- There are many situations in everyday life where the outcome is not known with certainty. For example; applying for a job or sitting an examination.
- We use words like "Chance", "the odds", "likelihood" etc but the most effective way of dealing with uncertainty is based on the concept of probability.
- Probability can be thought of as a number which measures the chance or likelihood that a particular event will occur.
- An example of the use of probability is in decision making. Decision making usually involves uncertainty. For example, should we invest in a company if there is a chance it will fail?
- Should we start production of a product even though there is a likelihood that the raw materials will arrive on time in poor? Having a number which measures the chances of these events occurring helps us to make a decision.
- Why are we interested in probability in this module? Many statistical methods use the idea of a probability distribution for this data.
- We have already looked at relative frequency distribution in Section 2. Probability distributions are based on the same concepts as relative frequency distributions. They are used to calculate probabilities of different values occurring in the data collected.

- We will examine probability distributions in more detail in Section 4. First we need to learn about the basic concepts of probability.
- **Sample Space**, S. For a given experiment the sample space, S, is the set of all possible outcomes.
- **Event**, E. This is a subset of S. If an event E occurs, the outcome of the experiment is contained in E.
- Probability concerns itself with random phenomena or probability experiments. These experiments are all different in nature, and can concern things as diverse as rolling dice or flipping coins.
- The common thread that runs throughout these probability experiments is that there are observable outcomes. If we collect all of the possible outcomes together, then this forms a set that is known as the sample space.

In this set theory formulation of probability the sample space for a problem corresponds to an important set. Since the sample space contains every outcome that is possible, it forms a setting of everything that we can consider. So the sample space becomes the universal set in use for a particular probability experiment.

A probability distribution is a table of values showing the probabilities of various outcomes of an experiment.

For example, if a coin is tossed three times, the number of heads obtained can be 0, 1, 2 or 3. The probabilities of each of these possibilities can be tabulated as shown:

Number of Heads	0	1	2	3
Probability	1/8	3/8	3/8	1/8

A discrete variable is a variable which can only take a countable number of values. In this example, the number of heads can only take 4 values (0, 1, 2, 3) and so the variable is discrete. The variable is said to be random if the sum of the probabilities is one.

Sample spaces abound and are infinite in number. But there are a few that are frequently used for examples in introductory statistics. Below are the experiments and their corresponding sample spaces:

- For the experiment of flipping a coin, the sample space is Heads, Tails and has two elements.
- For the experiment of flipping two coins, the sample space is (Heads, Heads), (Heads, Tails), (Tails, Heads), (Tails, Tails) and has four elements.
- For the experiment of flipping three coins, the sample space is (Heads, Heads, Heads), (Heads, Heads, Tails), (Heads, Tails, Heads), (Heads, Tails, Tails), (Tails, Heads, Heads), (Tails, Heads, Tails), (Tails, Tails, Heads), (Tails, Tails, Tails) and has eight elements.
- For the experiment of flipping n coins, where n is a positive whole number, the sample space consists of 2n elements. There are a total of $C(n, k)$ ways to obtain k heads and $n - k$ tails for each number k from 0 to n.

- For the experiment consisting of rolling a single six-sided die, the sample space is

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

- For the experiment of rolling two six-sided dice, the sample space consists of the set of the 36 possible pairings of the numbers 1, 2, 3, 4, 5 and 6.
- For the experiment of rolling three six-sided dice, the sample space consists of the set of the 216 possible triples of the numbers 1, 2, 3, 4, 5 and 6.
- For an experiment of drawing from a standard deck of cards, the sample space is the set that lists all 52 cards in a deck. For this example the sample space could only consider certain features of the cards, such as rank or suit.

Forming Other Sample Spaces

- These are the basic sample spaces. Others are out there for different experiments. It is also possible to combine several of the above experiments.
- When this is done, we end up with a sample space that is the Cartesian product of our individual sample spaces. We can also use a tree diagram to form these sample spaces.
- Probability theory is the mathematical study of randomness. A probability model of a random experiment is defined by assigning probabilities to all the different outcomes.

- Probability is a numerical measure of the likelihood that an event will occur. Thus, probabilities can be used as measures of degree of uncertainty associated with outcomes of an experiment. Probability values are always assigned on a scale from 0 to 1.
- A probability of 0 means that the event is impossible, while a probability near 0 means that it is highly unlikely to occur.
- Similarly an event with probability 1 is certain to occur, whereas an event with a probability near to 1 is very likely to occur.
- In the study of probability any process of observation is referred to as an experiment.
- The results of an experiment (or other situation involving uncertainty) are called the outcomes of the experiment.
- An experiment is called a random experiment if the outcome can not be predicted.
- Typical examples of a random experiment are
  - a role of a die,
  - a toss of a coin,
  - drawing a card from a deck.

  If the experiment is yet to be performed we refer to possible outcomes or possibilities for short. If the experiment has been performed, we refer to realized outcomes or realizations.

- The set of all possible outcomes of a probability experiment is called a ***sample space***, which is usually denoted by $\boldsymbol{S}$.
- The sample space is an exhaustive list of all the possible outcomes of an experiment. We call individual elements of this list ***sample points***.
- Each possible outcome is represented by one and only one sample point in the sample space.

For each of the following experiments, write out the sample space.

- Experiment: Rolling a die once
  - Sample space $S = \{1, 2, 3, 4, 5, 6\}$
- Experiment: Tossing a coin
  - Sample space $S = \{Heads, Tails\}$
- Experiment: Measuring a randomly selected persons height (cms)
  - Sample space $S =$ The set of all possible real numbers.
- An event is a specific outcome, or any collection of outcomes of an experiment.
- Formally, any subset of the sample space is an event.
- Any event which consists of a single outcome in the sample space is called an ***elementary*** or ***simple event***.

- Events which consist of more than one outcome are called ***compound events.***
- For example, an elementary event associated with the die example could be the "die shows 3".
- An compound event associated with the die example could be the "die shows an even number".
- The complement of an event A is the set of all outcomes in the sample space that are not included in the outcomes of event A.
- We call the complement event of A as A^c.
- The complement event of a die throw resulting in an even number is the die throwing an odd number.
- Question: if there is a 40% chance of a randomly selected student being male, what is the probability of the selected student being female?

Set theory is used to represent relationships among events.

Union of two events:

The union of events A and B is the event containing all the sample points belonging to A or B or both. This is denoted $A \cup B$, (pronounce as "A union B").

Intersection of two events:

The intersection of events A and B is the event containing all the sample points common to both A and B. This is denoted $A \cap B$, (pronounce as "A intersection B").

In general, if A and B are two events in the sample space S, then

- $A \subseteq B$ (A is a subset of B) = 'if A occurs, so does B
- $\varnothing$ (the empty set) = an impossible event
- S (the sample space) = an event that is certain to occur

Consider the experiment of rolling a die once. From before, the sample space is given as $S = \{1, 2, 3, 4, 5, 6\}$. The following are examples of possible events.

- A = score $< 4 = \{1, 2, 3\}$.
- B = 'score is even' $= \{2, 4, 6\}$.
- C = 'score is 7' $= 0$
- $A \cup B$ = 'the score is < 4 or even or both' $= \{1, 2, 3, 4, 6\}$
- $A \cap B$ = 'the score is < 4 and even $= \{2\}$
- A^c ='event A does not occur' $= \{4, 5, 6\}$

Chapter 1

Mathematical Fundamentals

The factorial function

The factorial function (symbol: !) just means to multiply a series of descending natural numbers.
Examples:

- $4! = 4 \times 3 \times 2 \times 1 = 24$
- $7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5,040$
- $1! = 1$
- $0! = 1$

Importantly

$$n! = n \times (n-1)! = n \times (n-1) \times (n-2)!$$

For Example

$$6! = 6 \times 5! = 6 \times 5 \times 4!$$

A factorial is a positive whole number, based on a number n , and which is written as "$n!$". The factorial $n!$ is defined as follows:

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

Remark $n! = n \times (n-1)!$

Example:

- $3! = 3 \times 2 \times 1 = 6$
- $4! = 4 \times 3! = 4 \times 3 \times 2 \times 1 = 24$

Remark $0! = 1$ not 0.

