University of Venda

School of Mathematical and Natural Sciences

Department of Statistics

Module Code: STA1142

Module Description: Introductory Probability

Year: 2021

Lecturer: Mrs. T . H Tshisikhawe

1.COURSE CONTENT

- Introduction
- Counting techniques
- Probability and Relative frequencies, properties
- Addition rule and mutually exclusive events
- Conditional probability, Bayes Theorem and independence
- Random variables and probability distributions
- Binomial, Poisson and Normal distributions, Binomial and Normal tables.

2. **MODULE ASSESSMENT**

2.1 Continuous Assessment

2.1.1 Test 1 40% Assignment 1 10%

2.1.2 Test 2 40% Assignment 2 10%

2.2 Examination [**refer to School calendar**]

2.3 Assessment Dates will be decided later.

3. READING LIST

Although lecture notes are provided, it is important that you reinforce this material by referring to more detailed texts.

Recommended Text

"Probability and Statistics for Scientists and Engineers" by R.E Walpole & R.H Myers, any edition. (519.02462 PRO)

STA 1142/1542: Introductory Probability

Statistical (or random) Experiment – is a process or a course of action which has several possible outcomes. The outcome that occurs cannot be predicted before the action is performed.

Some examples:

- Toss a coin
- Roll a die
- Pick a card out of a regular pack of cards
- Write a test which has 10 questions
- Ask consumer if she prefers product A or B.
- The body mass of a new born baby.

Sample space – is the set of all possible outcomes of a statistical experiment. Each outcome in a sample space is called an element or a member of the sample space. The outcomes can either be discrete (then they can be listed individually) or continuous (then they fall within an interval and cannot be listed individually).

Discrete sets – we explicitly write down every single possible outcome. List them between curly brackets and separate them with semi colons.

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

Standard Deck of 52 Playing Cards

Clubs	Spades	Hearts	Diamonds
A♣	A♠	A♥	A♦
2♣	2♠	2♥	2♦
3♣	3♠	3♥	3♦
4♣	4♠	4♥	4♦
5♣	5♠	5♥	5♦
6♣	6♠	6♥	6♦
7♣	7♠	7♥	7♦
8♣	8♠	8♥	8♦
9♣	9♠	9♥	9♦
10♣	10♠	10♥	10♦
Jack♣	Jack♠	Jack♥	Jack♦
Queen♣	Queen♠	Queen♥	Queen♦
King♣	King♠	King♥	King♦

Continuous sets – if the outcome of the experiment can be any value within an interval, then the sample space is the interval. Call the results of the experiment x, and if x can assume any value in interval [a, b] (including a and b), then we write the sample space as follows:

S={x: a $\leq x \leq b$}

Null set (or empty set) – a null set, Φ, is a set which contains no elements.

Event – an event is a subset of a sample space.

Venn diagram – is a pictorial illustration of a sample space and events.

Complement – if A is an event with respect to the sample space S, the complement of A (denoted by A') is the subset of all elements that are not in A.

Intersection ($A \cap B$) – the intersection of two events, A and B, is the event containing all elements that are common to A and B.

Mutually exclusive events – two events are said to be mutually exclusive if they have no common elements.

Union ($A \cup B$ **)** – the union of two events, A and B, is the event containing all the elements that belong to A or B or both.

Probability of Events

1.1.Introduction

- ✓ Most people have some vague ideas about what probability of an event means. The interpretation of the word probability involves synonyms such as chance, odds, uncertainty, prevalence, risk, expectancy etc
- ✓ There are many applications of probability theory. We are studying probability theory because we would like to study mathematical statistics.
- ✓ Statistics is concerned with the development of methods and their applications for collecting, analyzing and interpreting quantitative data in such a way that the reliability of a conclusion based on data may be evaluated objectively by means of probability statements.
- ✓ Probability theory is used to evaluate the reliability of conclusions and inferences based on data. Thus, probability theory is fundamental to mathematical statistics. For an event
- ✓

A of a discrete sample space S, the probability of A can be computed by using the formula

$$P(A) = \frac{N(A)}{S(A)}$$

here $N(A)$ denotes the number of elements of A and $N(S)$ denotes the number of elements in the sample space S. For a discrete case, the probability of an event A can be computed by counting the number of elements in A and dividing it by the number of elements in the sample space S.

1.2. Counting Techniques

There are three basic counting techniques. They are multiplication rule, permutation and combination.

1.2.1 Multiplication Rule.

If $E1$ is an experiment with $n1$ outcomes and $E2$ is an experiment with n_2 possible outcomes, then the experiment which consists of performing $E1$ first and then $E2$ consists of $n_1 n_2$ possible outcomes.

Example 1.1. Find the possible number of outcomes in a sequence of two tosses of a fair coin. Answer: The number of possible outcomes is $2 \cdot 2 = 4$. This is evident from the following tree diagram. (We can also show this using a tree diagram)

Example 1.2. Find the number of possible outcomes of the rolling of a die and then tossing a coin. Answer: Here $n_1 = 6$ and $n_2 = 2$. Thus by multiplication rule, the number of possible outcomes is 12.

Example 1.3. How many different license plates are possible if Kentucky uses three letters followed by three digits.

1.2.2. Permutation

Consider a set of 4 objects. Suppose we want to fill 3 positions with objects selected from the above 4. Then the number of possible ordered arrangements is 24 and they are

a b c	b a c	c a b	d a b
a b d	b a d	c a d	d a c
a c b	b c a	c b a	d b c
a c d	b c d	c b d	d b a
a d c	b d a	c d b	d c a

a d b b d c c d a d c b

The number of possible ordered arrangements can be computed as follows: Since there are 3 positions and 4 objects, the first position can be filled in 4 different ways. Once the first position is filled the remaining 2 positions can be filled from the remaining 3 objects.

Thus, the second position can be filled in 3 ways. The third position can be filled in 2 ways. Then the total number of ways 3 positions can be filled out of 4 objects is given by $(4)(3)(2) = 24$

In general, if r positions are to be filled from n objects, then the total number of possible ways they can be filled are given by $n(n-1)(n-2)\cdots(n-r-1)$

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

The number of permutations [order is important] of n distinct objects taken r at a time is:

$$_nP_r = \frac{n!}{(n-r)!} \quad \text{, where} \, r = 0,1,2,\ldots\ldots,n.$$

Thus, nPr represents the number of ways r positions can be filled from n objects.

Definition 1.1. Each of the nPr arrangements is called a permutation of n objects taken r at a time. Example 1.4. How many permutations are there of all three of letters a, b, and c?

$$3P3 = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

Example: Four names are drawn from the 24 members of a club for the

1.2.3. Combination

In permutation, order is important. But in many problems, the order of selection is not important and interest canters only on the set of r objects.

The number of combinations [order is not important] of r objects selected from a set of n distinct objects is:

$$^{n}C_{r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}, \text{ where } r = 0,1,2,\ldots\ldots,n.$$

Example 1.7. How many committees of two chemists and one physicist can be formed from 4 chemists and 3 physicists?

1.3.Probability Measure

A random experiment is an experiment whose outcomes cannot be predicted with certainty. However, in most cases the collection of every possible outcome of a random experiment can be listed.

Definition A sample space of a random experiment is the collection of all possible outcomes.

Example: What is the sample space for an experiment in which we select a rat at random from a cage and determine its sex?

