}{2}\right)^2=\frac{5}{16}$$

(b) $\Pr(\frac{1}{4}\le Y\le 1)=F(1)-F\left(\frac{1}{4}\right)$

$$=1-\left(\frac{1}{4}\right)^2=\frac{15}{16}$$

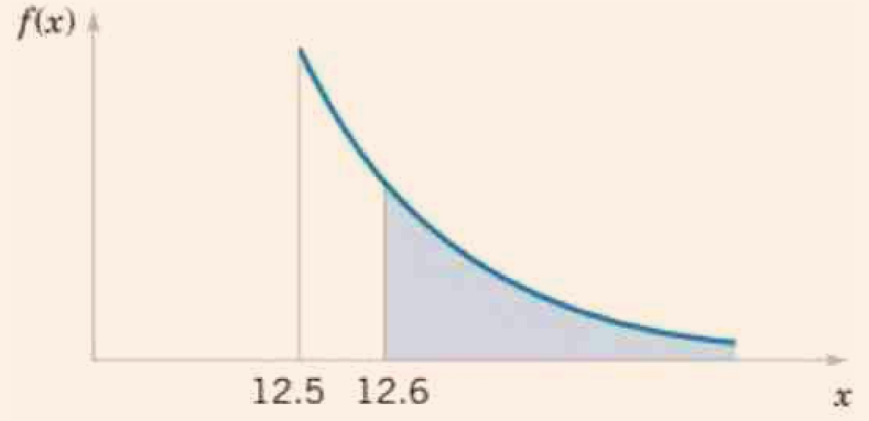
Activity 4

Suppose that a continuous random variable X denotes the diameter of a hole drilled in a sheet metal component measured in millimeters, which can be modeled by

$$f(x)=20e^{-20(x-12.5)}, x\ge 12.5$$

(a) If a part with a diameter larger than 12.6 millimeters is scrapped, what percentage of the parts is scrapped?
(b) What percentage of parts is between 12.5 and 12.6 millimeters?

Solution:
The density function and the requested probability are shown in the figure below



The pdf for Activity 4

(a) We are required to calculate Pr(X > 12.60). This is shown in the shaded region

$$\Pr(X>12.6)=\int_{12.6}^{\infty} f(x)dx$$

$$=\int_{12.6}^{\infty} 20e^{-20(x-12.5)}dx=-e^{-20(x-12.5)}\Big|_{12.6}^{\infty}$$

$$=-e^{-20(\infty)}+e^{-20(12.6-12.5)}$$

$$=0+e^{-2}=0.135$$

Thus the percentage of the parts scrapped is 13.5%

(b) $\Pr(12.5<X<12.6)=\int_{12.5}^{12.6} 20e^{-20(x-12.5)}dx=-e^{-20(x-12.5)}\Big|_{12.5}^{12.6}$

$$=-e^{-20(12.6-12.5)}+e^{-20(12.5-12.5)}$$

$$=e^{0}-e^{-2}=1-0.135=0.865$$

The percentage between 12.5 and 12.6 millimeters is 86.5%

Note: Because the total area under f(x) equals 1, we can alternatively calculate Pr(12.5 < X < 12.6) from 1 - Pr(X > 12.6) = 1-0.135 = 0.865

Summary

We have defined Continuous random variable. We have illustrated probability density function for continuous random variables was defined graphically. We gave numerical example. We also defined cumulative density function and some appropriate continuous probability distribution were selected to calculate specific probabilities.

Self Assessment Questions

1. State the conditions under which f(x) for a random variable X can be referred to as a continuous density function
2. The probability density function of the length of a metal rod is $f(x) = 2$ for $2.3 < x < 2.8$ meters. If the specifications for this process are from 2.25 to 2.75 meters, what proportion of the bars fail to meet the specifications?
3. Consider the function

$$f(x) = \begin{cases} c & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

   (a) For what value of c is f(x) a legitimate pdf
   (b) Find the cdf of the random variable X for the above cdf
4. The pdf of the time T (in minutes) it takes a bank teller to serve a customer is defined by

$$f(t) = \begin{cases} \frac{1}{6} & 2 \le t \le 8 \\ 0 & \text{otherwise} \end{cases}$$

   (a) What is the cdf of T?
   (b) What is the probability that a customer is served in less than 5 minutes?

Tutor Marked Assignment

1. Suppose that a random variable X has the pdf given as

$$f(x) = \begin{cases} \frac{x}{8}, & 3 < x < 5 \\ 0, & \text{otherwise} \end{cases}$$

Determine the following probabilities
   (a) $\Pr(X < 4)$
   (b) $\Pr(X < 3.5 \text{ or } X > 4.5)$
2. Suppose that Y is a continuous random variable having the pdf given as

$$f(y) = \begin{cases} A(2y - y^2), & 0 < y < 2 \\ 0 & \textit{otherwise} \end{cases}$$

   Find (a) the value of A (b) $\Pr(Y>1)$
3. Consider a random variable X for which the density function has the form ($c > 0$)

$$f(x) = \begin{cases} ce^{-cx}, & \text{for } x \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

   Find (a) $\Pr(0 < X \le 1)$ (b) $\Pr(X > 3)$

4. Suppose that the random variable X represents the waiting time of a customer at a ticket counter with the cumulative distribution given by

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x/3}, & 0 \le x < 3 \\ \frac{1 - e^{-x/3}}{2}, & x \ge 3 \end{cases}$$

Determine the probability that X is
   (a) more than two minutes and
   (b) between two and six minutes.

References

1. Montgomery, D.C. and Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. *John Wiley & Sons, Inc.*
2. Roussas. G.G. A Course in Mathematical Statistics. Academic Press, New York
3. Soong. T.T. Fundamentals of Probability and statistics for Engineers.(2004). *John Wiley & Sons Inc.*

Further Reading

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=2ahUKEwie4bya-trnAhVEz4UKHX1EBu8QFjACegQIBBAB&url=https%3A%2F%2Ffaculty.arts.ubc.ca%2Fhkasahara%2FEcon326%2Fnotes_exp.pdf&usg=AOvVaw3obFIPGy_ThQZ0MB1YdH0l

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwie4bya-trnAhVEz4UKHX1EBu8QFjAAegQIARAB&url=http%3A%2F%2Fwww.d.umn.edu%2F~zliu%2Fmath3611%2Fc04_mathexp.pdf&usg=AOvVaw0U0eeGdjh5Hxv_6pDo3Cfv

UNIT 4 MATHEMATICAL EXPECTATION FOR CONTINUOUS RANDOM VARIABLES

Introduction

I welcome you to the unit. We will define the mathematical expectations of a continuous random variable similar to a discrete random variable. I want you to note that the only difference is that the integration in the continuous case replaces the summation in the discrete random variables.

Learning Outcomes

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

1. Define expectation for continuous random variables
2. State and prove the properties of expectation of a continuous random variable
3. Compute and interpret expectation, variance, and standard deviation for continuous random variables.

Main Content

Expectation of random variables and functions of random variables 2mins

SAQ 1
SAQ 2
SAQ 3
SAQ 4
SAQ 5

For continuous-type random variables, the definitions associated with mathematical expectation are the same as those in the discrete case except that integrals replace summations. For illustrations, let us compute X to be a continuous random variable with a pdf f(x).

The expectation, expected value or mean of X is given as

$$\mu = E(X) = \int_{Rx} xf(x)dx$$

f(x)dx represents the probability that X is in an infinitesimal range of width dx around x. Thus we can interpret the formula for E(X) as a weighted integral of the values x of X, where the weights are the probabilities f(x)dx.

Definition: The expectation of any function of continuous random variable X, denoted h(x) can be given as

$$E(h(X)) = \int_{R_x} h(x)f(x)dx$$

Properties of Mathematical expectation for continuous random variables

The properties of mathematical expectation described for discrete random variables are also applicable in the continuous case with the replacement of the summation symbol $\sum$ with integral $\int$. Thus the mathematical expectation also satisfies the following properties.

(i) If a is a constant, then E(a)= a

(ii) If a is a constant and h(x) is a function then

$$E[ah(X)] = aE[h(X)]$$

(iii) If a_1 and a_2 are constants and h_1 and h_2 are functions, then

$$E[a_1(h_1(X)+a_2(h_2(X)] = a_1E[(h_1(X)+a_2E[(h_2(X)]$$

Proofs

(i) We can express E(a) as

$$E(a) = \int_{Rx} af(x)dx$$

$$= a\int_{Rx} f(x)dx$$

$$= a \text{ since } \int_{Rx} f(x)dx = 1$$

(ii) Similarly like the discrete case

Knowing that

$$E[ah(X)] = \int_{R_x} ah(x)f(x)dx$$

$$= a\int_{R_x} h(x)f(x)dx$$

$$= aE[h(X)]$$

(iii) $E[a_1(h_1(X)+a_2(h_2(X)] = \int_{Rx}[a_1h_1(x)+a_2h_2(x)]f(x)dx]$

$$= \int_{Rx}[a_1(h_1(x)f(x)dx] + \int_{Rx}[a_2(h_2(x)f(x)dx]$$

$$= a_1\int_{Rx}[(h_1(x)f(x)dx] + a_2\int_{Rx}[(h_2(x)f(x)dx]$$

By applying proof (ii), we have

$$E[a_1(h_1(X)+a_2(h_2(X)] = a_1E[(h_1(X) + a_2E[(h_2(X)$$

The variance is given by

$$\sigma^2 = E(X-\mu)^2 = \int_{Rx}(x-\mu)^2 f(x)dx$$

The standard deviation of X is

$$\sigma = \sqrt{E(X-\mu)^2}$$

Activity 1

Let the probability density function of a continuous random variable X be given as $f(x) = \frac{3x^2}{8}$ in the range [0,2]

(a) Find E(X)

(b) Can we say that X has a mean in the right half of this range?

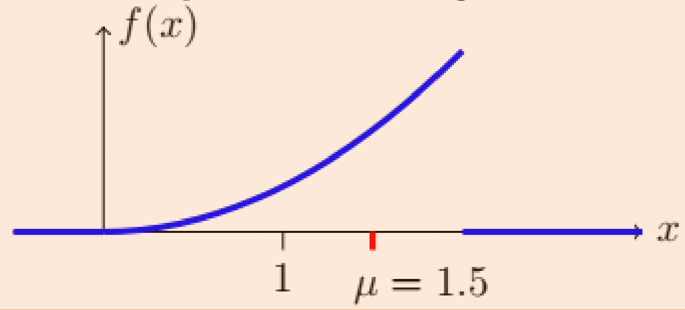
Solution:

(a) $E(X) = \int_{Rx} xf(x)dx$

$$= \int_0^2 x\frac{3x^2}{8}dx = \int_0^2 \frac{3x^3}{8}dx$$

$$= \left.\frac{3x^4}{32}\right|_0^2 = \frac{3}{2}$$

(b) Yes. Since the probability density increases as x increases over the range, the average value of x should be in the right half of the range and this is as represented in the diagram below



Activity 2

Let the pdf of a continuous random variables be given by

$$f(x) = \begin{cases} \frac{4}{81}(9x - x^3), 0 \le x \le 3 \\ 0, \quad otherwise \end{cases}$$

Find the mean and variance of X

$$E(X) = \int_0^3 xf(x)dx$$

$$= \frac{4}{81}\int_0^3 x(9x - x^3)dx$$

$$= \frac{4}{81}\int_0^3 (9x^2 - x^4)dx$$

$$= 1.6$$

$$E(X^2) = \frac{4}{81}\int_0^3 x^2(9x - x^3)dx$$

$$= 3$$
$$Var(X) = E(X^2)-(E(X))^2 = 3-(1.6)^2$$
$$= 0.44$$

Activity 3

Let the compute continuous random variable X to denote the current measured in a thin copper wire in milliamperes. Assume that the range of X is [0, 20 mA], and assume that the probability density function of X is f(x) = 0.05, find the mean current of the copper wire and the standard deviation.

Solution:
The mean of X is

$$E(X) = \int_0^{20} x f(x)dx$$

$$= \int_0^{20} 0.05\, x)dx = \left.\frac{0.05x^2}{2}\right|_0^{20}$$

$$= \frac{0.05[(20)^2 - 0}{2} = 10$$

$$E(X^2) = \int_0^{20} 0.05x^2)dx = \left.\frac{0.05x^3}{3}\right|_0^{20} = 133.3333$$

$$Var(X) = 133.3333 = 10^2 = 33.33$$

Summary

Expectations of random variables was defined for continuous random variables. A few properties were given and some numerical examples were illustrated.

Self Assessment Questions

1. Define the mathematical expectation of a continuous random variable X
2. State and prove the properties of expectation of a continuous random variable. Find E(X), Var(X)
3. If f(x)=e-2x , for x≥0, find the mean of ex
4. The waiting time X (in minutes) of a customer waiting to be served at a ticket counter has the density function

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

Determine the average waiting time and the variance

5. Suppose that the cdf of the lifetime of electrical equipment is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x/5} & 0 \leq x < \infty \end{cases}$$

Find the expected lifetime of the equipment?

Tutor Marked Assignment

1. If f(x) = a(2x +1)(3x-2) , 2 ≤ x ≤ 3 is a probability density function, determine the value of constant a. Hence or otherwise find E(X) and V(X)?
2. Consider a continuous random variable with pdf

$$f(x) = \begin{cases} \frac{1}{4} & 2 \leq x \leq 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the expected value of X and its variance

3. A shopping cart contains ten books whose weights are as follows: There are four books with a weight of 1.5 kg each, one book with a weight of 2 kg, two books with a weight of 2.5 kg each, and three books with a weight of 3.0 kg each.
   (a) What is the mean weight of the books?
   (b) What is the variance of the weights of the books?

References

1. Hogg , R.V. and Tanis, E.A. (1977). Probability and statistical Inference.
2. Montgomery, D.C. and Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. *John Wiley & Sons, Inc.*
3. Soong, T.T. Fundamentals of Probability and statistics for Engineers.(2004). *John Wiley & Sons Inc.*

Further Reading

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=14&ved=2ahUKEwj1qJf8tNLnAhXjnVwKHd_jCR84ChAWMAN6BAgDEAE&url=http%3A%2F%2Fwww.columbia.edu%2F~kr2248%2F4109%2Fchapter4.pdf&usg=AOvVaw3Z36G692z7Wg4UKJ5WpDJN

UNIT 5 JOINT DISTRIBUTIONS OF TWO DISCRETE RANDOM VARIABLES

Introduction

I welcome you to another unit. In the previous units, we have been considering a single random variable. Often time, in many random experiments, it is necessary that we consider two random variables simultaneously. I want you to know that the joint probability distribution of two random variables is often called a bivariate distribution.