A factorial is a positive whole number, based on a number n , and which is written as "$n!$". The factorial $n!$ is defined as follows:

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

Remark $n! = n \times (n-1)!$

Example:

- $3! = 3 \times 2 \times 1 = 6$
- $4! = 4 \times 3! = 4 \times 3 \times 2 \times 1 = 24$

Remark $0! = 1$ not 0.

- factorials

$$n! = (n) \times (n-1) \times (n-2) \times \ldots \times 1$$

  - $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$
  - $3! = 3 \times 2 \times 1$

- Zero factorial

$$0! = 1$$

Binomial Coefficients

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

- $\binom{6}{2} = 15$
- $\binom{5}{2} = 10$
- $\binom{4}{0} = 1$
- $\binom{4}{3} = 4$

Exercises

Evaluate the following:

(i) $\binom{5}{2}$

(ii) $\binom{5}{0}$

(iii) $\binom{6}{3}$

(iv) $\binom{6}{6}$

(v) $\binom{10}{1}$

(vi) $\binom{10}{9}$

1.1 Counting Problems

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

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

1.2 Permutations and Combinations

Often we are concerned with computing the number of ways of selecting and arranging groups of items.

- A ***combination*** describes the selection of items from a larger group of items.
- A ***permutation*** is a combination that is arranged in a particular way.

- Suppose we have items A,B,C and D to choose two items from.
- AB is one possible selection, BD is another. AB and BD are both combinations.
- More importantly, AB is one combination, for which there are two distinct permutations: AB and BA.

Combinations: The number of ways of selecting k objects from n unique objects is:

$$^{n}C_{k} = \frac{n!}{k! \times (n-k)!}$$

In some texts, the notation for finding the number of possible combination is written

$$^{n}C_{k} = \binom{n}{k}$$

How many ways are there of selecting two items from possible 5?

$$^{5}C_{2} \left(\text{ also } \binom{5}{2} \right) = \frac{5!}{2! \times 3!} = \frac{5 \times 4 \times 3!}{2 \times 1 \times 3!} = 10$$

Discuss how combinations can be used to compute the number of rugby matches for each group in the Rugby World Cup.

The number of different permutations of r items from n unique items is written as $^{n}P_{k}$

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

Example: How many ways are there of arranging 3 different jobs, between 5 workers, where each worker can only do one job?

$$^{5}P_{3} = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$$

1.3 Permutations

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

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

Worked Example A committee of 4 must be chosen from 3 females and 4 males.

- In how many ways can the committee be chosen.
- In how many can 2 males and 2 females be chosen.
- Compute the probability of a committee of 2 males and 2 females are chosen.
- Compute the probability of at least two females.

A committee of 4 must be chosen from 3 females and 4 males.

- In how many ways can the committee be chosen.
- In how many cans 2 males and 2 females be chosen.
- Compute the probability of a committee of 2 males and 2 females are chosen.
- Compute the probability of at least two females.

Part 1

We need to choose 4 people from 7:
This can be done in

$$^7C_4 = \frac{7!}{4! \times 3!} = \frac{7 \times 6 \times 5 \times 4!}{4! \times 3!} = 35 \text{ ways.}$$

Part 2

With 4 men to choose from, 2 men can be selected in

$$^4C_2 = \frac{4!}{2! \times 2!} = \frac{4 \times 3 \times 2!}{2! \times 2!} = 6 \text{ ways.}$$

Similarly 2 women can be selected from 3 in

$$^3C_2 = \frac{3!}{2! \times 1!} = \frac{3 \times 2!}{2! \times 1!} = 3 \text{ ways.}$$

When implementing combination calculations in `R`, we use the `choose()` function.

```
> choose(5,0)
[1] 1
> choose(5,1)
[1] 5
> choose(5,2)
[1] 10
> choose(5,3)
[1] 10
> choose(5,4)
[1] 5
> choose(5,5)
[1] 1
```

Part 2

Thus a committee of 2 men and 2 women can be selected in $6 \times 3 = 18$ ways.

Part 3

The probability of two men and two women on a committee is

$$\frac{\text{Number of ways of selecting 2 men and 2 women}}{\text{Number of ways of selecting 4 from 7}} = \frac{18}{35}$$

Part 4

- The probability of at least two females is the probability of 2 females or 3 females being selected.
- We can use the addition rule, noting that these are two mutually exclusive events.
- From before we know that probability of 2 females being selected is 18/35.
- We have to compute the number of ways of selecting 1 male from 4 (4 ways) and the number of ways of selecting three females from 2 (only 1 way)
- The probability of selecting three females is therefore $\frac{4\times1}{35} = 4/35$
- So using the addition rule

$$Pr(\text{ at least 2 females }) = Pr(\text{ 2 females }) + Pr(\text{ 3 females })$$

$$Pr(\text{ at least 2 females }) = 18/35 + 4/35 = 22/35$$

Part 1

We need to choose 4 people from 7:
This can be done in

$$^7C_4 = \frac{7!}{4! \times 3!} = \frac{7 \times 6 \times 5 \times 4!}{4! \times 3!} = 35 \text{ ways.}$$

Part 2

With 4 men to choose from, 2 men can be selected in

$$^4C_2 = \frac{4!}{2! \times 2!} = \frac{4 \times 3 \times 2!}{2! \times 2!} = 6 \text{ ways.}$$

Similarly 2 women can be selected from 3 in

$$^3C_2 = \frac{3!}{2! \times 1!} = \frac{3 \times 2!}{2! \times 1!} = 3 \text{ ways.}$$

(Hint Multiplication Rule)

Part 2

Thus a committee of 2 men and 2 women can be selected in $6 \times 3 = 18$ ways.

Part 3

The probability of two men and two women on a committee is

$$\frac{\text{Number of ways of selecting 2 men and 2 women}}{\text{Number of ways of selecting 4 from 7}} = \frac{18}{35}$$

Part 4

- The probability of at least two females is the probability of 2 females or 3 females being selected.
- We can use the addition rule, noting that these are two mutually exclusive events.
- From before we know that probability of 2 females being selected is 18/35.

Part 4

- We have to compute the number of ways of selecting 1 male from 4 (4 ways) and the number of ways of selecting three females from 2 (only 1 way)
- The probability of selecting three females is therefore $\frac{4 \times 1}{35} = 4/35$
- So using the addition rule

$$Pr(\text{ at least 2 females }) = Pr(\text{ 2 females }) + Pr(\text{ 3 females })$$

$$Pr(\text{ at least 2 females }) = 18/35 + 4/35 = 22/35$$

Computing binomial Coefficients with R

When implementing combination calculations in `R`, we use the `choose()` function.

```
> choose(5,0)
[1] 1
> choose(5,1)
[1] 5
> choose(5,2)
[1] 10
> choose(5,3)
[1] 10
> choose(5,4)
[1] 5
> choose(5,5)
[1] 1
```

Permutations

9B.1 Permutation

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

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

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

$$\binom{6}{3} = \frac{6!}{(6-3)!3!} = \frac{6!}{3! \times 3!}$$

$$\frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4 \times 3!}{3! \times 3!} = \frac{120}{6} = 120$$

Type of Permutations

There are two types of permutation:

1. Repetition is Allowed: such as the lock above. It could be "333".
2. No Repetition: for example the first three people in a running race. You can't be first and second.

Summary

- If the order doesn't matter, it is a Combination.
- If the order does matter it is a Permutation.

Permutations : Worked Example

How many anagrams (permutations of the letters) are there of the following words

1. ANSWER
2. PERMUTE
3. ANAGRAM
4. LITTLE

Part 1 : ANSWER

Examples:

ASNWRE, SANERW, REWSAN, ...