Answer: The sample space of this experiment is $S = \{M, F\}$ where M denotes the male rat and F denotes the female rat.

Example What is the sample space for an experiment in which we roll a pair of dice, one red and one green?

Answer: The sample space S for this experiment is given by

S = {(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)}

This set S can be written as $S = \{(x, y)|\ 1 \leq x \leq 6, 1 \leq y \leq 6\}$ where x represents the number rolled on red die and y denotes the number rolled on green die.

Definition 1.4. Each element of the sample space is called a sample point.

Definition 1.5. If the sample space consists of a countable number of sample points, then the sample space is said to be a countable sample space.

Definition 1.6. If a sample space contains an uncountable number of sample points, then it is called a continuous sample space. An event A is a subset of the sample space S. It seems obvious that if A and B are events in sample space S, then $A \cup B, A^c, A \cap B$ are also entitled to be events.

Example: Describe the sample space of rolling a die and interpret the event {1, 2}. Answer: The sample space of this experiment is S = {1, 2, 3, 4, 5, 6}. The event {1, 2} means getting either a 1 or a 2.

Example: First describe the sample space of rolling a pair of dice, then describe the event A that the sum of numbers rolled is 7.

Answer: The sample space of this experiment is $S = \{(x, y)|\ x, y = 1, 2, 3, 4, 5, 6\}$ and $A = \{(1, 6), (6, 1), (2, 5), (5, 2), (4, 3), (3, 4)\}$.

Definition 1.8. Let S be the sample space of a random experiment. A probability measure

$P : \mathcal{F} \to [0, 1]$ is a set function which assigns real numbers to the various events of S satisfying

$$P(A) \geq 0 for\ all\ events\ A$$

$$P(S) = 1$$

$$P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

if $A_1, A_2, A_3, \ldots, A_k$, ... are mutually disjoint events of S. Any set function with the above three properties is a probability measure for S. For a given sample space S, there may be more than one probability measure. The probability of an event A is the value of the probability measure at A, that is $Prob(A) = P(A)$.

Theorem

If $\emptyset$ is a empty set (that is an impossible event), then $P(\emptyset) = 0$.

1.4.Some Properties of the Probability Measure

Probability Reference List

The following properties hold for all events A, B.

- ✓ $P(\emptyset) = 0$.
- ✓ $0 \leq P(A) \leq 1$.
- ✓ Complement: $P(A^c) = 1 - P(A)$.
- ✓ Probability of a union: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

For three events A, B, C:

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

If A and B are mutually exclusive, then $P(A \cup B) = P(A) + P(B)$.

- ✓ Conditional probability: $P(A \mid B) = P(A \cap B)/P(B)$.
- ✓ Multiplication rule: $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$.
- ✓ The Partition Theorem: if $B_1, B_2, \ldots, B_m$ form a partition of Ω, then $P(A) = \sum_{i=1}^{m} P(A \cup B_i) = \sum_{i=1}^{m} P(A|B_i)P(B_i)$ for any event A.

  As a special case, B and B^c partition Ω, so:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$
$$= P(A \mid B)P(B) + P(A \mid B^c)P(B^c) \; for\ any\ A, B.$$

- ✓ Bayes' Theorem: $P(B \mid A) = \frac{P(A \mid B)P(B)}{P(A)}$. More generally, if $B_1, B_2, \ldots, B_m$ form a partition of Ω, then $P\left(B_j|A\right) = \frac{P\left(A|B_j\right)P(B_i)}{\sum_{i=1}^{m} P\left(A|B_j\right)P(B_i)}$ for any j
- ✓ Chains of events: for any events $A_1, A, \ldots, A_n$,

$$P(A1 \cap A2 \cap \ldots \cap An) = P(A1)P(A2 \mid A1)P(A3 \mid A2 \cap A1) \ldots P(An \mid An - 1 \cap \ldots \cap A1)$$

2. Conditional Probability and Bayes' Theorem

2.1. Conditional Probability

Suppose a bag contains 6 balls, 3 red and 3 white. Two balls are chosen (without replacement) at random, one after the other. Consider the two events R, W:

R is event "first ball chosen is red"

W is event "second ball chosen is white"

We easily find $P(R) = \frac{3}{6} = \frac{1}{2}$. However, determining the probability of W is not quite so straight forward. If the first ball chosen is red then the bag subsequently contains 2 red balls and 3 white. In this case $P(W) = \frac{3}{5}$. However, if the first ball chosen is white then the bag subsequently contains 3 red balls and 2 white. In this case $P(W) = \frac{2}{5}$. What this example shows is that the probability that W occurs is clearly dependent upon whether or not the event R has occurred. The probability of W occurring is conditional on the occurrence or otherwise of R. The conditional probability of an event B occurring given that event A has occurred is written $P(B|A)$. In this particular example

$$P(W|R) = \frac{3}{5} \text{ and } P(W|R') = \frac{2}{5}$$

Consider, more generally, the performance of an experiment in which the outcome is a member of an event A. We can therefore say that the event A has occurred. What is the probability that B then occurs? That is what is $P(B|A)$? In a sense we have a new sample space which is the event A. For B to occur some of its members must also be members of event A. So, for example, in an equi-probable space, $P(B|A)$ must be the number of outcomes in $A \cap B$ divided by the number of outcomes in A. That is

$$P(B|A) = \frac{number\ of\ outcomes\ in\ A \cap B}{number\ of\ outcomes\ in\ A}.$$

Now if we divide both the top and bottom of this fraction by the total number of outcomes of the experiment we obtain an expression for the conditional probability of B occurring given that A has occurred

Definition: Let S be a sample space associated with a random experiment. The conditional probability of an event A, given that event B has occurred, is defined by

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

provided $P(B) > 0$.

This conditional probability measure $P(A/B)$ satisfies all three axioms of a probability measure. That is,

- $P(A/B) \geq 0$ for all A
- $P(B/B) = 1$
- If $A1, A2, \ldots, Ak, \ldots$ are mutually exclusive events, then

$$P\left(\bigcup_{k=1}^{\infty} A_k \,|B\right) = \sum_{k=1}^{\infty} P(A_k|B)$$

Thus, it is a probability measure with respect to the new sample space B.

Examples:

1. A box contains six 10 Ω resistors and ten 30 Ω resistors. The resistors are all unmarked and are of the same physical size.

(a) One resistor is picked at random from the box; find the probability that:

(i) It is a 10 Ω resistor.

(ii) It is a 30 Ω resistor.

(b) At the start, two resistors are selected from the box.

Find the probability that:

(i) Both are 10 Ω resistors.

(ii) (ii) The first is a 10 Ω resistor and the second is a 30 Ω resistor.

(iii) (iii) Both are 30 Ω resistors.