Learning Outcomes

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

1. Define joint probability density function of two discrete random variables and the associated marginal and conditional probabilities.
2. Find a marginal and conditional probability densities functions from the joint probability density functions, say X and Y.
3. Show whether or not two random variables X and Y are independent

Main Content

Joint Distribution of a Discrete Random Variables

4mins

SAQ 1

SAQ 2

SAQ 3

We shall start this section with an illustrative example. Let us suppose we toss a pair of fair dice, in which our interest is to define the following:

X = the outcome on the first die = {1, 2, 3, or 4}

Y = the outcome on the second die = {1, 2, 3, or 4}

Suppose that we wish to find the probability that X takes on a particular value x, and Y takes on a particular value y. This can be written as $\Pr(X = x, Y = y)$

We can denote the support (range space of X as)

$R_y = \{1, 2, 3, 4\}$, and for Y as

$R_x = \{1, 2, 3, 4\}$

If we let (x, y) denote one of the possible outcomes of one toss of the pair of dice, then we observe that (1, 1) is a possible outcome, as is (1, 2), (1, 3) and (1, 4) are others. In similar manner, we can continue to enumerate all of the possible outcomes, we see that the joint support R has 16 possible outcomes given as

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

Since the dice are fair each of the 16 possible outcomes are equally likely with Pr(X=x,Y=y) =1/16

1st die	f(x,y)	2nd die (Y)				f(x)
		1	2	3	4	4/16
	1	1/16	1/16	1/16	1/16	4/16
	2	1/16	1/16	1/16	1/16	4/16
	3	1/16	1/16	1/16	1/16	4/16
	4	1/16	1/16	1/16	1/16	4/16
	f(y)	4/16	4/16	4/16	4/16	1

We can thus formally define a discrete joint probability density function (pdf) or probability mass function (pmf).

Definition: Let X and Y be two discrete random variables, and let R denote the two-dimensional support of X and Y. Then, the function f(x, y) = Pr(X = x , Y = y) is a **joint probability density function** (pdf.) if it satisfies the following three conditions:

(1) $0 \le f(x, y) \le 1$)

(2) $\sum_{(x,y)\in R}\sum f(x,y) = 1$

(3) $\Pr[(X,Y)\in A] = \sum_{(x,y)\in A}\sum f(x,y)$ where A is a subset of the support R.

Meanings of the above three conditions

(1) Condition 1: Each joint probability must be a valid probability number between 0 and 1 (inclusive).

(2) Condition 2: The sum of the probabilities over the entire support S must equal 1.

(3) Condition 3: This tells us that in order to determine the probability of an event A, one simply sums up the probabilities of the (x, y) values in A.

SAQ 1

Marginal probability density function

Suppose that X and Y have the joint probability density function f(x,y) with space R. The pdf of X alone is called the marginal pdf of X and it is defined by $f(x) = \Pr(X = x) = \sum_{y\in R_y} f(x,y), x \in R_x$, where, for each x in the support $R_x = \{1, 2, 3, 4\}$, , the summation is taken over all possible values of y

Similarly, the pdf of Y alone is called the marginal pdf of Y and it is defined by $f(y) = \Pr(Y = y) = \sum_{x\in R_x} f(x,y), y \in R_y$, where, for each y in the support $R_y = \{1, 2, 3, 4\}$, the summation is taken over all possible values of x.

SAQ 1

Independence of X and Y

Looking at the joint probabilities density function in the table, it is observed that

$\Pr(X = x, Y = y) = f(x,y) = 1/16$, and

$\Pr(X = x)\times\Pr(Y = y) = f(x)\times f(y) = (4/16)\times(4/16) = 1/16$

for all $x \in R_x, y \in R_y$

When this happens, then X and Y are said to be independent. We can then formally define independence of two random variables X and Y.

Definition: The random variables X and Y will be said to be independent if and only if: $\Pr(X = x, Y = y) = \Pr(X = x)\times\Pr(Y = y)$ $f(x,y) = f(x)\times f(y)$, for all $x \in R_x, y \in R_y$. Otherwise, they are said to be dependent.

Activity 1

Let the joint pdf of random variables X and Y be given as

$$f(x,y) = \frac{x+y}{21},\ x = 1,2,3;\ y = 1,2$$

(a) Find the

(i) marginal pdf f(x) of X

(ii) marginal pdf f(y) of Y

(b) Show whether or not X and Y are independent

Solution:

(a) (i) $f(x) = \sum_{y=1}^{2}\frac{x+y}{21}$

$$= \frac{x+1}{21} + \frac{x+2}{21}$$

$$= \frac{2x+3}{21},\ x = 1,2,3 \text{ ; and}$$

(ii) $f(y) = \sum_{x=1}^{3}\frac{x+y}{21}$

$$= \frac{x+y}{21} + \frac{x+y}{21} + \frac{x+y}{21}$$

$$= \frac{3y+6}{21},\ y = 1,2$$

(b) $f(x,y) = \dfrac{x+y}{21}$

$$f(x)f(y) = \left(\frac{2x+3}{21}\right)\times\left(\frac{2x+3}{21}\right)$$

$$= \frac{6xy + 12x + 9y + 18}{441}$$

$$= \frac{6x(y+2) + 9(y+2)}{441}$$

$$= \frac{(y+2)(6x+9)}{441} \neq \frac{x+y}{21}$$

Thus X and Y are not independent

Activity 2

Consider two random variables X and Y with joint pdf given as in the table

	Y = 0	Y = 1	Y = 2
X = 0	1/6	¼	1/8
X = 1	1/8	1/6	1/6

(a) Find Pr(X = 0, Y ≤ 1)
(b) Find the marginal pdfs of X and Y
(c) Find Pr(Y = 1|X = 0)

Solution:
(a) Pr(X = 0, Y ≤ 1) = p(0, 0) + p(0, 1) + p(0, 2) =1/6 + 1/4 = 5/12
(b) Since the range space of X is R_x = {0,1}, then to obtain the marginal pdf of X, we need to find
p_x (0) = p(0, 0)+p(0, 1) + p(0, 2) = 1/6 + 1/4 + 1/8 = 13/24,
and
p_x (1) = p(1, 0) + p(1, 1) + p(1, 2) = 1/8 + 1/6 + 1/6 = 11/24
We thus obtain

$$p_x(x) = \begin{cases} \frac{13}{24} & x = 0 \\ \frac{11}{24} & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Similarly, for the marginal pdf of Y, the range space of Y = R_y = {0, 1, 2} and we need to find
P_y (0) = p(0,0)+p(1,0)+p(0,2) = 1/6+1/8 = 7/24,
P_y (1) = p(0,1)+p(1,1) = 1/4+1/6 = 5/12
and
P_y (2) = p(0,1)+p(1,1) = 1/8+1/6 = 7/24

so that we obtain

$$p_y(y) = \begin{cases} \frac{7}{24} & y = 0 \\ \frac{5}{24} & y = 1 \\ \frac{7}{24} & y = 2 \\ 0 & \text{otherwise} \end{cases}$$

(c) Find Pr(Y = 1|X = 0) = $\frac{\Pr(X=0, Y=1)}{\Pr(X=0)} = \frac{\frac{1}{4}}{\frac{13}{24}} = \frac{6}{13}$

Conditional Distribution

Definition: The conditional probability density function of X given Y = y is defined by

$$f(x \mid y) = \frac{f(x,y)}{f(y)}, \quad f(y) > 0 \text{ ; and}$$

the conditional probability density function of Y given X= x is defined by

$$f(y \mid x) = \frac{f(x,y)}{f(x)}, \quad f(x) > 0$$

From activity 1 above,

$$f(x \mid y) = \frac{x+y}{21} \Big/ \frac{3y+6}{21}$$

$$f(x \mid y) = \frac{x+y}{3y+6}, \quad x = 1,2,3; \ y = 1,2$$

Also $$f(y \mid x) = \frac{x+y}{21} \Big/ \frac{2x+3}{21}$$

$$f(x \mid y) = \frac{x+y}{2x+3}, \quad x = 1,2,3; \ y = 1,2$$

Summary

Let us have a recap on what we have done, we defined the joint probability density (mass) function of two discrete random variables. We solved numerical problems here. We also discussed independence of the random variables and demonstrated using numerical examples.

Self Assessment Questions

1. Suppose that the random variables W and Z with range spaces R_W and R_Z be jointly distributed with pdf given as f(w,z). Define the
   (a) marginal probability density function of W
   (b) marginal probability density function of Z
2. Let the joint pdf of random variables X and Y be given as

$$f(x,y) = \frac{xy^2}{30}, \ x = 1,2,3; \ y = 1,2$$

Find the
   (a) Find the marginal probability density function of X
   (b) Find the marginal probability density function of Y
   (c) Are X and Y independent?
3. Suppose that X1 and X2 have the joint pdf

$$f(x_1, x_2) = \frac{x_1 + 2x_2}{12}, \ x_1 = 1,2,; \ x_2 = 1,2$$

Find the conditional density function of X_1 given X_2 denoted $f(x_1|x_2)$
4. Let U and W have the joint pdf

$$f(u,w) = \frac{u+w}{32}, \ u = 1,2; \ w = 1,2,3,4$$

Find $\Pr(1 \le W \le 3 | U = 1)$

Tutor Marked Assignment

1. Let X and Y have the joint p.d.f given as

$$f(x,y) = \begin{cases} A(x+y) & ,x = 1,2 \quad y = 1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

Find
   (a) the value of A
   (b) the marginal distributions of X and Y.
   (c) Pr(Y>X)
   (d) Pr (X+Y<4)
2. Let the joint pmf of random variables W and Z be given as follows

p(w,z)	W			
Z	0	1	2	3
-1	1/8	0	0	0
1	0	1/8	2/8	1/8
2	0	1/8	1/8	0
3	0	1/8	0	0

Compute the following probabilities
   (a) p(0, -1) (b) Pr(W = 1) (c) Pr(Z = 2)
3. Suppose that random variables X and Y have joint pdf given by
   $f(x,y) = (2x+y)$ x = 1, 2; y = 1, 2
where k is a constant
   (a) What is the value of k?
   (b) Find the marginal probability mass functions of X and Y
   (c) Are X and Y independent?

References

(1) Douglas C. Montgomery, G.C. Runger. Applied Statistics and Probability for Engineers. John Wiley & Sons. Inc.
(2) Dekking, F.M. Kraaikamp, C. Lopuhaa, H.P. and Meester. L.E. A Modern Introduction to Probability and Statistics (Understanding Why and How) Springer.

Further Reading

https://online.stat.psu.edu/stat414/node/104/

UNIT 6 JOINT DISTRIBUTIONS OF TWO CONTINUOUS RANDOM VARIABLES

Introduction

In the previous section, we investigated joint probability distributions of discrete random variables, that is, random variables whose support contains a countable number of outcomes. In the discrete case, the number of outcomes in the support can be either finite or countably infinite. In this section, as the title suggests, we are going to investigate probability distributions of continuous random variables, that is, random variables whose support R contains an infinite interval of possible outcomes.

Learning Outcomes

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

1. Define joint probability density function of two continuous random variables and the associated marginal and conditional probabilities.
2. Find a marginal and conditional probability density function of the random variables from the joint probability density function of say X and Y.
3. Show whether or not two joint random variables X and Y are independent

Main Content

Joint Distribution of a Continuous Random Variables

Definition: Let X and Y be two continuous random variables, and let S denote the two-dimensional support of X and Y. Then, the function f(x, y) is a **joint probability density function** (pdf) if it satisfies the following three conditions:

$f(x,y) > 0$

$\int_{Ry}\int_{Rx} f(x,y)dxdy = 1$

$\Pr[(X,Y) \in A] = \int\int_A f(x,y)dxdy$ where $\{(X,Y)\}$ is an event in

the *xy*-plane.

Meanings of the above three conditions

Condition 1: The function f(x,y) must be a nonnegative two-dimensional surface floating above the xy-plane.

Condition 2: The integral of the probabilities over the entire support R must equal 1.

Condition 3: This tells us that in order to determine the probability of an event A, one simply integrate the probabilities of the (x, y) over the space defined by the event A.

Definition: Marginal density function of continuous random variable

Let $f(x,y)$, for $x \in R_X$, $y \in R_Y$, be the joint probability density function of random variables X and Y. The marginal probability density function of X is given as

$$f(x) = \int_{R_Y} f(x,y)\,dy, \text{ for } x \in R_X, \text{ and}$$

the marginal probability density function of Y is given as

$$f(y) = \int_{R_X} f(x,y)\,dy, \quad y \in R_Y$$

For example, if X and Y have joint probability density function given as

$$f(x,y) = \begin{cases} 2, \; 0 \le x \le y;\; x \le y \le 1 \\ 0, \text{ otherwise} \end{cases}$$

then the marginal pdf of X is

$$f(x) = \int_x^1 2\,dy = [2y]_x^1$$
$$= 2(1-x),\; 0 \le x \le 1$$

and the marginal pdf of Y is

$$f(y) = \int_0^y 2\,dx = [2y]_0^y$$
$$= 2y,\; 0 \le y \le 1$$

Conditional distributions for continuous density functions

SAQ 1

Let the joint density function of random variables X and Y be $f(x,y)$, for $x \in R_X$, $y \in R_Y$ such that f(y) > 0, then the conditional probability density function of X given Y is defined as

$$f(x \mid y) = \frac{f(x,y)}{f(y)}, \text{ for } x \in R_X$$

and the conditional probability density function of Y given X is

$$f(y \mid x) = \frac{f(x,y)}{f(x)}, \text{ for } y \in R_Y \text{ if and only if } f(x) > 0$$

From our last example,

$$f(x \mid y) = \frac{2}{2y} = \frac{1}{y}, \quad 0 \le x \le y$$

and $f(y \mid x) = \dfrac{2}{2(1-x)} = \dfrac{1}{1-x}, \quad 0 \le y \le 1$

Independence of Random Variables X and Y

Similar to the case of discrete random variables, continuous random variables X and Y that are jointly distributed will be said to be independent if

$f(x,y) = f(x) \times f(y)$ for all $x \in R_X, y \in R_Y$ Otherwise, they will be said to be dependent.