Since ANSWER has 6 distinct letters, the number of permutations (anagrams) is

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = \mathbf{720}$$

Part 2 : PERMUTE

- The word PERMUTE has 7 letters, but only 6 different letters.
- There are $7!$ ways to arrange 7 letters.

- However, interchanging the two Es does not result in a new permutation. There would be two identical anagrams.

  PERMUTE, MUTEPER, PETEMUR, ..
  PERMUTE, MUTEPER, PETEMUR, ..

- The number of permutations (anagrams) is half of 7! .

$$\frac{7!}{2} = \frac{5040}{2} = \mathbf{2520}$$

Part 3 : ANAGRAM

- The word ANAGRAM has 7 letters, but there are three As.
- From before, there are 7! ways to arrange 7 letters.
- How many new permutations are found by re-arranging the As?

(i) ANAGRAM

(ii) ANAGRAM

(iii) ANAGRAM

(iv) ANAGRAM

(v) ANAGRAM

(vi) ANAGRAM

- We divide 7! by 3! to account for the identical anagrams.

$$\frac{7!}{3!} = \frac{5040}{6} = \mathbf{840}$$

1.3.1 Permutations

Part 2 : PERMUTE

- We re-express the answer from part 2 as follows:

$$\frac{7!}{2!} = \frac{5040}{2} = \mathbf{2520}$$

Part 4 : LITTLE

- The word LITTLE has 6 letters, but there are two Ls and two Ts.
- From before, there are 6! ways to arrange 6 letters.
- Again, interchanging the two Ls and Ts does not result in a new permutation.

$$\frac{6!}{2! \times 2!} = \frac{720}{4} = \mathbf{180}$$

1.3.2 Permutations

- In how many permutations are there of counting a subset of k elements, when there are n elements in total.
- The number of permutations of a set of n elements is denoted n! (pronounced n factorial.)

1.3.3 Permutation Formula

A formula for the number of possible permutations of k objects from a set of n. This is usually written ${}^{n}P_{k}$.

Formula:

$$ {}^{n}P_{k} = \frac{n!}{(n-k)!} = n.(n-1).(n-2)\ldots.(n-k+1) $$

Example:

How many ways can 4 students from a group of 15 be lined up for a photograph?

Answer:

There are ${}^{15}P_{4}$ possible permutations of 4 students from a group of 15.

$$ {}^{15}P_{4} = \frac{15!}{11!} = 15 \times 14 \times 13 \times 12 = 32760 $$

There are 32760 different lineups.

1.4 Combinations and Permutations

Often we are concerned with computing the number of ways of selecting and arranging groups of items.

- A ***combination*** describes the selection of items from a larger group of items.
- A ***permutation*** is a combination that is arranged in a particular way.

- Suppose we have items A,B,C and D to choose two items from.
- AB is one possible selection, BD is another. AB and BD are both combinations.
- More importantly, AB is one combination, for which there are two distinct permutations: AB and BA.

Combinations: The number of ways of selecting k objects from n unique objects is:

$$^nC_k = \frac{n!}{k! \times (n-k)!}$$

In some texts, the notation for finding the number of possible combination is written

$$^nC_k = \binom{n}{k}$$

How many ways are there of selecting two items from possible 5?

$$^5C_2 \left(\text{ also } \binom{5}{2} \right) = \frac{5!}{2! \times 3!} = \frac{5 \times 4 \times 3!}{2 \times 1 \times 3!} = 10$$

Discuss how combinations can be used to compute the number of rugby matches for each group in the Rugby World Cup.

The number of different permutations of r items from n unique items is written as ^nP_k

$$^nP_k = \frac{n!}{(n-k)!}$$

Example: How many ways are there of arranging 3 different jobs, between 5 workers, where each worker can only do one job?

$$^5P_3 = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60$$

Combinations

In mathematical terms, a combination is an subset of items from a larger set such that the order of the items does not matter.

1.4.1 Permutations

- The notion of permutation relates to the act of permuting (rearranging) objects or values.
- Informally, a permutation of a set of objects is an arrangement of those objects into a particular order.
- For example, there are six permutations of the set $\{1, 2, 3\}$, namely (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1).
- As another example, an anagram of a word is a permutation of its letters.

If the probability of C is 70% then the probability of C' is 30%

Worked Examples: Permuations and Combinations

Formula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1},$$

which can be written using factorials as whenever $k \leq n$

Example 1

$$\binom{5}{2} = \frac{5!}{2!\,(5-2)!} = \frac{5.4.3!}{2!.3!} = \frac{5.4}{2.1} = 10$$

Example 2

$$\binom{5}{0} = \frac{5!}{0!\,(5-0)!} = \frac{5!}{0!.5!} = \frac{5!}{2!} = 1$$

Recall $0! = 1$

Example 3 : Urn Question

Suppose an urn contains seven white, four black and three red beads. Three beads are picked at random without replacement. Find the probability that all three beads are the different in colour. at least two beads are the same colour.

- A bag contains 2 red, 3 green and 2 blue balls. Two balls are drawn at random. What is the probability that none of the balls drawn is blue?
- An IT consultant is responsible for three software engineering projects X, Y and Z. He knows that the probability of completing project X in time is 0.99, for project Y this probability is 0.95 and for project Z it is 0.80.

a What assumption do you need to make in order to calculate the probability of completing all three projects in time, from the information given?

b Calculate the probability of completing all three projects in time.

c Calculate the probability that only projects X and Y will be completed on time.

- A doctor treating a patient issues a prescription for antibiotics and provides for two repeat prescriptions. The probability that the infection will be cleared by the first prescription is p_1 =0.6. The probability that successive treatments are successful, given that previous prescriptions were not successful are $p_2 = 0.5$, $p_3 = 0.4$. Calculate the probability that:

  a a patient will require the third prescription,

  b the patient is still infected after the third prescription,

  c the patient is cured by the second prescription, given that the patient is eventually cured.

- Two people look at the letters in the word discovery. Independently of each other, each person writes down two of the letters from the word discovery. What is the probability that

  (i) One person writes down two vowels and the other person

  (ii)

- Three cards are drawn, one after the other, without replacement, from a pack of 52 playing cards. Find the probability that the
- On completion of a programming project, three programmers from a team submit a collection of subroutines to an acceptance group.

  The following table shows the percentage of subroutines each programmer submitted and the probability that a subroutine submitted by each programmer will pass the certification test based on historical data.

Programmer	A	B	C
Proportion of subroutines submitted	0.40	0.35	0.25
Probability of acceptance	0.75	0.95	0.85

i. (3 marks) What is the proportion of subroutines that pass the acceptance test?

ii. (3 marks) After the acceptance tests are completed, one of the subroutines is selected at random and found to have passed the test. What is the probability that it was written by Programmer A?

2 In how many ways can a group of four people be selected from three men and four women? In how many of these groups are there more women than men?

3 In how many ways can a group of five be selected from ten people How many groups can be selected if two particular people from the

ten can not be selected in the same group?

Counting Sets using Venn Diagrams

4 The Venn Diagram shows the number of elements in each subset of set S. If $P(A) = 3/10$ and $P(B) = 1/2$, find the values of x and y

5 How many different four digit numners greater than 5000 can be formed from the digits **2,4,5,8,9** if each digit can only be used once in any given number. How many of these numbers are odd?

1.5 Counting

Given S is the set of all 5 digit binary strings, E is the set of a 5 digit binary strings beginning with a 1 and F is the set of all 5 digit binary strings ending with two zeroes.

(a) Find the cardinality of S, E and F.

(b) Draw a Venn diagram to show the relationship between the sets S, E and F. Show the relevant number of elements in each region of your diagram.

- Previously we have been studying discrete random variables, such as the Binomial and the Poisson random variables.
- Now we turn our attention to continuous random variables.

- Recall that a continuous random variable is one which takes an infinite number of possible values, rather than just a countable number of distinct values.
- Continuous random variables are usually measurements.
- Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

Remarks: This is for continuous distributions only.