2. A lecturer on a topic of public health is held and 300 people attended. They are classified in the following way:

Gender	Doctors	Nurses	Total
Female	90	90	180
Male	100	20	120
Total	190	110	300

If one person is selected at random, find the following probabilities:

(a) $P(\, a \, doctor \, is \, chosen)$;
(b) $P(\, a \, female \, is \, chosen)$;
(c) $P(\, a \, nurse \, is \, chosen)$;
(d) $P(\, a \, male \, is \, chosen)$;
(e) $P(\, a \, female \, nurse \, is \, chosen)$;
(f) $P(\, a \, male \, doctor \, is \, chosen)$

Independent events

If the occurrence of one event A does not affect, nor is affected by, the occurrence of another event B then we say that A and B are independent events. Clearly, if A and B are independent then

P(B|A) = P(B) and $P(A|B) \; = \; P(A)$

Then, using the Key Point formula $P(A \; \cup \; B) \; = P(B|A)P(A)$ we have, for independent events:

If A and B are independent events, then

$$P(A \cap B) = P(A)P{\times}(B)$$

In words

“The probability of events A and B occurring is the product of the probabilities of the events occurring separately”

Theorem

Let $A, B \subseteq S$ if A and B are independent then

$$P(A|B) = P(A)$$

Proof: $PA|B) = \frac{P(A \cap B)}{p(B)}$

$$= \frac{P(A).P(B)}{P(B)}$$

$$= P(A)$$

Theorem

If A and B are independent events. Then A^c and B are independent. Similarly A and $B^c$are independent.

Proof: We know that A and B are independent, that is

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

We want to show that A^c and B are independent, that is

$$P(A^c \cap B) = P(A^c) \times P(B)$$

Since
$$P(A^c \cap B) = P(A^c|B) \times P(B)$$

$$= [1 - P(A|B)]P(B)$$

$$= P(B) - P(A|B)P(B)$$

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

$$= P(B) - P(A)P(B)$$

$$= P(B)[1 - P(A)]$$

$$= P(B)P(A^c),$$

the events A^c and B are independent. Similarly, it can be shown that A and B^c are independent and the proof is now complete.

Remark. The concept of independence is fundamental. In fact, it is this concept that justifies the mathematical development of probability as a separate discipline from measure theory. Mark Kac said, “independence of events is not a purely mathematical concept.”

It can, however, be made plausible that it should be interpreted by the rule of multiplication of probabilities and this leads to the mathematical definition of independence.

Examples

1. Flip a coin and then independently cast a die. What is the probability of observing heads on the coin and a 2 or 3 on the die?
2. Describe which of the following pairs (A and B) of events arising from the experiments described are independent.
   (a) One card is drawn from each of two packs A = {a red card is drawn from pack 1} B = {a picture card is drawn from pack 2}
   (b) The daily traffic accidents in Hull involving pedal cyclists and motor cyclists are counted A = {three motor cyclists are injured in separate collisions with cars} B = {one pedal cyclist is injured when hit by a bus}
   (c) Two boxes contains 20 nuts each, some have a metric thread, some have a British Standard Fine (BSF) threads and some have a British Standard Whitworth (BSW) thread. A nut is picked out of each box. A = {nut picked out of the first box is BSF} B = {nut picked out of the second box is metric }
   (d) A box contains 20 nuts, some have a metric thread, some have a British Standard Fine (BSF) threads and some have British Standard Whitworth (BSW) thread. Two nuts are picked out of the box. A = {first nut picked out of the box is BSF} B = {second nut picked out of the box is metric}
3. A box contains three white cards and 3 black cards numbered as follows:

   white

   1 2 2

   Black

   1 1 2

One card is picked out of the box at random. If A is the event 'the card is black' and B is the event 'the card is marked 2', are A and B independent?

4. What is the probability of obtaining "six" and "six" on two successive rolls of a die?
5. A couple decides has two children. Let A be the event 'they have one boy and one girl' and B the event that 'they have at most one boy'. Are A and B independent?

SUMMARIES

Laws of Elementary Probability

Let a sample space *S* consist of the n simple distinct events $E_1, E_2, E_3 \ldots En$ and let A and B be events contained in *S*.

Then:

- $0 \leq P(A) \leq 1. P(A) = 0$ is interpreted as meaning that the event A cannot occur and $P(A) = 1$ is interpreted as meaning that the event A is certain to occur.
- $P(A) + P(A')= 1$ where the event A' is the complement of the event A
- $P(E_1) + P(E_2) + \cdots + P(E_n) = 1$ where $E_1, E_2, E_3 \ldots En$ form the sample space
- If A and B are any two events, then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- If A and B are two mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$
- If A and B are two independent events, then $P(A \cap B) = P(A) \times P(B)$.

2.2. Bayes' Theorem

There are many situations where the ultimate outcome of an experiment depends on what happens in various intermediate stages. This issue is resolved by the Bayes' Theorem.

In statistics and probability theory, the Bayes' theorem (also known as the Bayes' rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes' theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.

The theorem is named after English statistician, Thomas Bayes, who discovered the formula in 1763. It is considered the foundation of the special statistical inference approach called the Bayes' inference.

Besides statistics, the Bayes' theorem is also used in various disciplines, with medicine and pharmacology as the most notable examples. In addition, the theorem is commonly employed in different fields of finance. Some of the applications include but are not limited to, modeling the risk of lending money to borrowers or forecasting the probability of the success of an investment.

The theorem of Total probability (or Rule of Elimination):

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that P(Bi)≠0 for

i = 1, 2,…., k, then for any event A of S,

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)\,(A|B_i)$$

Bayes' Rule

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(A) \neq 0$,

$$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(Bi \cap A)} = \frac{P(B_r)P(A/B_r)}{\sum_{i=1}^{k} P(B_i)P(A/B_i)}$$

Examples:

1. Two boxes containing marbles are placed on a table. The boxes are labelled B_1 and B_2. Box B_1 contains 7 green marbles and 4 white. Box B_2 contains 3 green marbles and 10 yellow marbles. The boxes are arranged so that the probability of selecting box B_1 is 1/3 and the probability of selecting box B_2 is 2/3. Kathy is blindfolded and asked to select a marble. She will win a color TV if she selects a green marble.

(a) What is the probability that Kathy will win the TV (that is, she will select a green marble)?

(b) If Kathy wins the color TV, what is the probability that the green marble was selected from the first box?

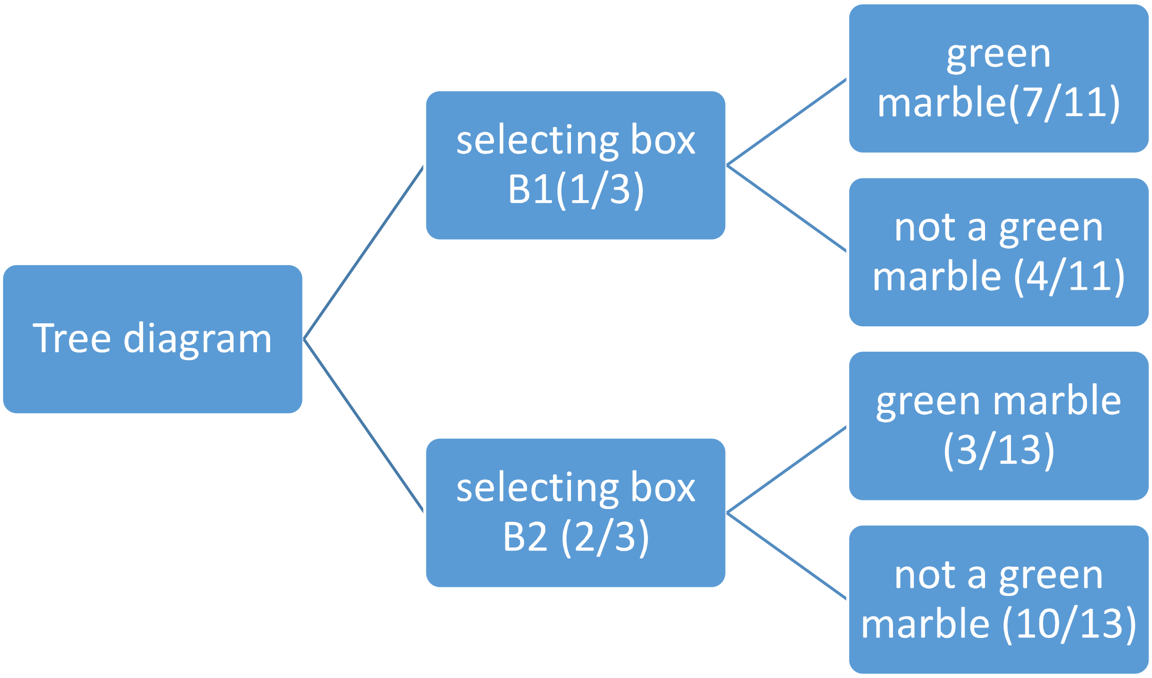
Answer: Let A be the event of drawing a green marble. The prior probabilities are P(B1) = 1/3 and P(B2) = 2/3 .