From the results above, we obtain

$$f(x) \times f(y) = 2(1-x) \times 2y$$
$$= 4y(1-x) \neq 2$$

Thus X and Y are not independent

Computing probabilities

From the last example, find $\Pr(\frac{1}{8} \le X \le \frac{1}{4} \mid y = \frac{3}{4}$

This can be obtained as

$$\Pr\left(\frac{1}{8} \le X \le \frac{1}{4} \middle| y = \frac{3}{4}\right) = \int_{1/8}^{1/4} f(x \mid y = \frac{3}{4})\,dx$$

but $f(x \mid y) = \dfrac{1}{y}$

Thus substituting this, we have

$$\Pr\left(\frac{1}{8} \le X \le \frac{1}{4} \middle| y = \frac{3}{4}\right) = \int_{1/8}^{1/4} \frac{1}{3/4}\,dx$$
$$= \int_{1/8}^{1/4} \frac{4}{3}\,dx = \frac{4}{3}[x]_{1/8}^{1/4}$$
$$= \frac{4}{3}\left[\frac{1}{4} - \frac{1}{8}\right] = \frac{1}{6}$$

Mean and Variance

The mean is given as

$$E(X) = \int_{R_X} x f(x)\,dx$$
$$= \int_0^1 2x(1-x)\,dx$$

$$=\int_0^1 (2x-2x^2)dx$$

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

$$=1-\frac{2}{3}-0=\frac{1}{3}$$

$$E(Y)=\int_{R_y} yf(y)dy$$

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

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

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

$$=\frac{2}{3}-0=\frac{2}{3}$$

To obtain the Variance, one needs to obtain the $E(X^2)$, $E(Y^2)$ and then obtain the variances for X and that of Y respectively as $E(X^2)-(E(X))^2$ and $E(Y^2)-(E(Y))^2$,

$$E(X^2)=\int_0^1 2x^2(1-x)dx$$

$$=\int_0^1 (2x^2-2x^3)dx$$

$$=\left[\frac{2}{3}x^3-\frac{1}{2}x^4\right]_0^1$$

$$=\frac{2}{3}-\frac{1}{2}-0=\frac{1}{3}$$

$$\text{Var}(X)=\frac{1}{3}-\left(\frac{1}{3}\right)^2=\frac{2}{9}$$

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

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

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

$$=\frac{1}{2}$$

$$\text{Var}(Y)=\frac{1}{2}-\left(\frac{2}{3}\right)^2$$

$$=\frac{1}{2}-\left(\frac{4}{9}\right)=\frac{5}{18}$$

Conditional mean

The conditional mean of X given Y = y can be obtained as

$$E(X\mid Y=y)=\int_0^y xf(x\mid y)dx$$

$$\int_0^y \frac{x}{y}dx=\left(\frac{x^2}{2y}\right)_0^y=\frac{y}{2},\ 0\le y\le 1$$

Similarly, the conditional mean of Y given X = x is

$$E(Y\mid X=x)=\int_x^1 yf(y\mid x)dy=\int_x^1 \frac{y}{(1-x)}dy$$

$$=\left(\frac{y^2}{2(1-x)}\right)_x^1=\left(\frac{1-x^2}{2(1-x)}\right)$$

$$=\left(\frac{(1-x)(1+x)}{2(1-x)}\right)=\frac{1+x}{2},\ 0\le x\le 1$$

From the above we can compute the conditional mean for a specified x value. For example,

$$E(Y\mid X=4)=\frac{1+4}{2}=2.5$$

Activity 1

Let us suppose that the joint p.d.f of random variables V and W is given as

$$f(v,w)=\begin{cases}k(2v+3w) & ,0\le v\le 1\ ,0\le w\le 1\\ 0 & \text{otherwise}\end{cases}$$

Find

(a) the value of k
(b) g (v) and g(w), the marginal density functions of V and W respectively
(c) Are V and W independent?

Solution:

$$f(v,w)=\begin{cases}k(2v+3w), & 0\le v\le 1,\ 0\le w\le 1\\ 0 & o/w.\end{cases}$$

For a pdf, it is required that

(a) $$k\int_0^1\int_0^1 (2v+3w)\,dv\,dw = 1$$

$$k\int_0^1 \left[2vw + \frac{3w^2}{2}\right]_0^1 dv = 1$$

$$k\int_0^1 \left[2v+\frac{3}{2}\right] dv = 1$$

$$k\left[v^2 + \frac{3v}{2}\right]_0^1 = 1$$

$$k\left[1+\frac{3}{2}\right] = 1$$

$$5k = 2$$

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

So that

$$f(v,w) = \begin{cases} \frac{2}{5}(2v+3w), & 0\le v\le 1,\ 0\le w\le 1 \\ 0 & otherwise \end{cases}$$

(b) The marginal density function of V is

$$g(v) = \frac{2}{5}\int_0^1 (2v+3w)\,dw$$

$$= \frac{2}{5}\left[2vw + \frac{3w^2}{2}\right]_0^1$$

$$= \frac{2}{5}\left[2v+\frac{3}{2}\right]$$

$$= \frac{1}{5}(4v+3), \qquad 0\le v\le 1$$

and the marginal density function of W is

$$g(w) = \frac{2}{5}\int_0^1 (2v+3w)\,dv$$

$$= \frac{2}{5}\left[v^2+3vw\right]$$

$$= \frac{2(1+3w)}{5}, \quad 0\le w\le 1$$

(c) V and W will be said to be independent if and only if $f(v,w) = g(v)g(w)$.

Now $g(v)g(w) = \dfrac{(4v+3)}{5}\dfrac{(2+6w)}{5} = \dfrac{2(4v+12vw+9w+3)}{25}$

and f(v,w) = $\frac{2}{5}(2v+3w)$

Above is an indication that V and W are not independent.

(d) The conditional density function of W given that V = x is

$$g(w|v) = \frac{f(v,w)}{g(v)}$$

$$= \frac{\frac{2}{5}(2v+3w)}{\frac{1}{5}(4v+3)}$$

$$= \frac{2(2v+3w)}{4v+3}$$

(e) $$\Pr\left(\frac{3}{10}\le w\le\frac{1}{2}\middle|v=\frac{1}{5}\right) = \int_{3/10}^{1/2} g\left(w\middle|v=\frac{1}{5}\right)$$

$$= 2\int_{3/10}^{1/2}\left(\frac{2/5+3w}{4/5+3}\right)dw.$$

$$= \frac{2}{3.8}\left[0.4w+\frac{3}{2}w^2\right]_{3/10}^{1/2} \qquad = 0.168$$

Summary

We defined the joint probability density function of two continuous random variables. We also defined marginal density. We solved the numerical problems here. We discussed the independence of the random variables and demonstrated using numerical examples. We also defined conditional distributions and calculated numerical probabilities on this. We determined Mean and Variance determined using the marginal densities.

Self Assessment Questions

1. State the conditions under which the function f(x,y) of two \ continuous random variables X and Y can be referred to as a joint probability density function
2. Suppose that the joint pdf of E and F is specified as follows

$$f(e,f)=\begin{cases} ce^2f, & e^2 \le f \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine the value of c
(b) Find Pr(E ≥ F)
(c) Find the marginal densities of X
(d) Find the marginal of Y

3 Suppose that X and Y have joint probability density function
$f(x,y) = 4xy$, for $0 \le x \le 1$ and $0 \le y \le 1$
Obtain f(x) and f(y)

4. Let $f(x,y) = 2e^{-x-y}$, for $0 \le x \le y < \infty$ be the joint pdf of X and Y
(a) Find the marginal densities of X and Y
(b) Show whether or not X and Y are independent
(c) Determine the conditional probability density function of Y given X

Tutor Marked Assignment

1. Let us suppose that X and Y are jointly continuous with pdf

$$f(x,y)=\begin{cases} cx^2+\frac{xy}{3} & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine the value of the constant c for which $f(x,y)$ is a valid joint pdf
(b) Find the marginal pdfs of X and of Y
(c) Are X and Y independent?

2. Let X and Y be two jointly continuous random variables with pdf

$$f(x,y)=\begin{cases} x+ay^2 & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the constant *a*

(b) Find $\Pr(0 \le X \le \frac{1}{2}, 0 \le Y \le \frac{1}{2})$

3. Suppose we have X and Y that have supports 0 < x < 1 and 0 < y <1,

with joint pdf $f(x,y)=\frac{3}{2}(x^2+y^2)$

Find the marginal pdf of X and use it to compute $\Pr(X < \frac{1}{2})$

References

(1) Douglas C. Montgomery, George C. Runger. Applied Statistics and Probability for Engineers. John Wiley & Sons. Inc.
(2) F.M. Dekking, C. Kraaikamp, H.P. Lopuhaa, L.E. Meester. A Modern Introduction to Probability and Statistics (Understanding Why and How) Springer.
(3) Morris H. Degroot. Probability and Statistics. Addison-Wesley publishing company

Further Reading

https://online.stat.psu.edu/stat414/node/107/

Source: womenshealthmag.com

MODULE 4

Unit 1: Higher moments of random variables
Unit 2: Binomial distribution
Unit 3: Poisson distribution
Unit 4: Negative binomial and geometric distributions
Unit 5: Uniform distribution
Unit 6: Normal distribution

UNIT 1 HIGHER MOMENTS OF RANDOM VARIABLES

Introduction

I welcome you to another unit. I want you to know first that moments are used to describe the peculiarities of the distributions of the data. Using moments, we can measure the central tendency, dispersion, symmetry and peakedness of the data. In this unit, we will consider the first four moments of a random variable.

You are expected be able to determine the moments of discrete and continuous random variables. They should also be able to determine the coefficients of skewness andkurtosis from the results of moments.

Learning Outcomes

1. Describe the raw and central moments of a random variable and their relationships.
2. Determine the mean, variance, coefficients of skewness and kurtosis using the central moments
3. Define and interpret Chebychev's inequality
4. Calculate probability from Chebychevs inequality.

Main Content

Raw Moments of random variables 4mins

SAQ 1

The rth raw moment for a random variable X is the moment about the origin (0), and it is defined by

$$m_r = E(X^r) = \begin{cases} \sum\limits_{R_X} x^r f(x), \text{if X is discrete} \\ \int\limits_{R_X} x^r f(x)dx, \text{ if X is continuous} \end{cases}$$

For example, the first raw moment (when r =1) is the mean $E(X^1) = m_1 = \mu$, the second raw moment is $E(X^2) = m_2$, the third raw moment is $E(X^3) = m_3$, etc.

Definition: (Central Moment): The rth central moment about the mean μ of a random variable X is given as

$$\mu_r = E(X-\mu)^r = \begin{cases} \sum_{R_X} (x-\mu)^r f(x), & X \text{ is discrete} \\ \int_{R_X} (x-\mu)^r f(x)dx, & X \text{ is continuous} \end{cases}$$

For example, the first central moment about the mean $E(X-\mu)^1 = 0$, the second central moment is the variance

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

Relationship between raw and central moments

The central moment about the mean of a random variable X can be expressed in terms of the raw moments. For example, use $\mu = m_1$, the second central moment is

$$\begin{aligned} \mu_2 &= E(X-m_1)^2 = E(X^2 - 2(m_1)X + (m_1)^2) \\ &= E(X^2) - 2m_1 E(X) + (m_1)^2 \\ &= m_2 - 2(m_1)^2 + (m_1)^2 \\ &= m_2 - (m_1)^2 = \text{variance} \end{aligned}$$

The higher central moments can also be expressed in terms of the raw moments. Thus, extending the derivations in the second moment above, the third central moment can be expressed as

$$\mu_3 = E(X - m_1)^3 = m_3 - 3m_2m_1 + 2(m_1)^3,$$

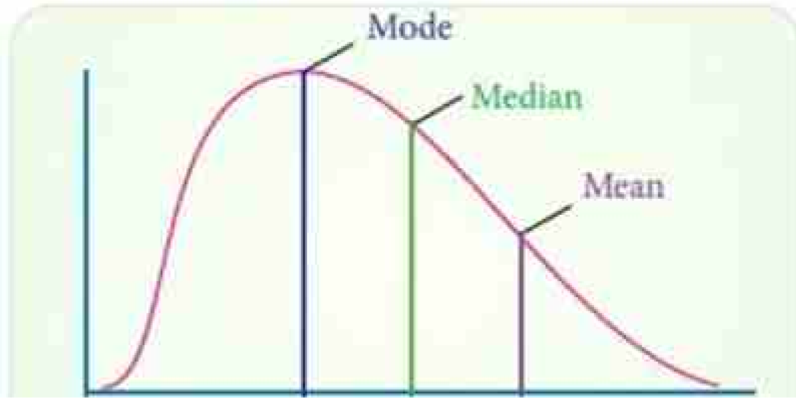
and the fourth central moment can be expressed as

$$\mu_4 = m_4 - 4m_3m_1 + 6m_2(m_1)^2 - 3(m_1)^4$$

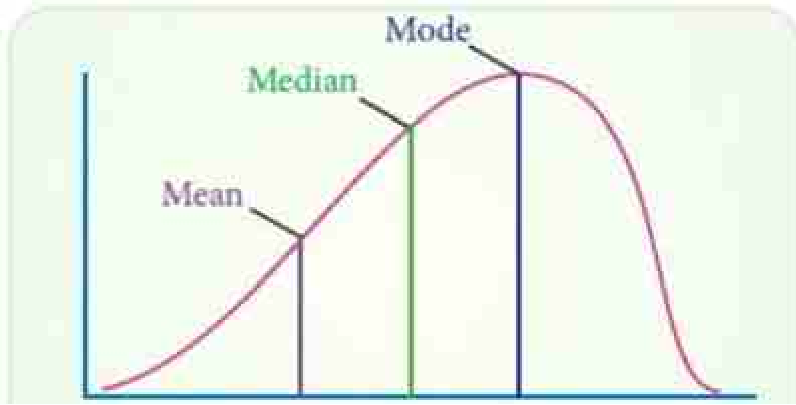
Skewness and Kurtosis

SAQ 2

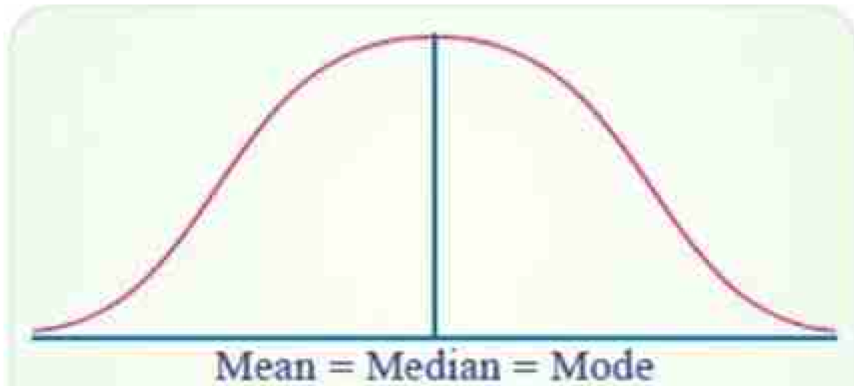
Skewness is the degree of asymmetry, or departure from symmetry, of a distribution. If the frequency curve of a distribution has a longer tail to the right of the central maximum than to the left as in (a), the distribution is said to be *skewed to the right*, or to have *positive skewness* as in (a) where the distribution has mode <median<mean. If the longer tail is to the left, it is said to be *skewed to the left*, or to have *negative skewness* as in (b), and mean < median < mode. However, if mean = median = mode as in (c), then the distribution is said to be symmetric or normal.



(a) **Right skewed (Positive skewness)**

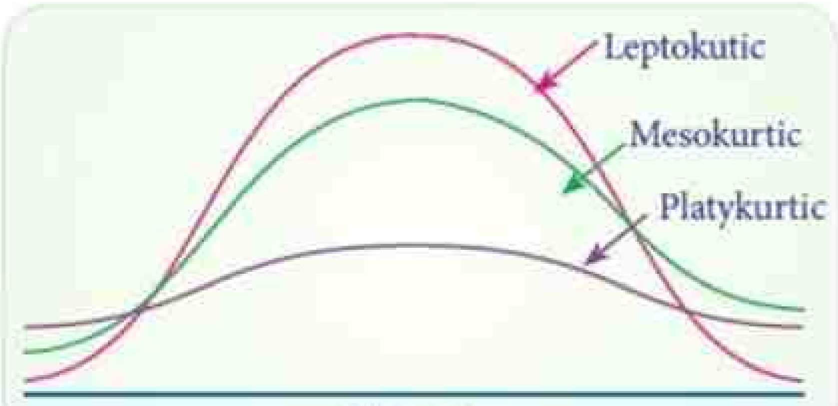


(b) **Left skewed (Negative skewness)**



(c) **Symmetric (Normal)**

Kurtosis in statistics is the degree of flatness or peakedness in the region about the mode of a frequency curve. The degree of kurtosis of distribution is measured relative to the peakedness of normal curve. In other words, measures of kurtosis tell us the extent of which a distribution is more peaked or flat-topped than the normal curve.