- The probability that a continuous random variable will take an exact value is infinitely small. We will usually treat it as if it was zero.
- When we write probabilities for continuous random variables in mathematical notation, we often retain the equality component (i.e. the "...or equal to..").
  For example, we would write expressions $P(X \leq 2)$ or $P(X \geq 5)$.
- Because the probability of an exact value is almost zero, these two expression are equivalent to $P(X < 2)$ or $P(X > 5)$.
- Also, the complement of $P(X \geq k)$ can be written as $P(X \leq k)$.
- Integration is not part of the syllabus, and it is assumed that students are not familiar with how to compute definite integrals.
- However, it is useful to know what the purpose of definite integrals are, because we will be using the results derived from definite integrals.
- It is assumed that students are familiar with functions.

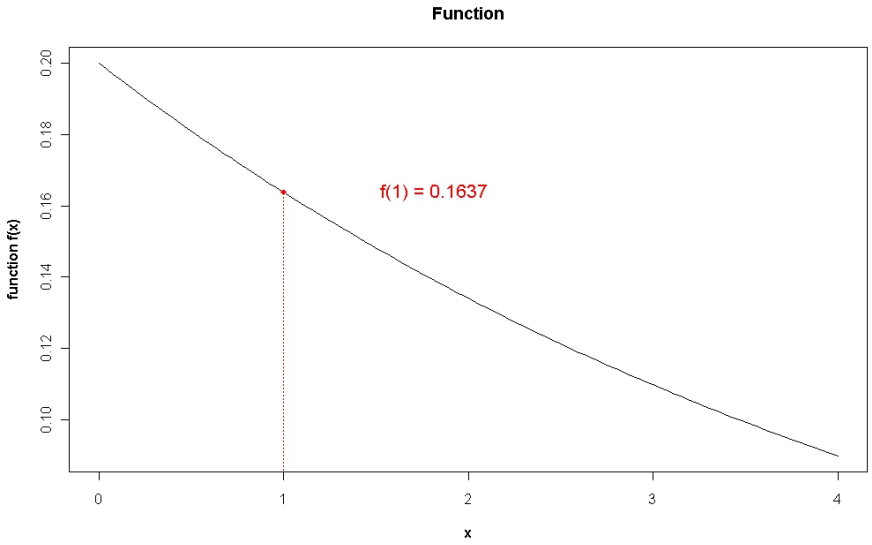
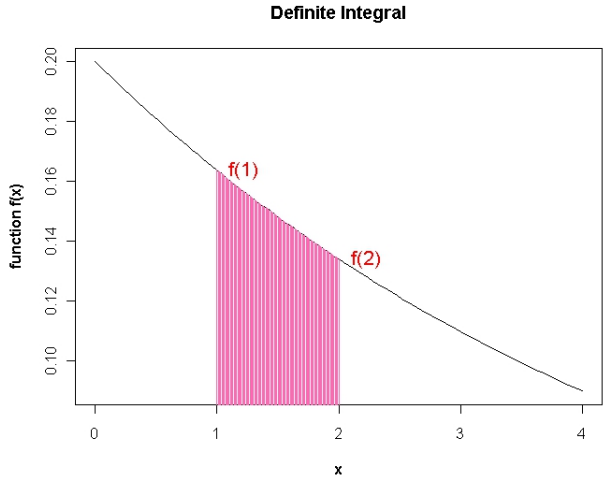


Figure 1.1:

Some function $f(x)$ evaluated at $x = 1$.



Definite integral of function is area under curve between X=1 and X=2.

- Definite integrals are used to compute the "***area under curves***".

- Definite integrals are defined by a lower and upper limit.
- The area under the curve between X=1 and X=2 is depicted in the previous slide.
- By computing the definite integral, we are able to determine a value for this area.
- Probability can be represented as an area under a curve.
- In probability theory, a ***probability density function*** (PDF) (or "density" for short) of a continuous random variable is a function that describes the relative likelihood for this random variable to occur at a given point.
- The PDF for a continuous random variable X is often denoted $f(x)$.
- The probability density function can be integrated to obtain the probability that the random variable takes a value in a given interval.
- The probability for the random variable to fall within a particular interval is given by the integral of this variable's density over the region.

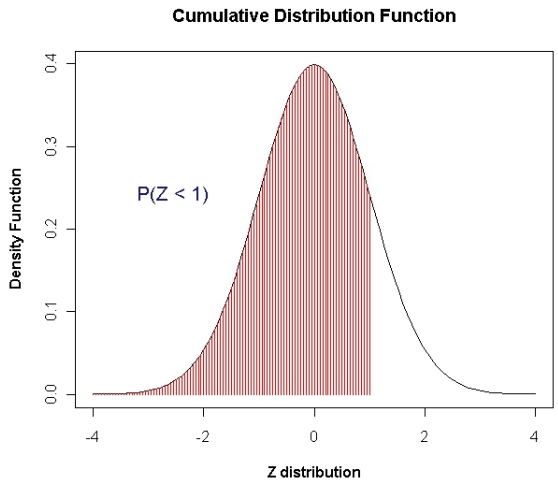
- The probability density function is non-negative everywhere, and its integral over the entire space is equal to one.
- A plot of the PDF is referred to as a '***density curve***'.
- A density curve that is always on or above the horizontal axis and has total area underneath equal to one.
- Area under the curve in a range of values indicates the proportion of values in that range.
- Density curves come in a variety of shapes, but the normal distribution's bell-shaped densities are perhaps the most commonly encountered.
- Remember the density is only an approximation, but it simplifies analysis and is generally accurate enough for practical use.

Recall:

- The ***cumulative distribution function*** (CDF), (or just distribution function), describes the probability that a continuous random variable X with a given probability distribution will be found at a value less than or equal to x.

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

- Intuitively, it is the "area so far" function of the probability distribution.



Cumulative Distribution Function $P(Z \leq 1)$
Here the random variable is called Z (we will see why later)

- Probability Density Function
- Cumulative Density Function

If X is a continuous random variable then we can say that the probability of obtaining a **precise** value x is infinitely small, i.e. close to zero.

$$P(X = x) \approx 0$$

Consequently, for continuous random variables (only), $P(X \leq x)$ and $P(X < x)$ can be used interchangeably.

$$P(X \leq x) \approx P(X < x)$$

Chapter 2

Compound Events

2.1 Probability Formulae

- Conditional probability:

$$P(B|A) = \frac{P\left(A \text{ and } B\right)}{P\left(A\right)}.$$

- Bayes' Theorem:

$$P(B|A) = \frac{P\left(A|B\right) \times P(B)}{P\left(A\right)}.$$

2.2 Basic of Compound Events

- pairwise disjoint sets
- The addition principle

Theorem

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Complement Rule

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

The complement rule in Probability
$P(C') = 1 - P(C)$
If the probability of C is 70% then the probability of C' is 30%

- The outcome of an experiment need not be a number, for example, the outcome when a coin is tossed can be 'heads' or 'tails'.
- However, we often want to represent outcomes as numbers.
- A ***random variable*** is a function that associates a unique numerical value with every outcome of an experiment.

- The value of the random variable will vary from trial to trial as the experiment is repeated.
- Numeric values can be assigned to outcomes that are not usually considered numeric.
- For example, we could assign a 'head' a value of 0, and a 'tail' a value of 1, or vice versa.

There are two types of random variable - discrete and continuous. The distinction between both types will be important later on in the course.