(a) The probability that Kathy will win the TV is

$$P(A) = P(A \cap B_1) + P(A \cap B_2)$$
$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2)$$
$$= \frac{7}{11} \times \frac{1}{3} + \frac{3}{13} \times \frac{2}{3}$$
$$= \frac{7}{33} + \frac{6}{39}$$
$$=$$

(b) Given that Kathy won the TV, the probability that the green marble was selected from B_1 is

Draw a tree diagram



$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)}$$

$$= \frac{\left(\frac{7}{11}\right)\left(\frac{1}{3}\right)}{\left(\frac{7}{11}\right)\left(\frac{1}{3}\right)+\left(\frac{3}{13}\right)\left(\frac{2}{3}\right)}$$

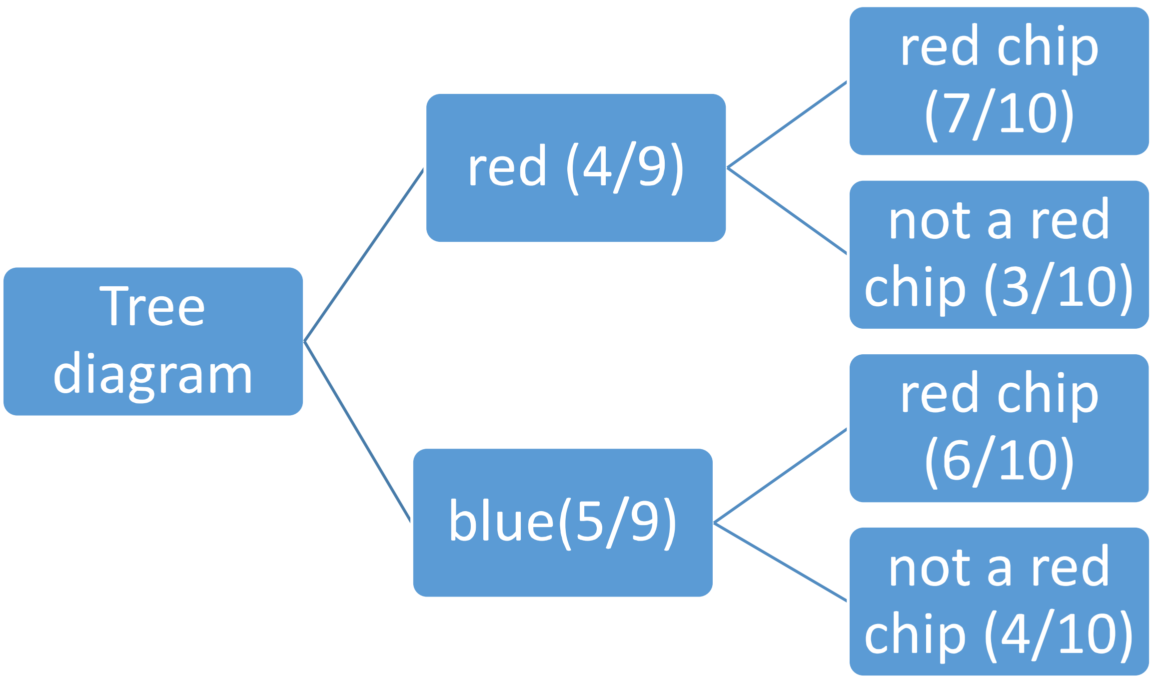
$$= \frac{91}{157}$$

Note that *$P(A/B_1)$* is the probability of selecting a green marble from B1 whereas *$P(B_1/A)$* is the probability that the green marble was selected from box B1.

(B1/A) = P(A/B1) P(B1) P(A/B1) P(B1) + P(A/B2) P(B2)

2. Suppose box A contains 4 red and 5 blue chips and box B contains 6 red and 3 blue chips. A chip is chosen at random from the box A and placed in box B. Finally, a chip is chosen at random from among those now in box B. What is the probability a blue chip was transferred from box A to box B given that the chip chosen from box B is red?

Answer: Tree diagram: Box A represented by first branches and box B represented by next branches



Let E represent the event of moving a blue chip from box A to box B. We want to find the probability of a blue chip which was moved from box A to box B given that the chip chosen from B was red. The probability of choosing a red chip from box A is P(R) = 4/9 and the probability of choosing a blue chip from box A is P(B) = 5/9 . If a red chip was moved from box

A to box B, then box B has 7 red chips and 3 blue chips. Thus the probability of choosing a red chip from box B is 7/10 . Similarly, if a blue chip was moved from box A to box B, then the probability of choosing a red chip from box B is 6/10 .

Hence, the probability that a blue chip was transferred from box A to box B given that the chip chosen from box B is red is given by

$$P(E|R) = \frac{P(R|E)P(E)}{P(R)}$$

$$= \frac{(\frac{6}{10})(\frac{5}{9})}{\left(\frac{7}{10}\right)\left(\frac{4}{9}\right)+(\frac{6}{10})(\frac{5}{9})}$$

$$= \frac{15}{29}$$

3. Sixty percent of new drivers have had driver education. During their first year, new drivers without driver education have probability 0.08 of having an accident, but new drivers with driver education have only a 0.05 probability of an accident. What is the probability a new driver has had driver education, given that the driver has had no accident the first year?

Answer: Let A represent the new driver who has had driver education and B represent the new driver who has had an accident in his first year. Let A^c and B^c be the complement of A and B, respectively. We want to find the probability that a new driver has had driver education, given that the driver has had no accidents in the first year, that is $P(A/B^c)$.