If a curve is more peaked than the normal curve, it is called '*leptokurtic*'. In such a case items are more closely bunched around the mode. On the other hand if a curve is more flat-topped than the normal curve, it is called '*platykurtic*'. The bell shaped (normal) curve itself is known as '*mesokurtic*'. We can find how much the frequency curve is flatter than the normal curve using measure of kurtosis.
The coefficient of skewness, denoted *sk* is mathematically given by

$$sk = \frac{\mu_3}{(\mu_2)^{3/2}}$$

Interpretation of Coefficient of Skewness and Kurtosis

Coefficient of skewness of zero (sk = 0) implies symmetric (normal) distribution, sk> 0 implies positive skewness while sk < 0 implies negative skewness.

Kurtosis measures the relative peakedness or flatness of a distribution relative to a normal distribution.
The coefficient of kurtosis, denoted *kurt* is given by

$$kurt = \frac{\mu_4}{(\mu_2)^2}$$

A distribution with positive kurtosis (*kurt*>0) is termed leptokurtic, with negative kurtosis (*kurt*<0) is termed platykurtic and with *kurt* = 0 is mesokurtic or normal distribution.

Activity 1

In testing three electronic components, the following is the probability distribution of X, the number of defectives in the test

p(0) =1/8
p(1) =3/8
p(2) =3/8
p(3) =1/8

Compute the first four raw and central moments

Solution:
The first four raw moments

$$m_1 = E(X) = (0\times1/8)+(1\times3/8)+(2\times3/8)+(3\times1/8)$$
$$= 3/8+6/8+3/8 = 12/8$$
$$m_2 = E(X^2) = (0\times1/8)+(1\times3/8)+(2^2\times3/8)+(3^2\times1/8)$$
$$= 3/8+12/8+9/8 = 24/8$$
$$m_3 = E(X^3) = (0\times1/8)+(1^3\times3/8)+(2^3\times3/8)+(3^3\times1/8)$$
$$= 3/8+24/8+27/8 = 54/8$$
$$m_4 = E(X^4) = (0\times1/8)+(1^4\times3/8)+(2^4\times3/8)+(3^4\times1/8)$$
$$= 3/8+48/8+81/8 = 132/8$$

The first four central moments

$$\mu_1 = \frac{12}{8} - \frac{12}{8} = 0$$
$$\mu_2 = \frac{24}{8} - \left(\frac{12}{8}\right)^2 = 0.75$$
$$\mu_3 = \frac{54}{8} - 3\left(\frac{24}{8}\right)\left(\frac{12}{8}\right) + 2\left(\frac{12}{8}\right)^3 = 0$$
$$\mu_4 = \frac{132}{8} - 4\left(\frac{54}{8}\right)\left(\frac{12}{8}\right) + 6\left(\frac{24}{8}\right)\left(\frac{12}{8}\right)^2 - 3\left(\frac{12}{8}\right)^4 = 1.312$$

Activity 2

From the pdf below,

$$f(x) = \begin{cases} 4(x-x^3), & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$$

(a) Find the first four raw and central moments
(b) Use the results to determine the coefficients of skewness and kurtosis and comment on the results

Solution:
(a) We first compute the raw moments as follows

$$m_1 = E(X) = \int_0^1 4x(x-x^3)dx$$
$$= 4\int_0^1 (x^2 - x^4)dx$$

$$=4\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_0^1$$

$$=4\left[\frac{1}{3}-\frac{1}{5}\right]_0^1 = \frac{8}{15}$$

$$= 0.5333$$

$$m_2 = E(X^2) \quad 4\int_0^1 x^2(x-x^3)dx$$

$$= 4\int_0^1 (x^3-x^5)dx$$

$$=4\left[\frac{x^4}{4}-\frac{x^6}{6}\right]_0^1$$

$$=4\left[\frac{1}{4}-\frac{1}{6}\right]= \frac{1}{3}$$

$$= 0.3333$$

$$m_3 = E(X^3) = 4\int_0^1 x^3(x-x^3)dx$$

$$= 4\int_0^3 (x^4-x^6)dx$$

$$=4\left[(\frac{x^5}{5}-\frac{x^7}{7})\right]_0^1$$

$$= 4\left[(\frac{1}{5}-\frac{1}{7})\right]= \frac{8}{35}$$

$$= 0.2286$$

$$m_4 = E(X^4) = 4\int_0^1 x^4(x-x^3)dx$$

$$= 4\int_0^1 (x^5-x^7)dx$$

$$=4\left[(\frac{x^6}{6}-\frac{x^8}{8})\right]_0^1$$

$$= 4\left[(\frac{1}{6}-\frac{1}{8})\right]_0^1= \frac{1}{6}$$

$$= 0.1667$$

The corresponding central moments given in terms of the raw moments as expressed above are obtained as below

$$\mu_2 = \frac{1}{3}-\left(\frac{8}{15}\right)^2 = 0.0489$$

$$\mu_3 = \frac{8}{35}-3\left(\frac{1}{3}\right)\left(\frac{8}{15}\right)+2\left(\frac{8}{15}\right)^3 = -0.0014$$

$$\mu_4 = \frac{1}{6}-4\left(\frac{8}{35}\right)\left(\frac{8}{15}\right)+6\left(\frac{1}{3}\right)\left(\frac{8}{15}\right)^2-3\left(\frac{8}{15}\right)^4 = 0.0052$$

$$\text{Coefficient of skewness} = sk = \frac{-0.0014}{(0.0489)^{3/2}} = -0.129$$

$$\text{Coefficient of kurtosis} = kurt = \frac{0.0052}{(0.0489)^2} = 2.175$$

The result of skewness above (< 0) suggests that the distribution has longer tail to the left (negative skewness) and kurtosis result (>0) suggests that the shape of the distribution is leptokurtic, which implies that the curve is more peaked than the normal curve.

Chebychev's Inequality

Definition: Consider a random variable X with finite mean μ and finite variance σ^2. Then for every k $\geq$1,

$$\Pr(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2} \quad \text{and} \quad \Pr(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2}$$

Interpretation: Chebychev's inequality states that the probability that X differs from its mean by at least k standard deviation is at most $\frac{1}{k^2}$ and that the probability that X differs from its mean by less than k standard deviation is at least 1 - $\frac{1}{k^2}$

Activity 3

A random variable X has mean 40 and variance 4, find

(i) $\Pr(|X-40| \geq 5)$

(ii) $\Pr(|X-40| < 5)$

Solution:

From the above definition, let μ=40 and σ^2=4, then

$\Pr(|X-40| \geq 5) = \Pr(|X-40| \geq 2k)$

(I) From the two sides of the equation, 2k = 5, so that k=5/2

Thus,

$$\Pr(|X-40|\geq 5)\leq \frac{1}{(5/2)^2}=\frac{4}{25}$$

(ii) $\Pr(|X-40|<5)\geq 1-\frac{4}{25}=\frac{21}{25}$

Summary

After a careful study of this unit, let us recap some of the things we have learnt. We presented definitions of raw and central moments both discrete and continuous random variables. The central moments were expressed in terms of raw moments and these were illustrated with numerical examples. The results from the central moments were used to compute coefficients of skewness and kurtosis. We also described Chebychev's inequality with illustrative examples.

Self Assessment Questions

1. Define the rth raw and central moments of a random variable X
2. Let X be a discrete random variable having support $R_x=\{2,1\}$ and probability mass function

$$p(x)=\begin{cases}3/4 & \text{if } x=1\\ 1/4 & \text{if } x=2\\ 0 & \text{otherwise}\end{cases}$$

   (a) Find the first two raw moments
   (b) Use the results in (a) to determine the variance of X
3. Suppose that the first four raw moments of a random variable X are
   $m_1 = 1.6$ $m_2 = 3.0$ $m_3 = 16.2$ $m_4 = 13.5$
   (a) Obtain the first four central moments μ_1, μ_2, μ_3, μ_4,
   $\mu_1 = 0$, $\mu_2 = 0.440$, $\mu_3 = -0.008$, $\mu_4 = -17.048$
   (b) Use the results in 3(a) to find the coefficients of skewness and kurtosis
4. State and interpret Chebychev's inequality
5. Suppose that the time (in minutes) required to serve a customer at the counter of a bank has an exponential distribution for which the mean is 4 minutes. Using Chebychev's inequality, determine the simple upper bound on the probability that the time required to serve a customer will be at least 25 minutes from the mean.

Tutor Marked Assignment

1. Let us consider the continuous random variable X which follows an exponential distribution with the probability density function given as

$$f(x)=\theta e^{-\theta x}$$

   (a) Find the mgf of X
   (b) Find the mean and variance of X
2. Let the probability mass function of a discrete random variable having support $R_X = \{1, 2, 3\}$ be given as

$$p(x)=\begin{cases}1/2 & \text{if } x=1\\ 1/3 & \text{if } x=2\\ 1/6 & \text{if } x=3\\ 0 & \text{otherwise}\end{cases}$$

   (a) Find the first four raw moments of X and the corresponding central moments
   (b) I want you to use the central moments in (a) to determine the coefficients of skewness and kurtosis

References

1. R.E. Walpole; R.H. Myers; S.L. Myers & Keying Ye . Probability & Statistics for Engineers & Scientists. ISBN 81-7808-613-1.
2. R.V. Hogg & E.A. Tanis. Probability and Statistical Inference.

Further Reading

Daley, D. J. and Vere-Jones, D. An introduction to the theory of point processes. Vol. {II}. Probability and its Applications (New York). Springer, New York.

UNIT 2 BINOMIAL DISTRIBUTION

Introduction

You are welcomed to another unit course. We will be discussing the nature and applications of discrete probability distribution known as Binomial distribution. An understanding of the situations in which the distribution arises enable us to choose it for a scientific phenomenon under consideration.

Learning Outcomes

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

1. Define and state binomial random variable and state its assumptions;
2. Derive the mean and variance of binomial distribution; and
3. Compute probabilities based on binomial distribution

Main Content

Bernoulli Distribution

4mins

Before we start the discussion of Binomial distribution, we shall briefly discuss Bernoulli distribution. This is a distribution in which the sample space of its random experiment has only two outcomes. For example, tossing a fair coin has two outcomes head (H) or tail (T). An item selected from a production process can be defective (D) or non-defective (ND). A patient admitted in the hospital can either be dead (D) or alive (A). It is generally termed as a success-failure experiment with probability of success p and the probability of failure 1-p

Definition: Let X be a discrete random variable, which can only assume two values, 1 or 0, then the probability density function of X can be given as

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

where p is the probability that X takes on value 1(probability of success) and 1-p is the probability of value 0 (probability of failure)

Mean and Variance of Bernoulli Distribution

The mean of Bernoulli Distribution is

$$E(X) = \sum_{x=0}^{1} x p^{x}(1-p)^{1-x}$$

Put 0 and 1 in the place of x in succession we have

$$= 0.p^{0}(1-p)^{1-0} + 1.p^{1}(1-p)^{1-1}$$

by simplifying , we have

$$= 0 + p = p$$

Variance of X is

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

Also put 0 and 1 in the place of x in succession we have

$$= \sum_{x=0}^{1} x^{2} p^{x}(1-p)^{1-x} - p^{2}$$

$$= 0.p^{0}(1-p)^{1-0} + 1.p^{1}(1-p)^{1-1} - p^{2}$$

By simplifying

$$= p - p^{2} = p(1-p)$$

Thus the mean and variance of the above Bernoulli distribution are p and p(1-p) respectively.

Binomial distribution

I want you to suppose that the Bernoulli experiment above is repeated n independent times with constant probability of occurrence of event (termed probability of success) p in each trial, and we let X denote the number of success in the n trials, then the resultant distribution is binomial.

Let us illustrate Binomial distribution with the following examples

1. Flip a coin 10 times. Let X be the number of heads obtained.
2. A machine produces 1% defective parts. Let X be the number of defective parts in the next 25 parts produced.
3. Each sample of air has a 10% chance of containing a particular rare molecule. Let X be the number of air samples that contain the rare molecule in the next 18 samples analyzed.
4. Of all bits transmitted through a digital transmission channel, 10% are received in error. Let X be the number of bits in error in the next five bits transmitted.

5. A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X be the number of questions answered correctly.
6. In the next 20 births at a hospital, let X the number of female births.
7. Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let X the number of patients who experience improvement.

In the above illustrative examples, each of the random experiments can be thought of as consisting of a series of repeated, random trials: 10 flips of the coin in experiment 1, the production of 25 parts in experiment 2, and so forth. The random variable in each case is a count of the number of trials that meet a specified criterion. The outcome from each trial either meets the criterion that X counts or it does not; consequently, each trial can be summarized as resulting in either a success or a failure. For example, in the multiple choice experiment, for each question, only the choice that is correct is considered a success. Choosing any one of the three incorrect choices results in the trial being summarized as a failure. The terms success and failure are just labels whose meanings are different from the grammatical interpretations.

The probability density function of Binomial Distribution

SAQ 1

SAQ 2

SAQ 3

Suppose now that n independent trials, each of which results in a “success” with probability p and in a “failure” with probability 1 – p, are to be performed. If X represents the number of successes that occur in the n trials, then X is said to be a binomial random variable with parameters (n, p), with the probability density function given as

$$f(x) = \Pr(X = x) = \binom{n}{x} p^{x}(1-p)^{n-x}, 0,1,\ldots,n$$

where $\binom{n}{x} = \dfrac{n!}{x!(n-x)!}$ is the number of different groups of I objects that can be chosen from a set of n objects.