Examples

- A coin is tossed ten times. The random variable X is the number of tails that are noted. X can only take the values $\{0, 1, ..., 10\}$, so X is a discrete random variable.
- A light bulb is burned until it burns out. The random variable Y is its lifetime in hours. Y can take any positive real value, so Y is a continuous random variable.
- A discrete random variable is one which may take on only a countable number of distinct values such as $\{0, 1, 2, 3, 4, .$

- Discrete random variables are usually (but not necessarily) counts.
- If a random variable can take only a finite number of distinct values, then it must be discrete.
- Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.
- A continuous random variable is one which takes an infinite number of possible values.
- Continuous random variables are usually measurements.
- Examples include height, weight, the amount of sugar in an orange, the time required to run a computer simulation.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

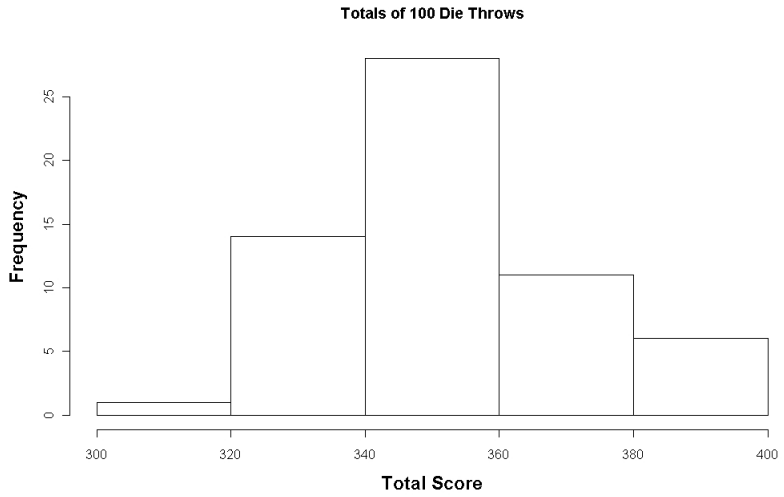
A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Bar-plots
- Histograms
- Boxplots
- Consider an experiment in which each student in a class of 60 rolls a die 100 times.
- Each score is recorded, and a total score is calculated.
- As the expected value of rolled die is 3.5, the expected total is 350 for each student.
- At the end of the experiment the students reported their totals.
- The totals were put into ascending order, and tabulated as follows (next slide).

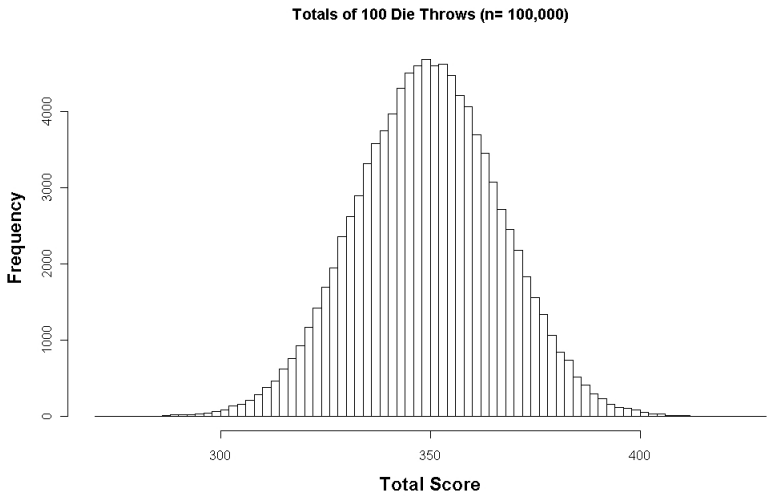
307	321	324	328	329	330	334	335	336	337
337	337	338	339	339	342	343	343	344	344
346	346	347	348	348	348	350	351	352	352
353	353	353	354	354	356	356	357	357	358
358	360	360	361	362	363	365	365	369	369
370	370	374	378	381	384	385	386	392	398

- What proportion of outcomes are less than or equal to 330?
(Answer: 10%)
- What proportion of outcomes are greater than or equal to 370?
(Answer: 16.66%)
- Compute an appropriate number of class intervals.
- As a rule of thumb, the number of class intervals is usually approximately the square root of the number of observations.
- As there are 60 observations, we would normally use 7 or 8 class intervals.
- To save time, we will just use 5 class intervals.



- Suppose that the experiment of throwing a die 100 times and recording the total was repeated 100,000 times.
- (If implemented on a computer, we would call this a simulation study)

- The histogram of data (with a class interval width of 2) is shown on the next slide.
- How should the shape of the histogram be described?
- "Bell-shaped" would be a suitable description.



A couple of remarks about the simulation study, some of which will be relevant later on.

- Approximately 68.7% of the values in the simulation study are between 332 and 367.
- Approximately 95% of the values are between 316 and 383.
- 2.5% of the values output are less than 316.
- 2.5% of the values study output are greater than 383.
- 175 values are greater than or equal to 400, whereas 198 values are less than or equal to 300.
- Results such as these are unusual, but they are not impossible.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the "success" and the other outcome the "failure".
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

MA4004 Section 1b - Probability Questions

Repeat 2008

Spring 2008 Question 1

a) One in 10, 000 people have a particular condition. Given that an individual has this condition, a test for this condition gives a positive result with probability 0.999. Given that an individual does not have this condition, this test gives a positive result with probability 0.001.

Suppose the individual tested is chosen at random from the population as a whole i) Calculate the probability that the test result is positive.

ii) Calculate the probability that the individual has the condition, given that the result of the test is positive. (6 marks)

Repeats 2007

a) A company obtains 2000 components/week from supplier A, 2000 components/week from supplier B and 1000 components/week from supplier C. 3

i) Calculate the probability that a randomly chosen component is defective.

ii) Calculate the probability that the component is from supplier A, given that it is not defective. (6 marks)

2. c) Calculate the probability of obtaining exactly 3 sixes when I roll a die 5 times. (4 marks)

Spring 2007

a) A company obtains 1500 components/week from supplier A, 1000 components/week from supplier B and 500 components/week from supplier C. 3

i) Calculate the probability that a randomly chosen component is defective.

ii) Calculate the probability that the component is from supplier C, given that it is defective. (6 marks) Question 3 c) Calculate the probability of obtaining exactly 2 sixes when I roll a die 5 times. (4 marks)

Repeats 2006 (c) A worker-operated machine produces a defective item with probability 0.01 if the worker follows the machines operating instructions exactly, and with probability 0.04 if he does not. If the worker follows the instructions 90

2.3 Independent Events

Competitors A and B fire at their respective targets. The probability that A hits a target is 1/3 and the probability that B hits a target is 1/5. Find the probability that:

i. (2 marks) A does not hit the target,

ii. (2 marks) both hit their respective targets,

iii. (2 marks) only one of them hits a target,

iv. (2 marks) neither A nor B hit their targets.

2.4 Mutually Exclusive Events

Mutually exclusive events are events that cannot happen at the same time.

$$P(AandB) = P(A) + P(B)$$

Probability

9B.2 The sample space of an experiment (S)

9B.3 The size of a sample space

9B.4 Indepedent Evcents (9.3.1)

2.5 Conditional Probability

What is the probability of one event given that another event occurs? For example, what is the probability of a mouse finding the end of the maze, given that it finds the room before the end of the maze?

This is represented as:

$$P[A|B]$$

or "the probability of A given B."

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

If A and B are independent of one another, such as with coin tosses or child births, then:

$$P[A|B] = P[A]$$

Thus, "what is the probability that the next child a family bears will be a boy, given that the last child is a boy."

This can also be stacked where the probability of A with several "givens."

$$P[A|B_1, B_2, B_3]$$

or "the probability of A given that B1, B2, and B3 are true?"