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)}$$

$$= \frac{P(B^c|A)P(A)}{P(B^c|A)P(A)+P(B^c|A^c)P(A^c)}$$

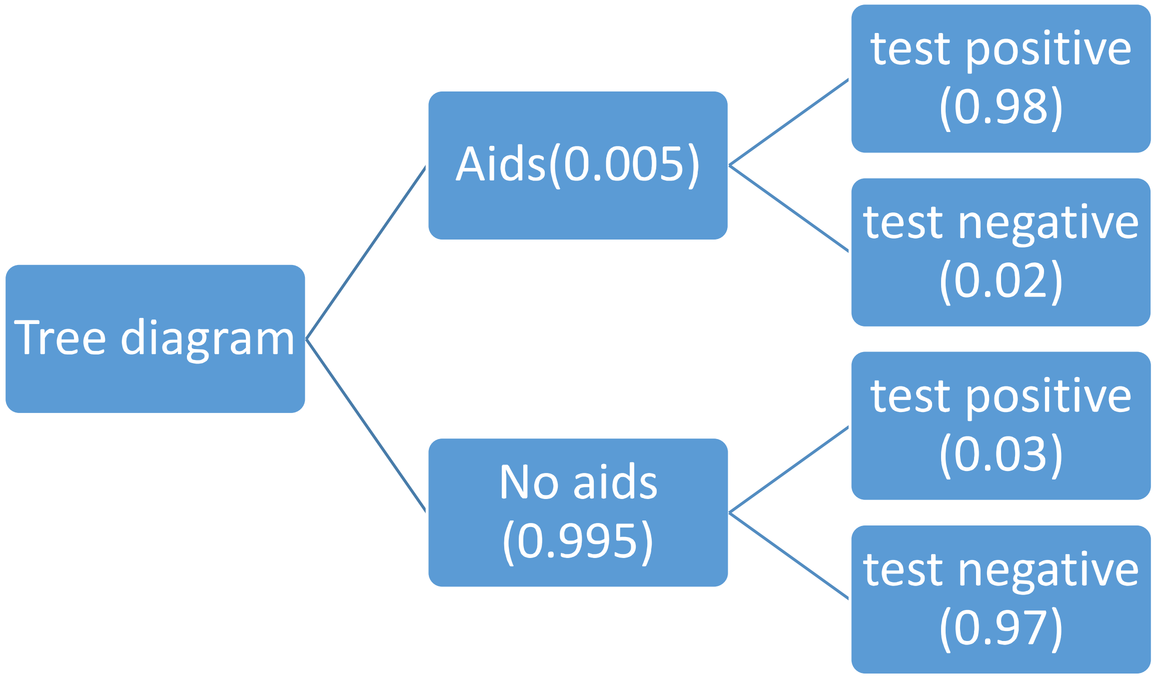
$$= \frac{[1-P(B|A)]P(A)}{[1-P(B|A)]P(A)+[1-P(B|A^c)][1-P(A)]}$$

$$= \frac{\left(\frac{60}{100}\right)\left(\frac{95}{100}\right)}{\left(\frac{40}{100}\right)\left(\frac{92}{100}\right)+\left(\frac{60}{100}\right)\left(\frac{95}{100}\right)}$$

$$= 0.6077$$

4. One-half percent of the population has AIDS. There is a test to detect AIDS. A positive test result is supposed to mean that you have AIDS but the test is not perfect. For people with AIDS, the test misses the diagnosis 2% of the times. And for the people without AIDS, the test incorrectly tells 3% of them that they have AIDS.
   (a) What is the probability that a person picked at random will test positive?
   (b) What is the probability that you have AIDS given that your test comes back positive?

Answer: Let A denote the event of one who has AIDS and let B denote the event that the test comes out positive.



(a) The probability that a person picked at random will test positive is given by

$$P(test\ positive) = (0.005)(0.98) + (0.995)(0.03)$$
$$= 0.0049 + 0.0298$$
$$= 0.035.$$

(b) The probability that you have AIDS given that your test comes back positive is given by

$$P(A|B) = \frac{favourable\ positive\ branches}{total\ positive\ branches}$$

$$= \frac{(0.005)(0.98)}{(0.005)(0.98)+(0.995)(0.003)}$$

$$= \frac{0,0049}{0.035}$$

$$= 0.14$$

Remark. This example illustrates why Bayes' theorem is so important. What we would really like to know in this situation is a first-stage result: Do you have AIDS? But we cannot get this information without an autopsy. The first stage is hidden. But the second stage is not hidden. The best we can do is make a prediction about the first stage. This illustrates why backward conditional probabilities are so useful.

3. Random Variables and Distribution Functions

3.1. Introduction

In many random experiments, the elements of sample space are not necessarily numbers. For example, in a coin tossing experiment the sample space consists of

$$S = \{Head, Tail\}.$$

Statistical methods involve primarily numerical data. Hence, one has to 'mathematize' the outcomes of the sample space. This mathematization, or quantification, is achieved through the notion of random variables.

Definition 1:

- A random variable is a quantity which can take on the values of a given set (sample space) with specified probabilities.
- A random variable is the aspect of a random experiment which we are interested in studying.

Definition 2:

A random variable is said to be discrete if it can only assume a countable number of probabilities.

3.2 Distribution Functions of Discrete Variables

Every random variable is characterized through its probability density function.

Definition 3.4. Let R_X be the space of the random variable X. The function $f : R_X \to \mathcal{R}$ defined by

$$f(x) = P(X = x)$$

is called the probability density function (pdf) of X.

Example: In an introductory statistics class of 50 students, there are 11 freshmen, 19 sophomores, 14 juniors and 6 seniors. One student is selected at random. What is the sample space of this experiment? Construct a random variable X for this sample space and then find its space. Further, find the probability density function of this random variable X.

Answer: The sample space of this random experiment is $S = \{Fr, So, Jr, Sr\}$.

Define a function X : S as follows: $X(Fr) = 1, X(So) = 2\ X(Jr) = 3, X(Sr) = 4$.

Then clearly X is a random variable in S.

The space of X is given by $R_X = \{1, 2, 3, 4\}$.

The probability density function of X is given by

$$f(1) = P(X = 1) = \frac{11}{50}$$

$$f(2) = P(X = 2) = \frac{19}{50}$$

$$f(3) = P(X = 3) = \frac{14}{50}$$

$$f(4) = P(X = 4) = \frac{6}{50}.$$

Example: A pair of dice consisting of a six-sided die and a four-sided die is rolled and the sum is determined. Let the random variable X denote this sum. Find the sample space, the space of the random variable, and probability density function of X.

Answer: The sample space of this random experiment is given by

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

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

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

$$(4,1)\ (4,2)\ (4,3)\ (4,4)\ (4,5)\ (4,6)\}$$

The space of the random variable X is given by $R_X = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$.

Therefore, the probability density function of X is given by

$$f(2) = P(X = 2) = \frac{1}{24}, \quad f(3) = P(X = 3) = \frac{2}{24}$$

$$f(4) = P(X = 4) = \frac{3}{24}, \qquad f(5) = P(X = 5) = \frac{4}{24}$$

$$f(6) = P(X = 6) = \frac{4}{24}, \qquad f(7) = P(X = 7) = \frac{4}{24}$$

$$f(8) = P(X = 8) = \frac{3}{24}, \qquad f(9) = P(X = 9) = \frac{2}{24}$$

$$f(10) = P(X = 10) = \frac{1}{24}$$

Example: A fair coin is tossed 3 times. Let the random variable X denote the number of heads in 3 tosses of the coin. Find the sample space, the space of the random variable, and the probability density function of X.

Answer: The sample space S of this experiment consists of all binary sequences of length 3, that is

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

The space of this random variable is given by $R_X = \{0, 1, 2, 3\}$. Therefore, the probability density function of X is given by

$$f(0) = P(X = 0) = \frac{1}{8}, \qquad f(1) = P(X = 1) = \frac{3}{8}$$

$$f(2) = P(X = 2) = \frac{3}{8}, \qquad f(3) = P(X = 3) = \frac{1}{8}.$$

This can be written as follows: $f(x) = \binom{3}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x}, x = 0, 1, 2, 3$

The probability density function f(x) of a random variable X completely characterizes it. Some basic properties of a discrete probability density function are summarized below.