For instance, if n = 5, x = 2, in a success (s)- failure(f) experiment, then there are $\binom{5}{2} = \dfrac{5!}{2!(5-2)!} = \dfrac{5\times4\times3!}{2\times3!} = 10$ choices of the two trials that are to result in successes, and these are any of the outcomes (s, s, f , f , f); (f , s, s, f , f); (f , f , s, f , s); (s, f , s, f , f); (f , s, f , s, f); (s, f , f , s, f); (f , s, f , f , s); (f , f , f , s, s); (s, f , f , f , s); (f , f , s, s, f)

Assumptions of Binomial Distribution

A random experiment consists of n Bernoulli trials such that:
(1) the trials are independent
(2) each trial results in only two possible outcomes, labeled as "success" and "failure"
(3) the probability of a success in each trial, denoted as p, remains constant

Activity 1

The probability that an insect survives a certain dose of insecticide is 0.75. If 10 insects are sprayed with the insecticide, what is the probability that
(i) 3 will survive
(ii) all will die
(iii) at least 1 will survive

Solution:
If X denotes the number of insects that survives, where p = 0.75, 1-p = 0.25 and n= 10, then

(i) $\Pr(X=3) = \binom{10}{3}(0.75)^3(0.25)^{10-3}$

$= 120(0.0000610)(0.421875) = 0.00308$

(ii) All will die implies none (0) will survive, thus the required probability is

$\Pr(X=0) = \binom{10}{0}(0.75)^0(0.25)^{10}$

$= (0.25)^{10}$

$= 9.536743\text{e-}07$

(iii) $\Pr(X=1) = 1-\Pr(X<1)$

$= 1-9.536743\text{e-}07$

$= 1.0000$

Mean and variance of binomial distribution

The mean and variance of a binomial distribution can be derived using direct method, method of moment or moment generating function. We shall use the direct method and students are advised to use the other two methods.

Consider $X^2 = (X(X-1) + X)$

and

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

where

$$\mu = E(X) = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)} p^x (1-p)^{n-x}$$

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

$$= np \sum_{x=0}^{n-1} \binom{n-1}{x-1} P^{x-1} (1-P)^{n-x}$$

$$= np, \quad \text{since} \quad \sum_{x=0}^{n-1} \binom{n-1}{x-1} P^{x-1} (1-P)^{n-x} = 1$$

Also. $E(X(X-1) = \sum_{x=0}^{n} x(x-1)\binom{n}{x} p^x (1-p)^{n-x}$

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

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

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

Thus

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

and the variance

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

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

$= np(1-p)$

For example to find the mean and the variance in the Activity

n= 10, p = 0.75, 1-p = 0.25

Mean is $10 \times 0.75 = 7.5$ and the variance is $10 \times 0.75 \times 0.25 = 1.875$

Activity 2

An occupant of a new bungalow has just installed 20 ceiling fans. Suppose that each has a probability 0.2 of functioning more than three months.
(a) What is the probability that at least five of these fans function more than three months?
(b) What is the average number of ceiling fans the occupant has to replace in three months?

Solution:
We assume that the light bulbs perform independently and that X is the number of bulbs functioning more than three months with probability p = 0.2 and 1-p =0.8, n= 20.

(a) The probability that at least five of these fans function more than three months is given by

$$\Pr(X\geq5) = 1-\Pr(X\leq4)$$

$$= 1-\sum_{x=0}^{4}\binom{20}{x}(0.2)^x(0.8)^{20-x}$$

$$= 1-(0.012+0.058+0.137+0.205+0.218)$$

$$= 0.370$$

(b) The average number of ceiling fans that functions more than three months is obtained as

$$E(X) = np$$

$$= 20(0.2) = 4$$

Therefore, the average number to be replaced in three months

$$20 - 4 = 16$$

Summary

Definition of binomial random variable was given and the assumption were stated. The mean and variance were derived. Numerical examples of all above were given.

Self Assessment Questions

1. State the assumptions of a Binomial distribution
2. Derive the mean and variance of a Binomial distribution
3. The probability that a certain component will survive electric shock is 3/4. If 5 components are selected at random, what is the probability that
   (a) at most 2 components will survive electric shock?
   (b) at least 3 components will survive electric shock?
4. Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant.
   (a) Find the probability that in the next 18 samples, exactly 2 contain the pollutant.
   (b) Determine the probability that at least four samples contain the pollutant.
5. The probability that a lab specimen contains high level of contaminants is 0.1. If five samples are checked independently, what is the probability that exactly
   (a) one sample will contain high level of contaminants.
   (b) no sample will contain high level of contaminants

Tutor Marked Assignment

1. Suppose that a random variable X has a binomial distribution with $n = 10$, $p = 0.5$, Determine the following probabilities.
   (a) $\Pr(X = 5)$
   (b) $\Pr(X \leq 2)$
   (c) $\Pr(X \geq 9)$
2. A person tested for COVID-19 could either be positive or negative. If 5 persons are selected at random for the test, what is the probability that
   (a) Between 2 and 4 will test positive
   (b) All will test negative
3. A Sales representative is selling a newly introduced product from house to house in a suburban town. The probability that he sells a product at any house he visits is 0.4. If he visits eight houses in a day,
   (a) State the probability mass function
   (b) Find the probability that he will sell a product to exactly five of these houses?
   (c) What is the expected number of product he will sell ?
4. A survey regarding youth welfare found that one out of 5 youths received their spending money from part-time jobs. If ten youths are randomly selected, find the probability that
   (a) exactly 3
   (b) at least 2 will have part-time jobs.

References

(1) Sheldon M. Ross. Introduction to Probability and Statistics for Engineers and Scientists, Elsevier academic Press.
(2) Montgomery, D. C. Runger. G. C. Applied Statistics and Probability for Engineers. John Wiley & Sons
(3) Harry Frank & Steven C. Althoen. Statistics Concepts and Applications

Further Reading

https://www.studocu.com/en-ca/document/the-university-of-british-columbia/fundamentals-of-biostatistics/lecture-notes/lecture-notes-lecture-3-binomial-and-poisson-distributions/218374/view

UNIT 3 POISSON DISTRIBUTION

Introduction

In this unit, we shall be discussing another discrete probability distribution used to model the number of events occurring in a given time period where such events occur at random within certain interval. Let X denote the random variable that counts the number of events in the interval, and suppose that the average number of events in the interval is λ. Let the interval be partitioned into n subintervals of small length, say, 1 unit each. If the subinterval chosen is small enough, then the probability that more than one event occurs in the subinterval is negligible. Furthermore, we can interpret the assumption that events occur at random to imply that every subinterval has the same probability of containing an event, say, p. Finally, if we assume that the probability that a subinterval contains an interval is independent of other subintervals, we can model the distribution of X as a Poisson random variable.

A few examples of Poisson random variable are number of:

(1) particles of contamination in semiconductor manufacturing,
(2) flaws in rolls of textiles,
(3) calls to a telephone exchange,
(4) power outages, and
(5) plane crashes

Learning Outcomes

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

1. Define Poisson random variable and the associated probability density function
2. Computing numerical probabilities based on Poisson distribution
3. Derive the mean and variance of Poisson distribution
4. Describe the Poisson approximation to binomial

Main Content

Poisson random variable

3mins

SAQ 1

SAQ 2

SAQ 3

Poisson random variable can be generalized to include a broad array of random experiments. The reasoning of partitioning interval such that the average number of events in the interval is λ can thus be applied to any interval, including an interval of time, an area, or a volume. Some examples of experiments that can be modelled by Poisson distribution include:

(i) counts of particles of contamination in semiconductor manufacturing,
(ii) flaws in rolls of textiles,
(iii) calls to a telephone exchange,
(iv) power outages, and
(v) atomic particles emitted from a specimen.

Definition: Let X be a discrete random variable with a non-negative integer support. X will be said to have a Poisson distribution with parameter λ>0 if the pdf is given as

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

where λ is the mean number of occurrence per unit.

Activity 1

During a laboratory experiment, the average number of radioactive particles passing through a point per millisecond is 4. What is the probability that

(a) exactly 6 particles will pass through the point in a given millisecond?
(b) at least 7 particles will pass through the point in a given millisecond?
(c) between 2 and 6 particles will pass through the point in a given millisecond?

Solution:
let X be the random variable denoting the number of particles that pass through the point per millisecond.
Given λ = 4, then the probability that x particles will pass through the point is given as

$$\Pr(X = x) = \frac{e^{-4}4^x}{x!} \quad x = 0,1,\ldots$$

Thus

$$\text{(a)}\ \Pr(X=6) = \frac{e^{-4}4^6}{6!} = \frac{e^{-4}4^6}{6\times5\times4\times3\times2\times1}$$

$$= \frac{0.0183\times4096}{720}$$

$$= 0.1041$$

$$\text{(b)}\ \Pr(X\geq7) = \sum_{x=7}^{\infty}\frac{e^{-4}4^x}{6!}$$

$$= 1 - e^{-4}\sum_{x=0}^{6}\frac{4^x}{6!}$$

$$= 1 - e^{-4}\left[\frac{4^0}{0!}+\frac{4^1}{1!}+\frac{4^2}{2!}+\frac{4^3}{3!}+\frac{4^4}{4!}+\frac{4^5}{5!}+\frac{4^6}{6!}\right]$$

= 1-0.0183(1+4+8+10.6667+10.6667+8.5333+5.68889)
= 1- 0.8886 = 0.1114

Units conversion in Poisson Distribution

It is important to use consistent units in the calculation of probabilities, means, and variances involving Poisson random variables. The following example illustrates unit conversions.
For example, if the average number of errors per page of manuscript is 3.4, then the
(1) average number of errors in 10 pages of wire is 3.4×10 =34
(2) average number of errors in 50 pages is 3.4×50 =170
(3) average number of errors in 100 pages is 3.4×100 =340

Thus, if a Poisson random variable represents the number of counts in some interval, the mean of the random variable must equal the expected number of counts in the same length of interval.

Activity 2

Suppose that flaws occur at random along the length of a thin copper wire. Let X denote the random variable that counts the number of flaws in a length of L millimeters of wire and suppose that the
average number of flaws per millimeters is λ =2.3. Find the probability of:
(a) exactly 2 flaws in 1 millimeter of wire.
(b) exactly 10 flaws in 5 millimeters of wire

(c) at least 1 flaw in 2 millimeters of wire

Solution:
(a) Let X denote the number of flaws in 1 millimeter of wire. Then, E(X) = 2.3
The probability of exactly 2 flaws in 1 millimeter of wire is

$$\Pr(X=2)=\frac{e^{-2.3}2.3^2}{2!}$$

$$=\frac{0.1003\times 5.29}{2}=0.2651$$

(b) Let X denote the number of flaws in 5 millimeters of wire. Then, E(X) = 2.3×5 = 11.5
Therefore, the probability of exactly 10 flaws in 5 millimeters of wire is

$$\Pr(X=10)=\frac{e^{-11.5}11.5^{10}}{10!}=0.113$$

(c) Let X denote the number of flaws in 2 millimeters of wire. Then, E(X) = 2.3×2 = 4.6
Therefore, the probability of at least 1 flaw in 2 millimeters of wire is

$$\Pr(X\ge 1)=1-\Pr(X=0)=1-e^{-4.6}$$

$$=1-0.0101=0.9899$$

Mean and variance of Poisson Random Variable

I want you note that one main characteristic of Poisson distribution is that the mean and the variance are equal and these are 2.3, 11. For example, in activity 2, the means = variances are 2.3, 11.5 and 4. 6 for (a), (b) and (c) respectively.

To derive the Mean and Variance of Poisson Distribution

$$\text{For Poisson}(\lambda),\ E(X)=\sum_{x=0}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x!}=\sum_{x=1}^{\infty} xe^{-\lambda}\frac{\lambda^x}{x(x-1)!}$$

$$=\lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!}=\lambda e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^y}{y!},\text{ where } y=x\text{-}1$$

Noting that $=\sum_{x=0}^{\infty}\frac{\lambda^y}{y!}$ converges to e^{λ}, so that from **, we have that

$$E(X)=\lambda e^{-\lambda}e^{\lambda}=\lambda$$

For the variance, let us have
$X^2 = X^2 - X + X = X(X-1) + X$
Then $E(X^2) = E(X(X-1)+E(X) = E(X(X-1))+\lambda$, where

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

$$=\lambda^2e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}=\lambda^2e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^y}{y!}\text{ ,where } y=x\text{-}2$$

$= \lambda^2$ for similar reason as in E(X) above

Now $E(X^2) = \lambda^2+\lambda$

Therefore variance $\sigma^2(X) = \lambda^2+\lambda - \lambda^2 = \lambda$

The Poisson Approximation to Binomial Distribution

If in the binomial distribution, n is large while the probability p of occurrence of an event is close to zero, so that the 1-p is close to 1, then the event is called rare event. In practice, we shall consider an event as rare if the number of trials is at least 50 (n≥50), while np<5. For such case, the binomial distribution is closely approximated by the Poisson distribution with λ=np.

Activity 3

Let us suppose that the probability of a transistor manufactured by a certain firm being defective is 0.015. What is the probability that there is no defective transistor in a batch of 100 using
(i) Binomial distribution
(ii) Poisson approximation to Binomial

Solution:
Let us denote that X is the number of flaws in 1 millimeter of wire. Then, E(X) = 2.3

(i) Using Binomial distribution:
Let X be the number of defective transistors in 100.
Given n=100, p = 0.015
Then the required probability is

$$\Pr(X=0)=\binom{100}{0}(0.015)^0(0.985)^{100}$$

$$=0.221$$

(ii) Using Poisson approximation to binomial

Let $\lambda=np=100(0.015)$

$= 1.5$

The required probability is

$$\Pr(X=0) = \frac{e^{-1.5}1.5^{0}}{0!} = 0.223$$

Clearly, from above, the two results are the same

Summary

Definition of Poisson random variable was given and the assumption were stated. The mean and variance were derived. Numerical examples were given. Poisson approximation to binomial was discussed and numerical examples was given.

Self Assessment Questions

1. Derive the mean and variance of Poisson distribution with average number of event m
2. Let us suppose that the average number of error per page of a manuscript as reported by a reviewer is 6. What is the probability that on a given page, between 2 and 5 errors will be found?
3. If approximately 2% of people in a room of 200 people are left-handed, find the probability that exactly five people are left-handed.
4. It is known from then past that in a certain plant, thee are on the average 4 industrial accidents per month
   (a) Give the probability mass function
   (b) Find the probability that in a given year, there will be less than 4 accidents

Tutor Marked Assignment

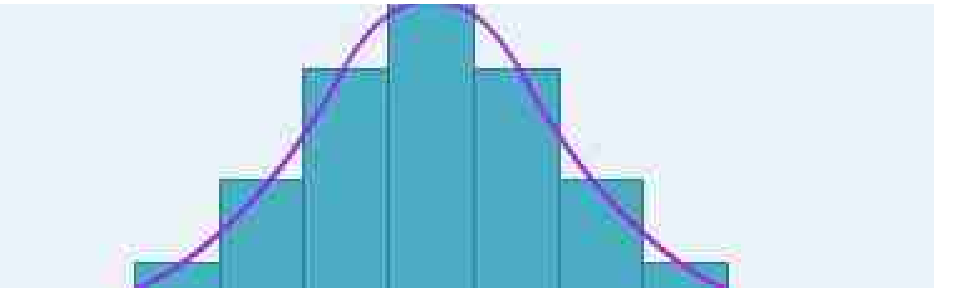
1. A sales firm receives on the average three call per hour on its toll-free number. For any given hour, find the probability that it will receive the following.
   (a) At most three calls
   (b) At least three calls
2. Let X be a Poisson distributed random variable with parameter λ. Given that $\Pr(X=2) = \Pr(X=3)$, find the
   (a) Pdf of X
   (b) $\Pr(X > 4)$
   (c) $\Pr(2 < X < 5)$
3. If 5% of the electronic bulbs manufactured by a company are defective, use Poisson approximation to Binomial to find the probability that in a sample of 100 bulbs,
   (a) none will be defective
   (b) five bulbs will be defective

References

(1) Ross, S.M. Introduction to Probability and Statistics for Engineers and Scientists, Elsevier academic Press.
(2) Montgomery, D. C. and Runger, G. C. Applied Statistics and Probability for Engineers. John Wiley & Sons
(3) Harry Frank, F. and Althoen. S. C. Statistics Concepts and Applications. Cambidge University Press.