- pairwise disjoint sets
- The addition principle

Suppose an electronics assembly subcontractor receives resistors from two suppliers: A and B

- Supplier A supplies 80% of the resistors

- Supplier B supplies 20% of the resistors

Suppose an electronics assembly subcontractor recieves resistors from two suppliers A and B

- Supplier A supplies 80% of the resistors
- *Probability that a randomly chosen resistor comes from A is 80 %*
- *P(A) = 0.80*
- Supplier B supplies 20% of the resistors
- *Probability that a randomly chosen resistor comes from B is therefore 20%*
- *P(B) = 0.20*
- We are giving information about the rate of faulty components from each supplier.
  (Faulty : resistor fails some quality test)

- 1% of the resistors supplied by A are faulty

- 3% of the resistors supplied by B are faulty

- We are giving information about the rate of faulty components from each supplier.
  (Faulty : resistor fails some quality test)
- *P(F) probability that randomly selecting compoent is faulty*
- 1% of the resistors supplied by A are faulty.
- *We write this as* $P(F|A) = 0.01$
- 3% of the resistors supplied by B are faulty
- *We write this as* $P(F|B) = 0.03$

Question 1:

- What is the probability that a randomly selected resistor fails the final test?
- In mathematical terms, compute P(F)

Law of Total Probability:

- Faulty Resistors are either from Supplier A or Supplier B.
- *Resistors MUST come from one of the two suppliers.*
- *A and B are mutually exclusive.*

$$P(F) = P(F \text{ and } A) + P(F \text{ and } B)$$

Conditional Probability

$$P(X|Y) = \frac{P(X \text{ and } Y)}{P(Y)}$$

Re-arranging

$$P(X \text{ and } Y) = P(X|Y) \times P(Y)$$

Therefore we can say

$$P(F \text{ and } A) = P(F|A) \times P(A)$$

$$P(F \text{ and } B) = P(F|B) \times P(B)$$

$$P(F \text{ and } A) = P(F|A) \times P(A)$$

$$P(F \text{ and } B) = P(F|B) \times P(B)$$

$$P(F \text{ and } A) = P(F|A) \times P(A) = 0.80 \times 0.01$$

$$P(F \text{ and } A) = 0.008$$

$$P(F \text{ and } B) = P(F|B) \times P(B) = 0.20 \times 0.03$$

$$P(F \text{ and } B) = 0.006$$

Recall:

$$P(F) = P(F \text{ and } A) + P(F \text{ and } B)$$

Session 09: Probability

9A.1 Counting Methods

9A.2 Counting using Sets

9A.3 Probability

9A.4 Independent Events

Chapter 3

Introduction to Random Variables

3.1 Random Variables

- The outcome of an experiment need not be a number, for example, the outcome when a coin is tossed can be 'heads' or 'tails'.
- However, we often want to represent outcomes as numbers.
- A ***random variable*** is a function that associates a unique numerical value with every outcome of an experiment.
- The value of the random variable will vary from trial to trial as the experiment is repeated.
- Numeric values can be assigned to outcomes that are not usually considered numeric.
- For example, we could assign a 'head' a value of 0, and a 'tail' a value of 1, or vice versa.

There are two types of random variable - discrete and continuous. The distinction between both types will be important later on in the course.

Examples

- A coin is tossed ten times. The random variable X is the number of tails that are noted. X can only take the values $\{0, 1, ..., 10\}$, so X is a discrete random variable.
- A light bulb is burned until it burns out. The random variable Y is its lifetime in hours. Y can take any positive real value, so Y is a continuous random variable.
- A discrete random variable is one which may take on only a countable number of distinct values such as $\{0, 1, 2, 3, 4, ...\}$.
- Discrete random variables are usually (but not necessarily) counts.
- If a random variable can take only a finite number of distinct values, then it must be discrete.
- Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.
- A continuous random variable is one which takes an infinite number of possible values.

- Continuous random variables are usually measurements.
- Examples include height, weight, the amount of sugar in an orange, the time required to run a computer simulation.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X. A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

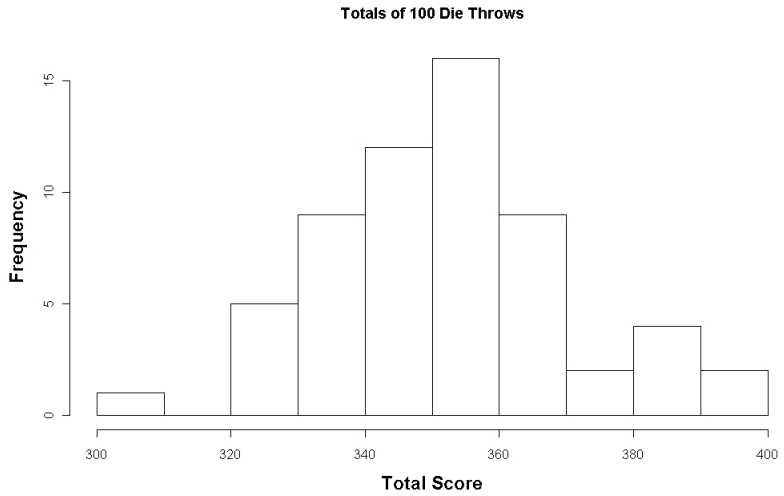
- (a) Find the distribution f of X.
- (b) Find the expectation E(X).
- Bar-plots
- Histograms
- Boxplots
- Consider an experiment in which each student in a class of 60 rolls a die 100 times.
- Each score is recorded, and a total score is calculated.
- As the expected value of rolled die is 3.5, the expected total is 350 for each student.

- At the end of the experiment the students reported their totals.
- The totals were put into ascending order, and tabulated as follows (next slide).

307	321	324	328	329	330	334	335	336	337
337	337	338	339	339	342	343	343	344	344
346	346	347	348	348	348	350	351	352	352
353	353	353	354	354	356	356	357	357	358
358	360	360	361	362	363	365	365	369	369
370	370	374	378	381	384	385	386	392	398

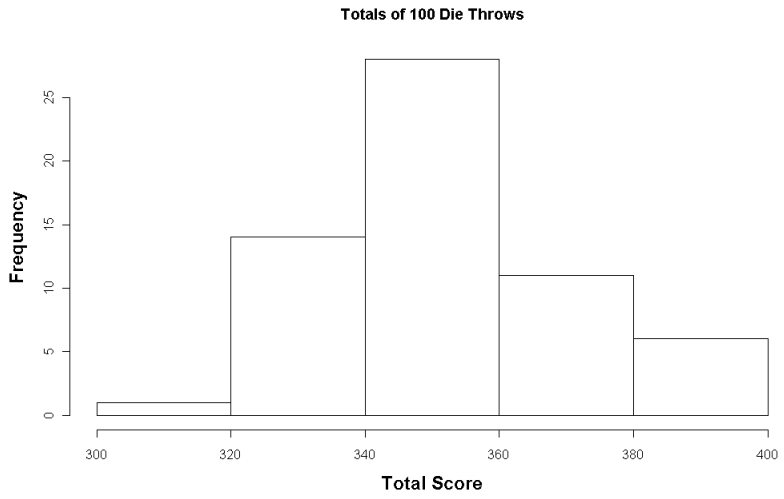
- What proportion of outcomes are less than or equal to 330? (Answer: 10%)
- What proportion of outcomes are greater than or equal to 370? (Answer: 16.66%)

For the die-throw experiment;

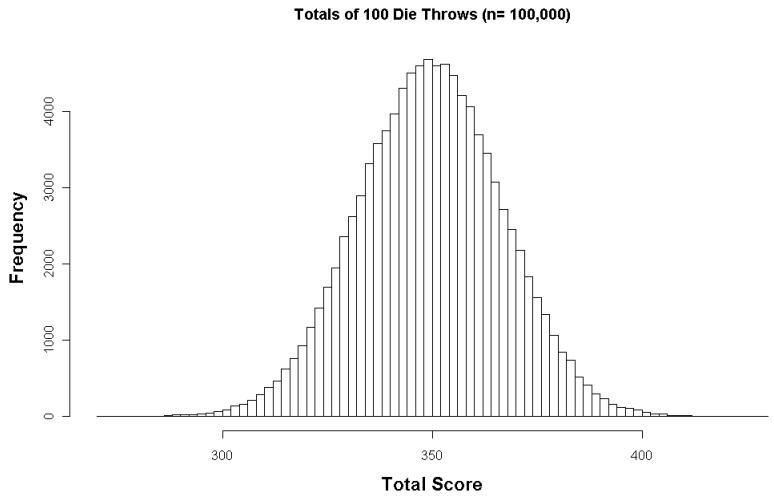


- Compute an appropriate number of class intervals.
- As a rule of thumb, the number of class intervals is usually approximately the square root of the number of observations.