Theorem

If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each x within the range of X is called the probability distribution of X with the following properties:

1. $f(x) \geq 0$ for each value within its domain.

2. $\sum_{x} f(x) = 1$, where the summation extends over all the values within its domain.

Example: If the probability of a random variable X with space $R_X = \{1, 2, 3, \ldots, 12\}$ is given by

$$f(x) = k(2x - 1)$$

then, what is the value of the constant k?

Definition

The cumulative distribution, F(x), of a discrete random variable, X, with probability function f(x) is:

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty$$

The values, F(x), of the cumulative function of a discrete random variable, X, satisfy the following conditions:

1. $F(-\infty) = 0$
2. $F(\infty) = 1$
3. If $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b.

Example: If the probability density function of the random variable X is given by

$$f(x) = \frac{1}{144}(2x - 1) \; for \; x = 1, 2, 3, \ldots, 12$$

then find the cumulative distribution function of X.

Theorem: Let X be a random variable with cumulative distribution function F(x). Then the cumulative distribution function satisfies the followings:

(a) $F(-\infty) = 0$
(b) $F(\infty) = 1$
(c) $F(x)$ is an increasing function, that is if $x < y$, then $F(x) \leq F(y)$ for all reals x, y

Theorem

If the space R_X of the random variable X is given by $R_X = \{x_1 < x_2 < x_3 < \cdots < x_n\}$, then

$$f(x_1) = F(x_1)$$

$$f(x_2) = F(x_2) - F(x_1)$$

$$f(x_3) = F(x_3) - F(x_2)$$

.. ..

.. ..

$$f(x_n) = F(x_n) - F(x_{n-1})$$

Example: Find the probability density function of the random variable X whose cumulative distribution function is

$$f(x) = \begin{cases} 0, & x < -1 \\ 0.25, & -1 \leq x < 1 \\ 0.50, & 1 \leq x < 3 \\ 0.75, & 3 \leq x < 5 \\ 1, & x \geq 5 \end{cases}$$

Also, find (a) $P(X \leq 3)$, (b) $P(X = 3)$, and (c) $P(X < 3)$.

3.3 Distribution Functions of Continuous Variables

A random variable X is said to be continuous if its space is either an interval or a union of intervals. The following definition formally defines a continuous random variable.

Definition

A random variable is said to be continuous if it can assume an uncountable number of values.

A function, f(x), defined over the set of all real numbers, is called a probability density function (pdf) of the continuous random variable, X, if and only if

$$p(a \le x \le b) = \int_a^b f(x)dx \qquad \text{for any real constants a and b with a} \le \text{b.}$$

The pdf satisfies the following conditions:

1. $f(x) \ge 0 \, for \, -\infty < x < \infty$
2. $\int_a^b f(x) = 1$

Example: Let X be a random variable with pdf

$$f(x) = \begin{cases} 2x^{-2}, & if \ 1 < x < 2 \\ 0 & otherwise \end{cases}$$

Is $f(x)$ a probability density function?

Example: For what value of the constant c, where a, b are real constants is $f(x)$ a probability density function?

$$f(x) = \begin{cases} c, & if \ a < x < b \\ 0 & otherwise \end{cases}$$

Exercises:

i. If the random variable X possesses the density function given by

$$f(x) = \begin{cases} cx, & if \ 0 < x < 2 \\ 0 & otherwise \end{cases}$$

then what is the value of c for which $f(x)$ is a probability density function? What is the cumulative distribution function of X.

ii. The length of time required by students to complete a 1-hour exam is a random variable with a pdf given by

$$f(x) = \begin{cases} cx^2 + x, & if \ 0 < x < 1 \\ 0 & otherwise \end{cases}$$

Find the value of c. Is X a probability density function?

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

$$f(x) = \begin{cases} \frac{x^2}{3}, & if -1 < x < 2 \\ 0 & otherwise \end{cases}$$

a. Verify that X is a probability distribution function.
b. Find P(0 < X ≤ 1).
c. Find F(x), and use it to evaluate P(0 < X ≤ 1).

Theorem:

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

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

Definition 6:

If X is a continuous random variable, the function given by

$$F(x) = P(X \leq x) = \int_{-\infty}^{\infty} f(t)dt \quad \text{for } -\infty < x < \infty$$

Is called the cumulative distribution of X. [$f(t)$ is the value of the probability density function of X at t]

Theorem:

If $f(x)$ is the pdf and $F(x)$, the cumulative distribution of X, then

$P\ (a \leq x \leq b) = F(b) - F(a)$, for any real constants a and b with $a \leq b$, and

$f(x) = \dfrac{dF(x)}{dx}$, where the derivative exists.

Example

1. If the random variable X has the following probability density function

$$f(x) = \begin{cases} 3x^2, 0 < x < 1 \\ 0, \quad elsewhere. \end{cases}$$

What is the CDF of X?

For 0 < x < 1

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} 3t^2 dt = x^3$$

The CDF is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ x^3, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

2. If a continuous random variable X has the pdf given by

$$f(x) = \begin{cases} \frac{x^3}{4}, 0 < x < 2 \\ 0, \quad elsewhere. \end{cases}$$

a) What is the cumulative distribution function of X?

b) Evaluate $P(1 < x \leq 2)$

Solution

a) For 0< x < 2

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \frac{t^3}{4} dt = \frac{t^4}{16} | = \frac{x^4}{16}.$$

The CDF is given by

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^4}{16}, & 0 < x < 2 \\ 1, & x \geq 2 \end{cases}$$

b) $P(1 < x \leq 2) = F(2) - F(1) = \frac{16}{16} - \frac{1}{16} = 15/16$

3. A continuous random variable X has a pdf $f(x)$, defined by

$$f(x) = \begin{cases} \frac{1}{4}, 0 \leq x < 1 \\ \frac{x^3}{5}, \quad 1 \leq x < 2 \\ 0, \quad elsewhere \end{cases}$$

Find the cumulative distribution function, $F(x)$.

Solution

For $0 \leq x < 1$

$F(x) = \int_0^x \frac{1}{4} dt = \frac{1}{4} t| = \frac{1}{4} x$, (by substituting the limits of integration)

For $1 \leq x < 2$

$F(x) = (\int_1^x \frac{t^3}{5} dt) + 1/4 = \frac{t^4}{20}| + \frac{1}{4} = \frac{x^4+4}{20}$, (by substituting the limits of integration).

The CDF is given by

$$F(x) = \begin{cases} 0, x < 0 \\ \frac{1}{4}x, 0 \leq x < 1 \\ \frac{x^4+4}{20}, 1 \leq x < 2 \\ 1, \quad x \geq 2 \end{cases}$$

4. MATHEMATICAL EXPECTATION

Definition 1:

Let X be a random variable with probability distribution *f(x).* The mean or expected value of X is defined as follows:

$$\mu = E(X) = \sum_x xf(x) \quad \text{if X is discrete;}$$

$$\mu = \text{E(X)} = \int_{-\infty}^{\infty} xf(x)dx \quad \text{if X is continuous.}$$

Definition 2:

Let X be a random variable with probability distribution *f(x)* and mean �. The variance of X is defined as follows:

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) \quad \text{if X is discrete;}$$

$$\sigma^2 = E[(X-\mu)^2] = \int (x-\mu)^2 f(x)dx \quad \text{if X is continuous.}$$

The positive square root of the variance, s, is called the standard deviation of X.