Further Reading

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=11&ved=2ahUKEwjw9reom9XnAhU1onEKHf6LCok4ChAWMAB6BAgCEAE&url=https%3A%2F%2Fmei.org.uk%2Ffiles%2Fpdf%2FPoisson_Distribution_8.pdf&usg=AOvVaw3WPiIq6pXi2jEAUIyhI2Lv
https://www.studocu.com/en-ca/document/the-university-of-british-columbia/fundamentals-of-biostatistics/lecture-notes/lecture-notes-lecture-3-binomial-and-poisson-distributions/218374/view



UNIT 4 NEGATIVE BINOMIAL AND GEOMETRIC DISTRIBUTIONS

Introduction

We are going to be considering another topic in this unit course. First I want you to note that Negative Binomial distribution has similar properties as for the binomial distribution, with exception that the trials will be repeated until a fixed number of successes occur. Therefore, while binomial distribution is the distribution of the number of successes in a fixed number of independent Bernoulli trials, negative binomial distribution is the distribution of the number of trials needed to get a fixed number of successes. It can also be defined as the number of failures before a fixed number of successes.

Learning Outcomes

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

1. Define negative binomial and geometric distributions and use them to calculate some probabilities.
2. Derive the formula for the negative binomial and geometric probability mass functions.
3. Explore the key properties, such as the mean and variance, of a negative binomial and geometric random variables.

Main Content

Negative Binomial Random variable

4mins

SAQ 1

SAQ 2

SAQ 3

The number X of trials before producing k successes in a negative binomial experiment is called a negative binomial random variable, and its probability distribution is called negative binomial distribution.

Illustrative example: Suppose the probability that a bit transmitted through a digital transmission channel is received in error is 0.1. Assume the transmissions are independent events, and let the random variable **X** denote the number of bits transmitted until the **fourth** error. Then, **X** has a negative binomial distribution with **r** = 4. Probabilities involving **X** can be found as follows. The Pr(X=10) is the probability that exactly three errors occur in the first nine trials and then trial 10 results in the fourth error. Because the trials are independent, the probability that exactly three errors occur in the first 9 trials and trial 10 results in the fourth error is the product of the probabilities of these two events namely,

$$\Pr(X=10)=\binom{10-1}{4-1}(0.1)^4(0.9)^{10-4},$$

$$\Pr(X=10)=\binom{9}{3}(0.1)^4(0.9)^6,$$

$$=\ 42\times0.0001\times0.53144=0.0022$$

Probability Density function of Negative Binomial Distribution

We can formally conceptualize negative binomial distribution as follows. In a series of Bernoulli trials (independent trials with constant probability of success), let the random variable X denote the number of trials until r successes, or equivalently, the number of failures before r successes occur. Then X is a negative binomial random variable with parameter 0<p<1 and r = 1, 2, 3, ……, with pdf given by

$$f(x)=\Pr(X=x)=\binom{x-1}{r-1}p^r q^{x-r},\quad r, r+1, r+2,\ldots,\ \text{where } q = 1\text{-}p.$$

Mean and Variance of Negative binomial random variable

If X is a negative binomial random variable with parameter p and r,

then the mean $\mu=E(X)=\frac{r}{p}$ and $\sigma^2=Var(X)=\frac{rq}{p^2}$

Activity 1

1. In an opinion poll on voters exiting from a polling booth and asking them if they voted for a female Presidential candidate. The probability (p) that a person voted for a female candidate is 20%
   (a) What is the probability that 15 voters will be asked before you can find 5 people who voted for a female candidate?
   (b) Compute the expected number of voters to be interviewed before finding 5 people who voted for female candidates and the variance.

Negative binomial and geometric distributions Unit 4 | Module 4 | STA 221

Solution:
Let X be the random variable denoting the number of voters interviewed before finding 5 voters who voted for a female presidential candidate. We have that number of success k =5 and the number of failure before obtaining the 5 successes is
x - k =15-5 =10.
P= 0.2, q =1-p = 0.8

(a) Thus the required probability is

$$\Pr(X=5)=\binom{14}{4}(0.2)^5(0.8)^{10}=0.034$$

(b) The expected number of voters is

$$E(X)=\frac{5}{0.2}=25 \text{ and } Var(X)=\frac{5(0.8)}{(0.2)^2}=100$$

2. Suppose that independent tosses of a biased coin are made until r heads are observed, where the probability of observing a head on any given toss is 0.25. Let X be the random variable denoting the number of trials required to obtain the rth head.
   (a) Give the appropriate probability density function for this experiment.
   (b) Given that r =3 in question 2(a) above, find the probability that
      (i) 3 heads are observed on the sixth trial.
      (ii) at least 6 trials are required to observe 3 heads.

Solution:
p = 0.25, 1-p =0.75

(a) $\Pr(X=x)=\binom{x-1}{2}0.25^r0.75^{x-r},\quad$ x=r, r+1, r+2,…

(b) when r =3, the pdf is given as

$$\Pr(X=x)=\binom{x-1}{2}0.25^2 0.75^{x-2},\quad 3, 4,\ldots$$

Required to find

(i) $\Pr(X=6)=\binom{x-1}{2}0.25^3 0.75^{x-3},$

$$=\binom{5}{2}0.25^3 0.75^3$$
$$=0.066$$

Activity 2

On each randomly dialed number, there is a 9% chance of reaching an adult who will complete the survey.

(a) What is the probability that the 3rd completed survey occurs on the 10th call?

(b) Determine the average number of telephone calls required to get 3 completed survey and the variance.

Solution:

(a) Let X be a random variable denoting the number of calls required to complete or more surveys.

We have r =3, X=10, p = 0.09

$$\Pr(X=10)=\binom{10-1}{3-1}0.09^3(1-0.09)^{10-3},$$

$$\Pr(X=10)=\binom{9}{2}0.09^3(1-0.09)^7,$$

$$= 0.0136$$

(b) $E(X)=\dfrac{3}{0.09}=33.3$ and $Var(X)=\dfrac{3(0.91)}{(0.09)^2}=337.03$

Geometric Distribution

Geometric distribution is a limiting distribution of negative binomial distribution. The experiment consists of one or more Bernoulli trials with all failures except the last one, which is a success. This means that one keeps repeating the experiment until you obtain the first success, then you stop. For example, you toss a coin a number of times independently, and you define the first time you obtain a head as a "success" so you stop tossing the coin. It might take six trials until you obtain the first success. Suppose we denote the head by H and the tail by T, then for the six trials we may have T, T, T, T, T, H. You can also think of the trials as failure, failure, failure, failure, failure, success.

Generally, we can define a geometric random variable X as the number of trials until the first success is obtained.

Conditions for a geometric distribution

(i) An experiment consists of repeated trials until first success occurs.
(ii) There are two possible outcomes for each trial.
(iii) Repeated trials are independent.
(iv) The probability of success p is constant for all trials

Other examples of geometric distribution include

(a) Terminals on an on-line computer system are attached to a communication line to the central computer system. The probability that any terminal is ready to transmit is 0.95. Let X = number of terminals polled until the first ready terminal is located.
(b) It is known that 20% of products on a production line are defective. Products are inspected until first defective is encountered. Let X = number of inspections to obtain first defective
(c) One percent of bits transmitted through a digital transmission are received in error. Bits are transmitted until the first error. Let X denote the number of bits transmitted until the first error.

Definition: Consider a repeated independent trials where the outcome of each trial can result in a success with probability p and a failure with probability q =1-p, then the probability distribution of a random variable X, the number of trials before the first successes occurs is given as

$$geom(x,p)=p^k q^{x-1}, \quad 1,2,3,\ldots$$

The random variable defined above denotes the number of trials required to obtain the first success
The mean and variance of geometric distribution are

$$\text{The mean } E(X)=\frac{1}{p}, \text{ and } Var(X)=\frac{1-p}{p^2}$$

Activity 1

A representative from the National Football League's Marketing Division randomly selects people on a random street in Kansas City, Kansas until he finds a person who attended the last home football game. Let p, the probability that he succeeds in finding such a person be 0.20, and, let X denote the number of people he selects until he finds his first success.

(a) What is the probability that the marketing representative must select 4 people before he finds one who attended the last home football game?

(b) Find the average number of people the marketing representative needs to select before he finds one who will attend the last home football game and the standard deviation

Solution:

Since X denotes the number of people he selected until he finds his first success, then using Geometric distribution, the required probability is

(a) $\Pr(X=4)=0.20(0.80)^{4-1}$
$=0.20(0.80)^3$
$= 0.1024$

(b) The average number is $\frac{1}{p}=\frac{1}{0.20}=5$

That is, we should expect the marketing representative to select 5 people before he finds one who will attend the last football game.

The variance is

$$\sigma = \text{Var(X)} = \sqrt{\frac{1-p}{p^2}}$$

$$=\frac{0.08}{0.20^2}=4.472$$

Summary

Upon completion of this course unit, we have defined Negative binomial and geometric random variables and their probability density functions stated. These were used to calculate some probabilities. We explored some properties of negative binomial and geometric distributions, including the mean and variance.

Self Assessment Questions

1. Distinguish between
   (a) Negative binomial and Binomial distributions
   (b) Negative binomial and geometric distributions
2. An oil company conducts a geological study that indicates that an exploratory oil well should have a 20% chance of striking oil. What is the probability that the first strike comes on the third well drilled?
3. An electronic scale in an automated filling operation stops the manufacturing line after three underweight packages are detected. Suppose that the probability of an underweight package is 0.001

and each fill is independent
   (a) What is the mean number of fills before the line is stopped?
   (b) What is the standard deviation of the number of fills before the line is stopped?
4. State the conditions for geometric distribution
5. A student's summer job is to call the university alumni for support for the university's scholarship fund. Studies indicate that the probability that each of the student's calls is answered is 1/3. What is the probability that the second call to be answered on one particular day is the student's sixth call?

Tutor Marked Assignment

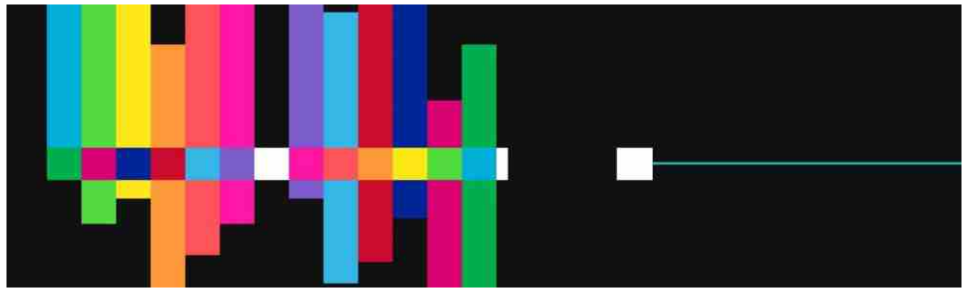
1. A test of weld strength involves loading welded joints until a fracture occurs. For a certain type of weld, 80% of the fractures occur in the weld itself, while the other 20% occur in the beam. A number of welds are tested and the tests are independent
   (a) If X is the number of test at which the first beam fracture is observed, find the probability that the first beam fracture happens on the third trial or later.
   (b) Suppose that the trial is continued, find the probability that the 3rd beam fracture (success) occurs on the 6th trial.
   (c) If X represent the number of trials until 3 beam fractures occur, find the mean and variance of X
2. Each of calls to a popular television station has a probability of 0.98 of obtaining a signal. Assume that calls are independent.
   (a) What is the probability that the first call obtains a signal ?
   (b) What is the probability that it requires more than five calls for a signal to be received?
   (c) What is the mean number of calls needed for a signal to be receive?
3. Suppose that X is a negative binomial random variable with p = 0.4 and r =8. Determine the following.
   (a) E(X)
   (b) Var(X)
   (c) Pr(X = 15)
4. A Sales representative is selling a newly introduced product from house to house in a suburban town. The probability that he sells a product at any house he visits is 0.4. If he visits eight houses in a day. What is the probability that he will sell his third product in the sixth house he visits?

References

1. Montgomery, D.C. and Runger, G.C. Applied Statistics and Probability for Engineers. Third Edition. John Wiley & Sons, Inc.
2. Roussas. G. G. A Course in Mathematical Statistics. Academic Press, New York
3. Soong. T.T. Fundamentals of Probability and statistics for Engineers.(2004). John Wiley & Sons Inc.

Further Reading

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=2ahUKEwi15t3ZtdXnAhUzrHEKHQ64BJoQFjAFegQIChAB&url=http%3A%2F%2Fwww.math.umd.edu%2F~millson%2Fteaching%2FSTAT400%2Fslides%2Farticle8.pdf&usg=AOvVaw0y4PNHw9jXtvPvDtzkqYSf
https://online.stat.psu.edu/stat414/node/80/



UNIT 5 UNIFORM DISTRIBUTION

Introduction

I welcome you to this course unit. Uniform distribution is often used to model events which are equally likely to occur. It is defined by two parameters, say a and b, where a = minimum value and b = maximum value. This can be denoted (a, b). For example, a deck of cards has within it uniform distributions because the likelihood of drawing a heart, a club, a diamond or a spade is equally likely. A coin also has a uniform distribution because the probability of getting either heads or tails in a coin toss is the same.

The uniform distribution can be visualized as a straight horizontal line, so for a coin flip returning a head or tail, both have a probability p = 0.50 and would be depicted by a line from the y-axis at 0.50.

Learning Outcomes

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

1. Define discrete uniform random variable and state the probability density function.
2. Describe the mean and variance of discrete uniform random variable and solve some numerical examples
3. Define continuous uniform random variable and state the probability density function
4. Derive the mean and the variance of continuous uniform random variable and solve some numerical problems.

Main Content

Discrete Uniform Distribution

2mins

A random variable X has a discrete uniform distribution if each of the n values in its range, say $x_1, x_2, \ldots, x_n$ has equal probability . The pdf is thus given as

$$f(x) = 1/n$$

Example

The first digit of a part's serial number is equally likely to be any one of the digits 0 through 9. If one part is selected from a large batch and X is the first digit of the serial number, then X has a discrete uniform distribution with probability 0.1 for each value in R = {0, 1, ..., 9}. That is,

f(x) = 0.1, for each value in R.

Mean and variance of Discrete uniform distribution

SAQ 1

SAQ 2

Suppose X is a discrete random variable on the consecutive integers a, a+1,a+2,...,b, for a ≤ b,
then the mean of X is

$$E(X)=\frac{b+a}{2},$$

and the variance of X is

$$Var(X)=\frac{(b-a+1)^2-1}{12}=\frac{n^2-1}{12}$$

where n = b - a + 1 denotes the number of values that X may take.