- As there are 60 observations, we would normally use 7 or 8 class intervals.
- To save time, we will just use 5 class intervals.



- Suppose that the experiment of throwing a die 100 times and recording the total was repeated 100,000 times.
- (If implemented on a computer, we would call this a simulation study)
- The histogram of data (with a class interval width of 2) is shown on the next slide.
- How should the shape of the histogram be described?
- "Bell-shaped" would be a suitable description.



A couple of remarks about the simulation study, some of which will be relevant later on.

- Approximately 68.7% of the values in the simulation study are between 332 and 367.
- Approximately 95% of the values are between 316 and 383.
- 2.5% of the values output are less than 316.
- 2.5% of the values study output are greater than 383.
- 175 values are greater than or equal to 400, whereas 198 values are less than or equal to 300.
- Results such as these are unusual, but they are not impossible.

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation E(X).

- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the "success" and the other outcome the "failure".
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

3.2 What is a Probability Distribution

A statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. This range will be between the minimum and maximum statistically possible values, but where the possible value is likely to be plotted on the probability distribution depends on a number of factors, including the distributions mean, standard deviation, skewness and kurtosis.

Thirty-eight students took the test. The X-axis shows various intervals of scores (the interval labeled 35 includes any score from 32.5 to 37.5). The Y-axis shows the number of students scoring in the interval or below the interval.

cumulative frequency distributionA can show either the actual frequencies at or below each interval (as shown here) or the percentage of the scores at or below each interval. The plot can be a histogram as shown here or a polygon.

Probability Distributions (Question 2 for End Of Year Exam)

- Discrete Probability Distributions
  - Binomial Probability Distribution (Week 3)
  - Geometric Probability Distribution (Week 3)
  - Poisson Probability Distribution (Week 3/4)
- Continuous Probability Distributions

– Exponential Probability Distribution (Week 4)
– Uniform Probability Distribution (Week 4)
– Normal Probability Distribution (Week 4/5)

3.3 Quantiles

The quantile (this term was first used by Kendall, 1940) of a distribution of values is a number xp such that a proportion p of the population values are less than or equal to xp. For example, the .25 quantile (also referred to as the 25th percentile or lower quartile) of a variable is a value (xp) such that 25% (p) of the values of the variable fall below that value.

Similarly, the 0.75 quantile (also referred to as the 75th percentile or upper quartile) is a value such that 75% of the values of the variable fall below that value and is calculated accordingly.

See

Chapter 4

Probability

4.1 Dice Questions

Suppose a pair of fair dice is thrown.

(a) What is the probability of getting a sum of 9 from two throws of a dice

Find the probability that the sum is 10 or greater if

(b) a 5 appears on the first die,

(c) a 5 appears on at least one of the dice.

Tickets numbered 1 to 20 are mixed up and then a ticket is drawn at random. What is the probability that the ticket drawn has a number which is a multiple of 3 or 5

Session 9 Probability

The complement rule in Probability

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

If the probability of C is 70% then the probability of C' is 30%

- Permutations where repetition is allowed:

$$n!$$

- Permutations where repetition is not allowed

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

4.2 Expected Values of Random Variables

If the random variable Z has a distribution which is standard normal, show that the expected value of e^{sZ} is given as follows:

$$E(e^{sZ}) = e^{\frac{s^2}{2}}$$

- In general, the expected value is computed using this formula

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

- The expected value of a ***transformed*** random variable is computed using this formula

$$E(tf(X)) = \int_{-\infty}^{\infty} tf(x) \times f(x)dx$$

- The probability density function for the standard normal distribution is

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

- The probability density function for the standard normal distribution is

$$f(z) = f(x, \mu = 0, \sigma = 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}^2}$$

$$E(e^{sZ}) = \int_{-\infty}^{\infty} e^{sx} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}^2}$$

$$E(e^{sZ}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[sx - \frac{(x)^2}{2}\right]} dx$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[-\frac{(x-s)^2}{2}\right]} dx$$

Mathematical Identity

- Proven in a separate video

$$\int_{-\infty}^{\infty} e^{-\frac{y}{2}^2} dy = \sqrt{2\pi}$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[-\frac{(x-s)^2}{2}\right]} dx$$

$$E(e^{sZ}) = e^{\frac{s^2}{2}} \times \frac{1}{\sqrt{2\pi}} \left[\sqrt{2\pi}\right]$$

4.3 Worked Example

Suppose an electronics assembly subcontractor recieves resistors from two suppliers A and B

Supplier A supplies 80% of the resistors

P(A) = 0.80 probability that a randomly chosen resistor comes from A

Supplier B supplies 20% of the resistors

P(B) = 0.20 probability that a randomly chosen resistor comes from B

- 1% of the resistors supplied by A are faulty (i.e. resistor fails the final test)
- 3% of the resistors supplied by B are faulty

Question: What is the probability that a randomly selected resistor fails the final test?

Compute P(F)

$$P(F) = P(FandA) + P(FandB)$$

A probability distribution is a mathematical approach to quantifying uncertainty.

There are two main classes of probability distributions: Discrete and continuous.

Discrete distributions describe variables that take on discrete values only (typically the positive integers), while continuous distributions describe variables that can take on arbitrary values in a continuum (typically the real numbers).

Binomial Distribution : Worked Example

A manufacturer of hospital equipment knows from experience that 5% of the production will have some type of minor default, and will require adjustment.

Number of independent trials n

Probability of a "success" p

A basic introduction to the concept

Example

Certain events happen at unpredictable intervals. But for some reason, no matter how recent or long ago last event was, the probability that another event will occur within the next hour is exactly the same (say, 10%). The same holds for any other time interval (say, second). Moreover, the number of events within any given time interval is statis-

tically independent of numbers of events in other intervals that do not overlap the given interval. Also, two events never occur simultaneously.

Then the number of events per day is Poisson distributed.

Formal definition

Let X be a stochastic variable taking non-negative integer values with probability density function

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

Then X follows the Poisson distribution with parameter λ.

Characteristics of the Poisson distribution

If X is a Poisson distribution stochastic variable with parameter λ, then

- The expected value $E[X] = \lambda$
- The variance $Var[X] = \lambda$

The Normal Distribution

Symmetric Intervals

Symmetric Intervals

$$P(-z \leq Z \leq z)$$

The Normal Distribution

The Symmetry Rule

4.3.1 Normal Distribution : The Symmetry Rule

From statistical tables, we could determine the following:

- $P(Z \leq 1.5)$
- $P(Z \geq 1.5)$

Consider the normally distributed random variable X

$$X \sim \mathcal{N}(\mu = 1000, \sigma^2 = 2500)$$

Paramters:

- $\mu = 1000$
- $\sigma = 50$

Questions

- $P(X \leq 925)$
- $P(X \geq 925)$

Z-score

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

$$X \sim \mathcal{N}(\mu = 1000, \sigma^2 = 2500)$$

- Mean $\mu = 1000$
- Standard Deviation $\sigma = 50$

$$P(X \leq 925) = P(Z \leq -1.5)$$

Applying the Symmetry Rule

$$P(Z \leq -1.5) = P(Z \geq 1.5) = 0.0668$$

Theerefore we can say

$$P(X \leq 925) = 0.0668$$

As a consequence of Property 1, it is possible to relate all normal random variables to the standard normal.

If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma}$$

is a standard normal random variable: $Z \sim N(0, 1)$. An important consequence is that the cdf of a general normal distribution is therefore

$$\Pr(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x - \mu}{\sigma\sqrt{2}}\right)\right)$$

.

Conversely, if Z is a standard normal distribution, $Z \sim N(0, 1)$, then

$$X = \sigma Z + \mu$$

is a normal random variable with mean μ and variance σ^2.

The standard normal distribution has been tabulated (usually in the form of value of the cumulative distribution function F), and the other normal distributions are the simple transformations, as described above, of the standard one. Therefore, one can use tabulated values of the cdf of the standard normal distribution to find values of the cdf of a general normal distribution.