Skewness of data

The distribution of data can be symmetrical, have positive skew or negative skew. The coefficient of skewness of a random variable X can be calculated as;

$$\beta_1 = \frac{\mu_3}{\sigma^3}$$

μ_3=$\sum_{x\in A}(x-\mu)^3 P(X=x)$, called the central moment of X, σ is the standard deviation.

If $\beta_1 = 0$, the distribution is symmetrical

If $\beta_1 > 0$, distribution is positively skewed and

If $\beta_1 < 0$, distribution is negatively skewed. The larger the value of β_1, the greater the skewness of the distribution.

Examples (discrete case)

1. Let X be the random variable denoting the number of spots that appear when a fair dice is rolled once. Find the mean, variance and standard deviation of the number of X.

Solution

The sample space, S = {1, 2, 3, 4, 5, 6} and each outcome has a probability of 1/6.

The probability distribution is shown below:

x	1	2	3	4	5	6
$f(x)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

The mean is computed as follows:

$$\mu = E(X)$$

$$= \sum_x xf(x)$$

= 1(1/6) +2(1/6) + 3(1/6) +4(1/6) +5(1/6) +6(1/6) = 3.5

The variance is computed as follows:

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

$$\sigma^2 = \sum_x x^2 f(x) - \mu^2$$

$= (1^2.1/6 + 2^2.1/6 + 3^2.1/6 + 4^2.1/6 + 5^2.1/6 + 6^2.1/6) - (3.5)^2$

=2.9

The standard deviation, $\sigma = \sqrt{2.9}$ = 1.7

Check the skewness of the distribution and comment.

2. The probabilities for the discrete random X is defined by:

   $f(x) = \frac{x^2-1}{n}, \quad for\ x = \ 2, 3, 4, 5$

   a) Determine n
   b) Construct the probability distribution of X
   c) Calculate the mean value of X
   d) Calculate the standard deviation of X
   e) Would you say that the distribution is symmetrical? If not, comment on the distribution.

Solution

a) By definition,

$$\sum_{x \in A} f_X\ (x) = 1$$

$$\Rightarrow\ P(2) + P(3) + P(4) + P(5) = 1$$

$$\Rightarrow\ \frac{2^1-1}{n} + \frac{3^2-1}{n} + \frac{4^1-1}{n} + \frac{5^1-1}{n} = 1$$

$$\Rightarrow\ \frac{3}{n} + \frac{8}{n} + \frac{15}{n} + \frac{24}{n} = 1,$$

$\Rightarrow$ n=50

b) The probability distribution of X is

X	2	3	4	5
P(x)	3/50	4/25	3/10	12/25

c) The mean value of X is calculated as:

   $\mu = E(X) = \sum_x xf(x)$

   = 2(3/50) +3(4/25) + 4(3/10) +5(12/25)

   = 4.2

d) The variance of X is calculated as:

   $\sigma^2 = E(X^2) - \mu^2 = \sum_x x^2\ f(x) - \mu^2$

   = $\{2^2(3/10) + 3^2(4/25) + 4^2(3/10) + 5(12/25)\} - 4.2^2$

   = 0.84

Standard deviation, $\sigma = \sqrt{0.84}$ = 0.9165

e) The coefficient of skewness of X is calculated as;

$$\beta_1 = \frac{\mu_3}{\sigma^3}$$

Now μ_3= the third central moment of X is computed as;

$$\mu_3 = \sum_{x \in A}(x-\mu)^3 P(X = x)$$

= $(2-4.2)^3(3/50) + (3-4.2)^3(4/25) + (4-4.2)^3(3/10) + (5-4.2)^3(12/25)$

= -0.672

Thus,

$$\beta_1 = \frac{\mu_3}{\sigma^3}$$

$$= \frac{-0.672}{(0.9165)^3}$$

$$= -0.8729$$

Since $\beta_1 < 0$, the data is negatively skewed,

Thus, the distribution is not symmetrical.

3. One thousand tickets are sold at R1 each for a colour television valued at R350. What is the expected value of the gain if a person purchases one ticket?

Solution: The problem can be set up as follows:

	win	lose
Gain x	R349	-R1
P(x)	1/1000	999/1000

Two things should be noted. First, for a win, the net gain is R349, since the person does not get the cost of the ticket (R1) back. Second, for a loss, the gain is represented by a negative number, in this case -R1. The solution, then is

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

= \$349.1/1000 + (-\$1). 999/1000

= -\$0.65

COMMENT

Note that the expectation is -R0.65. This does not mean that a person loses R0.65, since the person can only win a television set valued at R345 or lose R1 on the ticket.

What this expectation means is that the average of the losses is R0.65 for each of the 1000 tickets holders. Here is another way of looking at this situation. If a person purchased one ticket each week over a long period of time, the average loss would be R0.65 per ticket, since theoretically, on average, that person would win the set once for each 1000 tickets purchased.

Examples (continuous case)

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

$$f(x) = \begin{cases} \frac{x^2}{3}, -1 < x < 2 \\ 0, \quad elsewhere. \end{cases}$$

   a) Find the mean and variance of X
   b) Check the skewness of the distribution,

2. The Department of Energy (DOE) puts projects out on bid and generally estimates what a reasonable bid should be. Call the estimate b. The DOE has determined that the density function of the winning (low) bid is

$$f(y) = \begin{cases} \frac{5}{8b}, \ \frac{2b}{5} \leq y \leq 2b \\ 0, \quad elsewhere. \end{cases}$$

   a) Find the mean and variance of X
   b) Check the skewness of the distribution,

4. Binomial, Poisson and Normal distributions, Binomial and Normal tables

The Binomial Distribution

- The random experiment consists of *n* independent repeated Bernoulli trials.
- The probability of "success", p, remains constant from trial to trial
- The Binomial random variable, X = the number of "successes" out of the *n* trials.

The Binomial distribution is given by:

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \mathrm{x} = 0, 1, 2, \ldots\ldots, \mathrm{n}$$

n and *p* are called the parameters of the distribution, $\mu = np$ and $\sigma^2 = np(1-p)$

Examples:

1. A survey found that one out of five South Africans say he or she has visited a doctor in any given month. If 10 people are selected at random, find the probability that exactly 3 will have visited a doctor last month.

Solution

In this case, $n = 10, X = 3, p = 1/5 \; and \; q = 4/5$.

Hence, $P(X = 3) = \frac{10!}{(10-3)!3!}(1/5)^3(4/5)^{10-3} = 0{,}201$

Alternatively using the binomial table

$$b(x, n, p) = b(3, 10, 0{,}2) = 0{,}201$$

2. A survey from Teenage Research Unlimited found that 30% of teenage consumers receive their spending money from part-time jobs. If 5 teenagers are selected at random, find the probability that at least 3 of them will have part-time jobs.