Activity 1

A voice communication system for a business contains 48 external lines. At a particular time, the system is observed, and some of the lines are being used. Let the random variable X denote the number of lines in use. Then, X can assume any of the integer values 0 through 48. The mean of X is thus obtained as

$$E(X)=\frac{0+48}{2}=24,$$

and the variance of X is

$$Var(X)=\frac{(48-0+1)^2-1}{12}=200$$

If P is the proportion of the voice lines that are in use at a particular time from activity 11, find the mean and variance of P

Solution:
P = X/48,
so that E(P) = E(X/48)
= E(X)/48
= 24/48 = 0.5,
Var(P) = Var(X/48)
= Var(X)/482
= 200/482 = 0.087

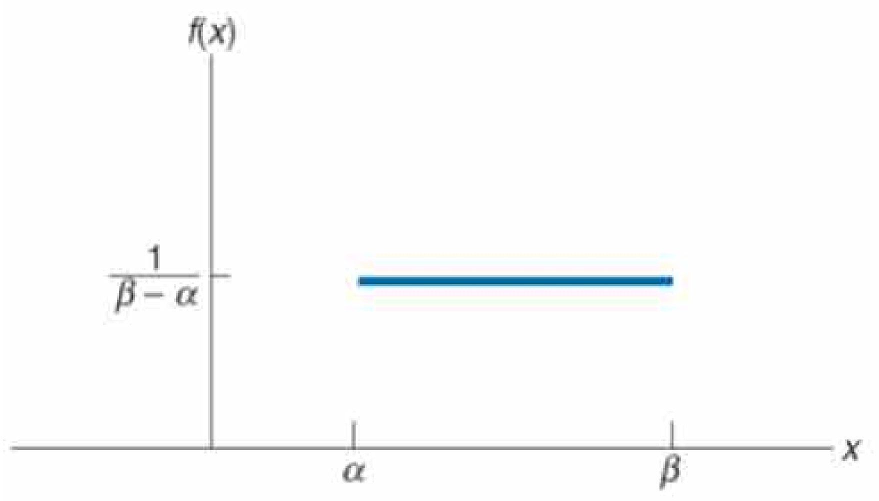
Continuous Uniform Distribution

SAQ 3

The simplest continuous distribution is analogous to its discrete counterpart. Thus a continuous random variable X is said to have a uniform distribution if the pdf is given as

$$f(x)=\frac{1}{\beta-\alpha},\alpha\le x\le\beta$$

The uniform distribution arises in a situation when we suppose a certain random variable is equally likely to be near any value in the interval [α, β]. The probability that X lies in any subinterval of [α, β] is equal to the length of that subinterval divided by the length of the interval [α, β]. A graph of the probability distribution of continuous uniform distribution is given below.

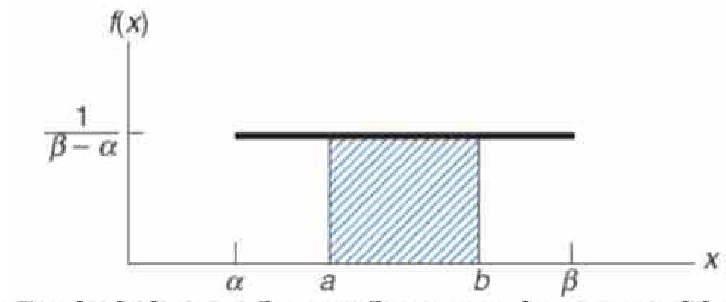


Graph of f (x) for a uniform [α, β].

It should be noted that the uniform distribution also meets the requirements of being a probability density function since

$$\int_{\alpha}^{\beta}\frac{x}{\beta-\alpha}dx=1$$

The probability of a continuous uniform random variable X can be illustrated as follows
If X is a continuous uniform random variable which lies between intervals a and b as given in the graph below, then



Probabilities of a uniform random variable.

$$\Pr\{a<X<b\} = \frac{1}{\beta-\alpha}\int_a^b dx$$
$$= \frac{b-a}{\beta-\alpha}$$

Mean and variance of a continuous uniform distribution

The mean of X is given as

$$E(X) = \int_\alpha^\beta \frac{x}{\beta-\alpha}dx$$
$$= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} = \frac{(\beta-\alpha)(\beta+\alpha)}{2(\beta-\alpha)}$$
$$= \frac{\beta+\alpha}{2}$$

and the variance of X is

$$Var(X) = \int_\alpha^\beta \frac{\left(\frac{x-(\alpha+\beta)}{2}\right)^2}{\beta-\alpha}dx = \frac{(\beta-\alpha)^2}{12}$$

Activity 2

Let the continuous random variable X denote the current measured in a thin copper wire in Milliamperes (mA). Assume that the range of X is [0, 20 mA], and assume that the probability density function of X is f(x) = 0.05, $0 \le x \le 20$.

(i)What is the probability that a measure of current is between 5 and 10 millampere?

(ii) Find the mean current and the variance.

Solution:

(i) $\Pr(0<x<10) = \int_5^{10} 0.05dx$

$$= [0.05x]_5^{10} = 0.25$$

(ii) a=0, b=20, so that E(X) = (0+20)/2 =10mA

and Var(X) = $\frac{(20-0)^2}{12} = 33.33mA^2$

The cumulative distribution function (CDF)

The cumulative density function of a continuous uniform random variable over the interval [a,b] is obtained as

$$F(x) = \int_a^x \frac{1}{b-a}du$$
$$= \frac{x-a}{b-a}$$

Therefore, the complete description of the cumulative distribution function of a continuous uniform random variable is

$$\begin{bmatrix} 0 & x<a \\ \frac{x-a}{b-a} & a\le x<b \\ 1 & x\ge b \end{bmatrix}$$

Activity 3

Obtain the cumulative density function of Activity 2

Solution:

$$F(x) = \int_a^x \frac{1}{b-a}du$$
$$= \int_0^x \frac{1}{20}du$$
$$= \left.\frac{u}{20}\right|_0^x$$

$$F(x) = \begin{bmatrix} 0 & x<0 \\ \frac{x}{20} & 0\le x<20 \\ 1 & x\ge 20 \end{bmatrix}$$

Summary

We have carefully defined discrete uniform distribution. We also stated the means and the variance and we solved some numerical problems. Also, we defined continuous uniform distribution for random variable and solved numerical problems.

Self Assessment Questions

1. Let the random variable X have a discrete uniform distribution on the integers, 1 < x < 3. Determine the mean and variance of X
2. The time that Joe, the teaching assistant, takes to grade a paper is

uniformly distributed between 5 minutes and 10 minutes. Find the mean and variance of the time he takes to grade a paper.

3. Consider the continuous random variable X which is uniformly distributed on the interval [a,b]. Derive the mean and the variance of X.
4. If X is uniformly distributed over the interval [0, 10], compute the following probabilities
   (a) Pr(2 < X < 9)
   (b) Pr(X < 5)
   (c) Pr(X > 6)

Tutor Marked Assignment

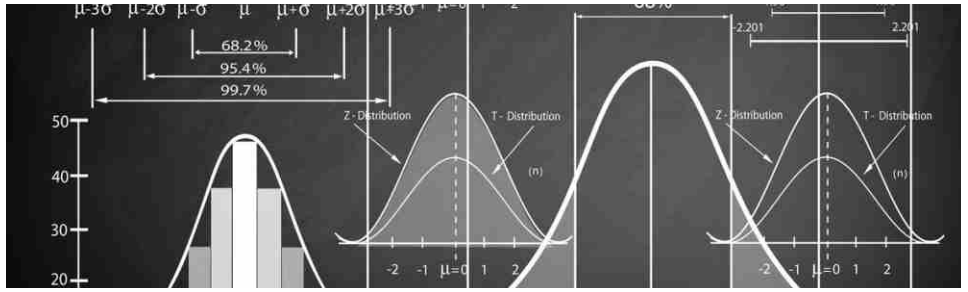
1. Suppose the time it takes a data collection operator to fill out an electronic form for a database is uniformly between 1.5 and 2.2 minutes.
   (a) What is the mean and variance of the time it takes an operator to fill out the form?
   (b) What is the probability that it will take less than two minutes to fill out the form?
   (c) Determine the cumulative distribution function of the time it takes to fill out the form
2. Let X be a discrete uniform random variable that denotes the outcome of the roll of a fair die.
   (a) State the pmf of X
   (b) Find the mean and the variance of X
3. Suppose that customers arrive a bank at a time that is uniformly distributed over the interval (0,30). If Y denotes the arrival time, find
   (a) Pr(25 ≤ Y ≤ 30)
   (b) E(Y)
   (c) Var(Y)

References

(1) Frank, H & Steven C. Althoen. Statistics Concepts and Applications. Cambridge University Press.
(2) T.T. Soong. Fundamentals of Probability and statistics for Engineers. John Wiley & Sons Inc.
(3) George G. Roussas. A Course in Mathematical Statistics. Academic Press, New York

Further Reading

https://www.pdfdrive.com/applied-statistics-and-probability-for-engineers-5e-e33357551.html



UNIT 6 NORMAL DISTRIBUTION

Introduction

I welcome you to this course unit, I want you to note that the normal distribution is the single most important distribution in statistics. If a random sample is taken from a normal distribution, then the distributions of various important functions of the of the observations in the sample can be derived explicitly and will themselves have simple forms. Therefore, it is mathematically convenient to be able to assume that the distribution from which a random sample is drawn is a normal distribution. Secondly, most of the random variables studied by scientists in physical experiments often have distributions that are approximately normal. Thirdly, central limit theorem. If a large random sample is taken from the distribution, even though the distribution is not itself approximately normal, the distribution of the sample mean will be approximately normal.

Learning Outcomes

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

1. Define normal distribution of a random variable and use it to compute probabilities
2. Derive the mean and variance using method of moment generating function
3. Describe normal approximation to Poisson distribution and use it to compute some probabilities
4. Describe normal approximation to Binomial distribution and use it to compute some probabilities

Main Content

Normal Distribution

3mins

SAQ 1

SAQ 2

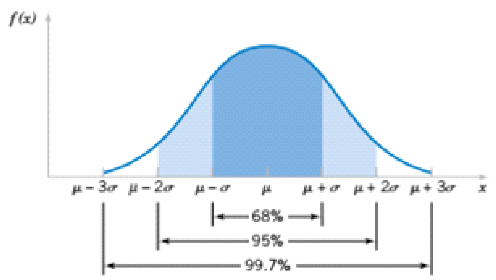
A random variable X is said to have a normal distribution with mean μ and variance σ^2 if the pdf is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)}{2\sigma^2}}, \ -8 < x < 8 \ ; \ -8 < \mu < 8 \ ; \ \sigma^2 > 0$$

The normal distribution above is often denoted $N(\mu, \sigma^2)$.

Some useful results concerning a normal distribution are summarized in the figure below. For any normal random variable, it is shown that

$\Pr(\mu-\sigma<X<\mu+\sigma) = 0.6827$
$\Pr(\mu-2\sigma<X<\mu+2\sigma) = 0.9545$
$\Pr(\mu-3\sigma<X<\mu+3\sigma) = 0.9973$



Probabilities associated with a normal distribution

Standard normal random variable

If X is a normal random variable with $E(X) = \mu$ and $Var(X) = \sigma^2$, then

$Z = \dfrac{X-\mu}{\sigma}$ is called the standard normal random variable with

$E(Z) = 0$ and $Var(Z) = 1$.

The cumulative distribution function of a standard normal random variable is denoted as $\Phi(z) = \Pr(Z \le z)$

In order to find the probability of interest, the normal random variable X needs to be transformed to the standard normal random variable. Examples are as illustrated below,

(1) $\Pr(X = a) = \Pr\left(\dfrac{X-\mu}{\sigma} \le \dfrac{a-\mu}{\sigma}\right)$

$= \Pr(Z \le z_a)$

$= \Phi(z_a)$

(2) $\Pr(X = a) = \Pr\left(\dfrac{X-\mu}{\sigma} \ge \dfrac{a-\mu}{\sigma}\right)$

$= \Pr(Z \ge z_a)$

$= 1 - \Pr(Z \le z_a)$

$= 1 - \Phi(z_a)$

(3) $\Pr(a = X = b) = \Pr\left(\dfrac{a-\mu}{\sigma} \le \dfrac{X-\mu}{\sigma} \le \dfrac{b-\mu}{\sigma}\right)$

$= \Pr(z_a \le Z \le z_b)$

$= \Phi(z_b) - \Phi(z_a)$

Activity 1

Assume that the current measurements in a strip of wire follow a normal distribution with mean 10 milliamperes and variance 4 (milliamperes^2).

(a) What is the probability that a current measurement

(i) exceeds 13 milliamperes?

(ii) is between 9 and 11 milliamperes

(b) Determine the value for which the probability that a current measurement is below this value is 0.98

Solution:

a. (i) Let X denote the current measurements in milliamperes.

The mean $\mu = 10$ and variance $\sigma^2 = 4$, so that $\sigma = 2$

Using example 2 above, the probability can be written as

$\Pr(X>13) = \Pr\left(\dfrac{X-10}{2} \ge \dfrac{13-10}{2}\right)$

$= \Pr(Z \ge 1.5)$

$= 1 - \Pr(Z \le 1.5)$

$= 1 - \Phi(1.5)$

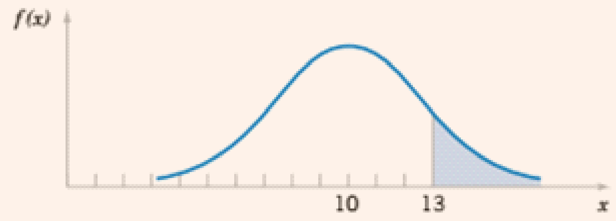
From the standard normal table in Appendix I, the required probability is

$= 1 - 0.93319$

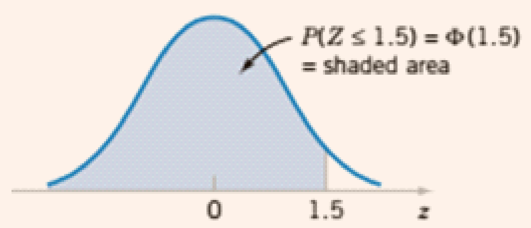
$= 0.6681$

The above probability is shown as the shaded area under the normal probability density function in the figures below

(a)



(b)



(ii) We shall compute the Z values as follows

$$\Pr(9 \le X \le 11) = \Pr\left(\frac{9-10}{2} \le \frac{X-10}{2} \le \frac{11-10}{2}\right)$$

$$= \Pr(-0.5 \le Z \le 0.5)$$

$$= \Phi(0.5) - \Phi(-0.5)$$

From the standard normal table in Appendix I, the required probability is

0.6915 - 0.3085 = 0.3830

(iii) We need the value of x such that Pr(X<x) = 0.98. By standardizing, this probability expression can be written as

$$\Pr\left(\frac{X-10}{2} < \frac{x-10}{2}\right)$$

$$= \Pr\left(Z < \frac{x-10}{2}\right) = 0.98$$

From the Table in Appendix I, we need to find the z value such that Pr(Z < z) = 0.98. The nearest probability from the table is

Pr(Z < 2.05) =0.97982

This implies that

$$\Pr\left(\frac{x-10}{2}\right) = 2.05$$

And from the standardized transformation, we reverse the expression to have

x= 2(2.05)+10 = 14.1 milliamperes

Derivation of Mean and Variance of Normal Distribution

I want you to bear in mind that the mean and variance of a normal distribution can be derived using the moment generating function approach.

Recalling the pdf

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

The moment generating function is given by

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

Define $Z = \frac{X-\mu}{\sigma}$, where Z~N(0,1)

and $X = Z\sigma + \mu$,

Thus $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$, and

$$M_x(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{t(z\sigma+\mu)} e^{-\frac{1}{2}z^2} dz$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zt\sigma - 2\mu t)} dz$$

$$= e^{\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zt\sigma)} dz \qquad (***)$$

By completing the square, (***) can be written as

$$= e^{\mu t} e^{\frac{1}{2}t^2\sigma^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z^2 - 2zt\sigma + t^2\sigma^2)} dz$$

$$= e^{\mu t} e^{\frac{1}{2}t^2\sigma^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z - t\sigma)^2} dz$$

$$M_x(t) = e^{\mu t} e^{\frac{1}{2}t^2\sigma^2}$$

(since the quantity under the integral is a pdf with value 1)

To obtain the mean and the variance

$$M'_x(t) = \frac{dM_x(t)}{dt} = (\mu + \sigma^2 t) e^{ut + \frac{1}{2}\sigma^2 t^2)}$$

E(X) = $M'_x(0) = (\mu + \sigma^2(0)) e^{u(0) + \frac{1}{2}\sigma^2(0)}$

$= \mu$

For variance

$$M''_x(t) = \frac{d^2 M_x(t)}{dt^2} = (\mu + \sigma^2 t) M'_x(t) + M_x(t) \frac{d(\mu + \sigma^2 t)}{dt}$$

$$= (\mu + \sigma^2 t)^2 e^{ut + \frac{1}{2}\sigma^2 t^2} + e^{ut + \frac{1}{2}\sigma^2 t^2} (\sigma^2)$$

$$M''_x(0) = \mu^2 + \sigma^2$$

$$\text{Var(X)} = M''_x(0) - (M'_x(0))^2 = \mu^2 + \sigma^2 - \mu^2$$
$$= \sigma^2$$

Normal Approximation to Poisson Distribution

The normal distribution can be approximated to the Poisson distribution when λ is large, best when λ > 20.
If X is the Poisson random variable with pdf given by

$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ $x = 0, 1, \ldots$, and the corresponding standardized random variable is $Z = \frac{X - \mu}{\sqrt{\sigma}} = \frac{X - \lambda}{\sqrt{\lambda}}$, knowing that the mean and variance of a Poisson distribution are the same.

then

$\Pr(a \le X \le b)$ can be obtained by the transformation

$$\Pr(z_a \le \frac{X - \lambda}{\sqrt{\lambda}} \le z_b)$$

$$\Pr(z_a \le Z \le z_b)$$

$$\Phi(z_b) - \Phi(z_a)$$

Activity 2

In a country there are 36 fatal accidents per year and the number of accidents per year follows a Poisson distribution. Use the normal approximation to find the probability that there are

(a) more than 48 accidents in a year.
(b) Between 40 and 50 accidents in a year

Solution:
Because λ > 20 a normal approximation can be used.
Let X be the random variable of the number of accidents per year.
For the normal approximation $\mu = \sigma^2 = \lambda = 36$
and $\sigma = \sqrt{\lambda} = 6$
(a) The probability that there are more than 48 accidents in a year is thus

$$\Pr(X>48) = \Pr(\frac{X - 36}{6} \ge \frac{48 - 36}{6})$$

$$\Pr(Z \ge 2) = 1 - \Pr(Z \le 2)$$
$$= 1 - \Phi(2)$$

This is obtained from the standard normal table in Appendix I as
= 1-0.9772 = 0.0228

$$\Pr(40 \le X \le 50) = \Pr(\frac{40 - 36}{6} \le \frac{X - 36}{6} \le \frac{50 - 36}{6})$$
$$= \Pr(0.67 \le Z \le 2.33)$$
$$= \Phi(2.33) - \Phi(0.67)$$
$$= 0.9901 - 0.7486$$
$$= 0.2415$$

Normal Approximation to Binomial Distribution

When the probability of success p is approximately 0.5, and as the number of trials n increases, the shape of the binomial distribution becomes similar to that of the normal distribution. The larger n is and the closer p is to 0.5, the more similar the shape of binomial is to that of the normal distribution. The approximation is best for p = 0.5 and variance npq ≥ 20.

Recall that the pdf of binomial random variable is given as

$$f(x) = \binom{n}{x} p^x (1 - p)^{n - x}, \quad x = 1, \ldots, n$$

and the mean and variance of X are respectively, $\mu = np$ and $\sigma^2 = np(1-p)$. Then the corresponding standardized random variable is

$\Pr(a \le X \le b)$ can be obtained by the transformation

$$\Pr(z_a \le \frac{X - np}{\sqrt{np(1 - p)}} \le z_b)$$

$$\Pr(z_a \le Z \le z_b)$$

$$\Phi(z_b) - \Phi(z_a)$$

Activity 3

A magazine reported that 48% of American drivers read the newspaper while driving. If 300 drivers are selected at random, find the probability that

(a) not more than 125 say they read the newspaper while driving.

(b) between 120 and 150 say they read the newspaper while driving

Solution:

Here n= 300, p = 0.48 (close to 0.5), thus normal approximation to binomial distribution is appropriate.

Now, μ= 300(0.48) = 144 and σ^2 = 300(0.48)(0.52) = 74.88

(a) Required to find $\Pr(X \le 125)$

$$= \Pr\left(\frac{X-144}{\sqrt{74.88}} \le \frac{125-144}{\sqrt{74.88}}\right)$$

$$= \Pr(Z \le -2.20) = 0.0139$$

$$= \Pr(120 \le X \le 150) = \Pr\left(\frac{120-144}{\sqrt{74.88}} \le \frac{X-144}{\sqrt{74.88}} \le \frac{150-144}{\sqrt{74.88}}\right)$$

$$= \Pr(-2.77 \le Z \le 0.69)$$

$$= \Phi(0.69) - \Phi(-2.77)$$

$$= 0.7549 - 0.0028 = 0.7521$$

Summary

We have successfully defined Normal random variable upon completion of this unit. Recall that I said it is used to solve some simple probability problems whose random variables follows normal distributions. We illustrated normal approximation to Poisson distribution with example. Finally we derived the mean and Variance using moment generating function approach.

Self Assessment Questions

1. Derive the mean and variance of a random variable X which follows a normal distribution
2. Each month, an American household generates an average of 28 pounds of newspaper for garbage or recycling. It is assumed that the amount generated is normally distributed and the standard deviation is 2 pounds. If a household is selected at random, find the probability of its generating
   (a) Between 27 and 31 pounds per month
   (b) More than 30.2 pounds per month
3. The length of an injection-molded plastic case that holds magnetic tape is normally distributed with a length of 90.2 millimeters and a standard deviation of 0.1 millimeter.
   (a) What is the probability that a part is longer than 90.3 millimeters or shorter than 89.7 millimeters?
   (b) What should the process mean be set at to obtain the greatest number of parts between 89.7 and 90.3 millimeters?
   (c) If parts that are not between 89.7 and 90.3 millimeters are scrapped, what is the yield for the process mean that you selected in part (b)?
4. A radioactive element disintegrates such that it follows a Poisson distribution. If the mean number of particles (λ) emitted is recorded in a 1 second interval as 69, use Normal approximation to Poisson distribution to evaluate the probability that:
   (i) less than 60 particles are emitted in 1 second.
   (ii) between 65 and 75 particles inclusive are emitted in 1 second.
5. From many years of observation, a biologist knows that the probability is only 0.65 that any given Arctic tern will survive the migration from its summer nesting area to its winter feeding grounds. A random sample of 500 Arctic terns were banded at their summer nesting area. Use normal approximation to binomial distribution to approximate the probability that between 310 and 340 of the banded Arctic terns will survive the migration?

Tutor Marked Assignment

1. Suppose that X has the normal distribution with mean 5 and standard deviation 2. Determine Pr(1<X<8)
2. The time it takes a cell to divide (called mitosis) is normally distributed with an average time of one hour and a standard deviation of 5 minutes.
   (a) What is the probability that a cell divides in less than 45 minutes?

(b) What is the probability that it takes a cell more than 65 minutes to divide?

(c) What is the time that it takes 99% of all cells to complete mitosis?

3. Suppose that X is a binomial random variable with n = 200 and p = 0.4

   (a) Approximate the probability that X is less than or equal to 70

   (b) Approximate the probability that X is greater than 70 and less than 90

4. Suppose that X is a binomial random variable with probability of success in a trial p = 0.05 and the number of independent trials n = 100. Use normal approximation to binomial distribution to find $Pr(X \geq 1)$

References

(1) Spiegel, M. R., Schiller, J and Srinivasan, R. A. Schaum's Outlines: Probability and Statistics. McGraw-Hill Company

(2) Woodroofe, M. Probability with Applications. McGraw-Hill.

Further Reading

https://www.pdfdrive.com/basic-probability-theory-robert-b-ash-department-of-mathematics-e8518230.html

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.00005	.00005	.00004	.00004	.00004	.00004	.00004	.00004	.00003	.00003
-3.8	.00007	.00007	.00007	.00006	.00006	.00006	.00006	.00005	.00005	.00005
-3.7	.00011	.00010	.00010	.00010	.00009	.00009	.00008	.00008	.00008	.00008
-3.6	.00016	.00015	.00015	.00014	.00014	.00013	.00013	.00012	.00012	.00011
-3.5	.00023	.00022	.00022	.00021	.00020	.00019	.00019	.00018	.00017	.00017
-3.4	.00034	.00032	.00031	.00030	.00029	.00028	.00027	.00026	.00025	.00024
-3.3	.00048	.00047	.00045	.00043	.00042	.00040	.00039	.00038	.00036	.00035
-3.2	.00069	.00066	.00064	.00062	.00060	.00058	.00056	.00054	.00052	.00050
-3.1	.00097	.00094	.00090	.00087	.00084	.00082	.00079	.00076	.00074	.00071
-3.0	.00135	.00131	.00126	.00122	.00118	.00114	.00111	.00107	.00104	.00100
-2.9	.00187	.00181	.00175	.00169	.00164	.00159	.00154	.00149	.00144	.00139
-2.8	.00256	.00248	.00240	.00233	.00226	.00219	.00212	.00205	.00199	.00193
-2.7	.00347	.00336	.00326	.00317	.00307	.00298	.00289	.00280	.00272	.00264
-2.6	.00466	.00453	.00440	.00427	.00415	.00402	.00391	.00379	.00368	.00357
-2.5	.00621	.00604	.00587	.00570	.00554	.00539	.00523	.00508	.00494	.00480
-2.4	.00820	.00798	.00776	.00755	.00734	.00714	.00695	.00676	.00657	.00639
-2.3	.01072	.01044	.01017	.00990	.00964	.00939	.00914	.00889	.00866	.00842
-2.2	.01390	.01355	.01321	.01287	.01255	.01222	.01191	.01160	.01130	.01101
-2.1	.01786	.01743	.01700	.01659	.01618	.01578	.01539	.01500	.01463	.01426
-2.0	.02275	.02222	.02169	.02118	.02068	.02018	.01970	.01923	.01876	.01831
-1.9	.02872	.02807	.02743	.02680	.02619	.02559	.02500	.02442	.02385	.02330
-1.8	.03593	.03515	.03438	.03362	.03288	.03216	.03144	.03074	.03005	.02938
-1.7	.04457	.04363	.04272	.04182	.04093	.04006	.03920	.03836	.03754	.03673
-1.6	.05480	.05370	.05262	.05155	.05050	.04947	.04846	.04746	.04648	.04551
-1.5	.06681	.06552	.06426	.06301	.06178	.06057	.05938	.05821	.05705	.05592
-1.4	.08076	.07927	.07780	.07636	.07493	.07353	.07215	.07078	.06944	.06811
-1.3	.09680	.09510	.09342	.09176	.09012	.08851	.08691	.08534	.08379	.08226
-1.2	.11507	.11314	.11123	.10935	.10749	.10565	.10383	.10204	.10027	.09853
-1.1	.13567	.13350	.13136	.12924	.12714	.12507	.12302	.12100	.11900	.11702
-1.0	.15866	.15625	.15386	.15151	.14917	.14686	.14457	.14231	.14007	.13786
-0.9	.18406	.18141	.17879	.17619	.17361	.17106	.16853	.16602	.16354	.16109
-0.8	.21186	.20897	.20611	.20327	.20045	.19766	.19489	.19215	.18943	.18673
-0.7	.24196	.23885	.23576	.23270	.22965	.22663	.22363	.22065	.21770	.21476
-0.6	.27425	.27093	.26763	.26435	.26109	.25785	.25463	.25143	.24825	.24510
-0.5	.30854	.30503	.30153	.29806	.29460	.29116	.28774	.28434	.28096	.27760
-0.4	.34458	.34090	.33724	.33360	.32997	.32636	.32276	.31918	.31561	.31207
-0.3	.38209	.37828	.37448	.37070	.36693	.36317	.35942	.35569	.35197	.34827
-0.2	.42074	.41683	.41294	.40905	.40517	.40129	.39743	.39358	.38974	.38591
-0.1	.46017	.45620	.45224	.44828	.44433	.44038	.43644	.43251	.42858	.42465
-0.0	.50000	.49601	.49202	.48803	.48405	.48006	.47608	.47210	.46812	.46414

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865	.99869	.99874	.99878	.99882	.99886	.99889	.99893	.99896	.99900
3.1	.99903	.99906	.99910	.99913	.99916	.99918	.99921	.99924	.99926	.99929
3.2	.99931	.99934	.99936	.99938	.99940	.99942	.99944	.99946	.99948	.99950
3.3	.99952	.99953	.99955	.99957	.99958	.99960	.99961	.99962	.99964	.99965
3.4	.99966	.99968	.99969	.99970	.99971	.99972	.99973	.99974	.99975	.99976
3.5	.99977	.99978	.99978	.99979	.99980	.99981	.99981	.99982	.99983	.99983
3.6	.99984	.99985	.99985	.99986	.99986	.99987	.99987	.99988	.99988	.99989
3.7	.99989	.99990	.99990	.99990	.99991	.99991	.99992	.99992	.99992	.99992
3.8	.99993	.99993	.99993	.99994	.99994	.99994	.99994	.99995	.99995	.99995
3.9	.99995	.99995	.99996	.99996	.99996	.99996	.99996	.99996	.99997	.99997