Solution

To find the probability that at least 3 have part-time jobs, it is necessary to find the individual probabilities for either 3, or 4 or 5, and then add them to get the total probability.

$$n = 5, p = 0{,}3, q = 0{,}7$$

$$P(X = 3) = \frac{5!}{(5-3)!\,3!}(0{,}3)^3(0{,}7)^{5-3} = 0{,}132$$

$$P(X = 4) = \frac{5!}{(5-4)!\,4!}(0{,}3)^4(0{,}7)^{5-4} = 0{,}028$$

$$P(X = 5) = \frac{5!}{(5-5)!\,5!}(0{,}3)^5(0{,}7)^{5-5} = 0{,}002$$

$$P(at\ least\ three\ teenagers\ have\ part-time\ jobs) = 0{,}132 + 0{,}028 + 0{,}002$$
$$= 0{,}162$$

Alternatively, using the binomial table

P(X≥ 3) = 1-P(X < 3)

= 1 – [P(X= 0) + P(X = 1) + P(X = 2)]

= 1-[b(0, 5, 0,3) + b(1, 5, 0,3) + b(2, 5, 0,3)]

= 1-(0,168 +0,360 + 0,309)

= 1- 0,837

= 0,163

3. The probability that a certain kind of component will survive a shock test is 3/5.

Find the probability that exactly 2 of the 4 components tested survive.

Solution

Assuming that the test are independent and p = 3/5, q = 1-3/5 =2/5, x = 2, n = 4.

We obtain

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

$$P(X = 2) = \frac{4!}{(4-2)!2!}\,(3/5)^2(2/5)^{4-2} = 216/625$$

Or b(2, 4, 3/5) = $\binom{4}{2}(3/5)^2(2/5)^{4-2}$ = 216/625

Or using the tables, $b(x, n, p) = b(2, 4, 0.6) = 0{,}346$.

The Poisson Distribution

Poisson experiments yield numerical values of a random variable X representing the number of outcomes occurring during a given time interval or in a specified region.

The Poisson distribution is given by:

$$p(x) = \frac{e^{\lambda}\lambda^{x}}{x!}, \quad \mathrm{x} = 0, 1, 2, \ldots\ldots$$

Where λ is the average number of outcomes per unit time or region.

$\mu = \lambda$ and $\sigma^2 = \lambda$

Examples

1. If there are 200 typographical errors randomly distributed in a 500-page manuscript, find the probability that a given page contains exactly three errors.

Solution

First, find the mean number of errors. Since there are 200 errors distributed over 500 pages, each page has an average of

$\lambda = \frac{200}{500} = \frac{2}{5} = 0{,}4$ error per page, X = 3

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

$$P(3; 0{,}4\ ) = \frac{(2{,}7183)^{-0{,}4}.\ (0{,}4)^{3}}{3!} = 0.0072$$

Thus, there is less than 1% probability that any given page will contain exactly three errors.

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

Solution

$\lambda = 4$, X = 6

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

$$P(6,4) = \frac{(2{,}7183)^{-4}.\ (4)^{6}}{6!}$$

= 0,1042

From the Poisson table, where X = 6 and $\lambda = 4$, $P(6,4) = 0{,}1042$.

The Normal Distribution

The normal probability distribution or the normal curve is a bell-shaped (symmetric) curve. Its mean is denoted by μ and its standard deviation by σ. A continuous random variable *x* that has a normal distribution is called a normal random variable. Not all bell-shaped curves represent a normal distribution curve.

Properties of Normal distribution.

1. The area under the normal curve is equal to 1.
2. The curve is symmetric about the mean.
3. Mean, median and mode of a normal distribution are equal.
4. Normal distributions are defined by two parameters, the mean (μ) and the standard deviation (σ).

The standard Normal distribution

The standard normal distribution is a special case of the normal distribution. For the standard normal distribution, the value of the mean is equal to zero, and the value of the standard deviation is equal to one.

The normal distribution with μ = 0 and σ = 1 is called the standard distribution. The random variable that possesses the standard normal distribution is denoted by *z*. the units marked on the horizontal axis of the standard normal curve are denoted by *z* and are called the *z* values or *z* scores. A specific value of *z* gives the distance between the mean and the point represented by *z* in terms of the standard deviation.

The z values on the right side of the mean are positive and those on the left side of the mean are negative. Although the values of z on the left side of the mean are negative, the area under the curve is always positive. The area under the standard curve between any two points can be interpreted as the probability that z assumes a value within that interval.

The normal distribution table gives the area to the left of a z value. To find the area to the right of z, first we find the area to the left of z, then we subtract this area from 1 which is the total area under the curve.

Examples:

a. Find the area under the standard normal curve to the left of z = 1.96.
b. Find the area under the standard normal curve from z = -2.15 to z = 0.
c. Find the following areas under the standard normal curve.
   i. Area to the right of z = 2.38
   ii. Area to the left of z = -1.57

Standardizing a Normal Distribution

In real world applications, a random variable may have a normal distribution with values of the mean and standard deviation that are different from 0 and 1, respectively. The first step in such a case is to convert the given normal distribution to the standard normal distribution. This procedure is called standardizing a normal distribution. The units of the normal distribution are denoted by x.

Converting an x value to a z value.

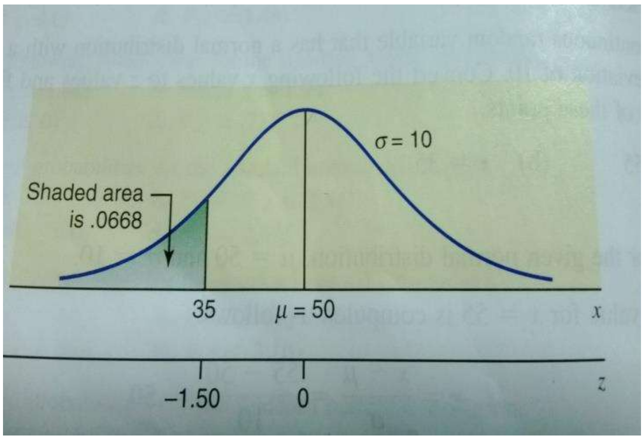
For a normal random variable x, a particular value of x can be converted to its corresponding z value by the formula:

$$z = \frac{x - \mu}{\sigma}$$

Where μ and σ are the mean and standard deviation of the normal distribution of x, respectively. Thus to find the z value for an x value, we calculate the difference between the given x value and the mean μ, and divide this by the standard deviation σ. If the value of x is equal to μ, then its z value is equal to zero. Note that we will always round z values to two decimal places.

Example

Let x be a continuous random variable that has a normal distribution with a mean of 50 and a standard deviation of 10. Convert the x value of 35 to z value and find the probability to the left of this point.



The z value for an x value that is greater than m is positive, the z value for an x value that is equal to μ is zero, and the z value that is less than μ is negative.

To find the area between two values of x for a normal distribution, we first convert both values of x to their respective z values. Then we find the area under the standard normal curve between those two z values. The area between the two z values gives the area between the corresponding x values.

- Applications of the Normal Distribution
- Determining the z and x values

How to find the corresponding value of z or x when an area under a normal distribution curve is known.

Example:

1. Find the point z such that the area under the standard normal curve to the left of z is 0.9251.

Solution: to find the required value of *z*, we locate 0.9251 in the body of the normal distribution table. Next we read the numbers in column and row for *z* that correspond to 0.9251, which is 1.4 and 0.04. Combining these two numbers, we obtain the required value of z = 1.44.

APPROXIMATIONS

Under certain conditions:

-The Binomial distribution can be used to approximate the Hypergeometric distribution

- The Poisson distribution can be used to approximate the Binomial distribution.

-The normal distribution can be used to approximate both the Binomial and Poisson distributions.

##