4.3.2 Statistics

1. Sample mean
$$\bar{x} = \frac{\sum x_i}{n}.$$
2. Sample standard deviation
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}.$$
3. Conditional probability:
$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}.$$

MA4704 Tutorial Sheet 2.

Question 1

A doctor treating a patient issues a prescription for antibiotics and provides for two repeat prescriptions. The probability that the infection

will be cleared by the first prescription is p1 =0.6. The probability that successive treatments are successful, given that previous prescriptions were not successful are p2 = 0.5, p3 = 0.4. Calculate the probability that

1. the patient is still infected after the third prescription 2. the patient is cured by the second prescription. 3. the patient is cured by the second prescription, given that the patient is eventually cured.

Question 2

A driver passes through 3 traffic lights. The chance he/she will stop at the first is 1/2 , at the second 1/3 and at the third independently of what happens at any of the other lights.

What is the probability that 4. the driver makes the whole journey without being stopped at any of the lights 5. the driver is only stopped at the first and third lights 6. the driver is stopped at just one set of lights. 7. the driver stopped at the second set of lights, given he/she stopped at one set of lights.

Question 3

The masses of 30 human males and 30 arabian stallions were observed. Their masses (in lbs) are given below

Humans 106, 120, 130, 138, 145, 151, 156, 161, 166, 171 176, 180, 185, 189, 194, 198, 203, 208, 212, 217 223, 228, 234, 240, 247, 255, 264, 276, 290, 313

Stallions 808, 824, 835, 843, 851, 857, 862, 868, 872, 877 881, 886, 890, 894, 898, 902, 906, 910, 914, 919 923, 928, 932, 938, 943, 949, 957, 965, 976, 992

a) Draw histograms for these samples and compare them with respect to shape, centrality and relative dispersion. b) Calculate the medians of these samples (from the raw data).

Question 4

The following data give the marks of 10 students in a test (out of 20 marks). Calculate i) the median ii) the mean iii) the range iv) the standard deviation v) The Inter-Quartile Range

12, 17, 7, 11, 18, 6, 14, 15, 11, 9.

Question 1 : Probability Distribution

Introduction

Consider playing a game in which you are winning when a ***fair die*** is showing 'six' and losing otherwise.

Part 1

If you play three such games in a row, find the probability mass function (pmf) of the number X of times you have won.

- Firstly: what type of probability distribution is this?

- Is this the distribution ***discrete*** or ***continuous***?
- The outcomes are whole numbers - so the answer is discrete.
- So which type of discrete distribution? (We have two to choose from. See first page of formulae)
- **Binomial:** characterizing the number of ***successes*** in a series of n ***independent trials***, with the ***probability of a success*** in each trial being p.
- **Poisson:** characterizing the ***number of occurrences*** in a ***unit space*** (i.e. a unit length, unit area or unit volume, or a unit period in time), where λ is the the number of occurrences per unit space.

4.3.3 Standardisation Formula

$$Z = (X - \mu)/\sigma \tag{4.1}$$

4.4 Discrete Random Variables

1. The probability distribution of discrete random variable X is tabulated below. There are 6 possible outcome of X, i.e. 0, 1, 2, 4 ,8 and 10.

x_i	0	1	2	4	8	10
$P(x_i)$	0.25	0.15	0.25	0.15	k	0.10

i. (1 marks) Compute the value for k.

ii. (3 marks) Determine the expected value $E(X)$.

iii. (2 marks) Evaluate $E(X^2)$.

iv. (3 marks) Compute the variance of random variable X.

2. Suppose X is a random variable with

- $E(X^2) = 3.6$
- $P(X = 2) = 0.6$
- $P(X = 3) = 0.1$

(a) The random variable takes just one other value besides 2 and 3. This value is greater than 0. What is this value?

(b) What is the variance of X?

3. Consider the random variables X and Y. Both X and Y take the values 0, 1 and 2. The joint probabilities for each pair are given by the following table.

	$X = 0$	$X = 1$	$X = 2$
$Y = 0$	0.1	0.15	0.1
$Y = 1$	0.1	0.1	0.1
$Y = 2$	0.2	0.05	0.1

Compute the $E(U)$ expected value of U, where $U = X - Y$.

- Suppose we have a set of **n** items.
- From that set, we create a subset of **k** items.
- The **order** in which items are selected is recorded. (The ordering of selected items is very important.)
- The total number of **ordered subsets** of **k** items chosen from a set of **n** items is

$$\frac{n!}{n-k!}$$

An ordered sequence of four digits is formed by choosing digits without repetition from the set $\{1, 2, 3, 4, 5, 6, 7\}$.

(i) the total number of such sequences; (780)

(ii) the number of sequences which begin with an odd number; (480) N(A)

(iii) the number of sequences which end with an odd number; (480) (NB)

(iv) the number of sequences which begin and end with an odd number;(240)

(v) the number of sequneces which begin with an odd number or end with an odd number or both; (720)

(vi) the number of sequences which begin with an odd number or end with an odd number but not both. (480)

A college teaches a range of courses including maths, physics and IT. Students choose a range of courses from these three subject areas. Currently 600 students are enrolled of whom 300 study maths courses, 120 study IT and 380 study physics courses.

- 40 students study courses from all three subject areas.
- 200 maths students study physics as well. 60 physics students also study IT and 70 IT students also study maths. 20 students study physics and IT, but not maths.
- How many students study none of these courses at all? (90)
- How many students study maths but not physics or IT? (70)
- How many students study both maths and physics but not IT? (160)
- How many students study courses from precisely two of these subject areas? (210)

Chapter 5

Discrete Probability Distributions

Overview 1) The binomial distribution 2) The Poisson distribution 3)

Section 3 : Probability

How to Compute Probability: Equally Likely Outcomes Sometimes, a statistical experiment can have n possible outcomes, each of which is equally likely. Suppose a subset of r outcomes are classified as "successful" outcomes.

The probability that the experiment results in a successful outcome (S) is:

P(S) = (Number of successful outcomes) / (Total number of equally likely outcomes) = r / n

Consider the following experiment. An urn has 10 marbles. Two marbles are red, three are green, and five are blue. If an experimenter randomly selects 1 marble from the urn, what is the probability that it will be green?

In this experiment, there are 10 equally likely outcomes, three of

which are green marbles. Therefore, the probability of choosing a green marble is 3/10 or 0.30.

- Conditional probability
- Independent events
- Repeated independent events

5.1 Discrete Probability Distributions

- Over the next set of lectures, we are now going to look at two important discrete probability distributions
- The first is the ***binomial*** probability distribution.
- The second is the Poisson probability distribution.
- In `R`, calculations are performed using the `binom` family of functions and `pois` family of functions respectively.
- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (16541705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).

- We will call one of the outcomes the "success" and the other outcome the "failure".
- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n, p)$$

- a probability mass function (pmf) is a function that gives the probability that a discrete random variable is exactly equal to some value.
- The probability mass function is often the primary means of defining a discrete probability distribution

5.2 The Binomial distribution

The binomial distribution is a discrete probability distribution that is applicable as a model for decisionmaking situations in which a sampling process can be assumed to conform to a Bernoulli process. A Bernoulli process is a sampling process in which

(1) Only two mutually exclusive possible outcomes are possible in each trial, or observation. For convenience these are called success and failure.

(2) The outcomes in the series of trials, or observations, constitute independent events.

(3) The probability of success in each trial, denoted by p, remains constant from trial to trial. That is, the process is stationary.

The binomial distribution can be used to determine the probability of obtaining a designated number of successes in a Bernoulli process. Three values are required: the designated number of successes (X); the number of trials, or observations (n); and the probability of success in each trial (p). Where $q = (1-p)$, the formula for determining the probability of a specific number of successes X for a binomial distribution is

Formula

5.3 Poisson Approximation

The Poisson Approximation of the binomial ditribution

Example

$$P(X2) = 1 - (0.134 + 0.27) = 0.596$$

$$P(X = 1) = 2000.010.99199$$

$$P(X = 1) = 0.270$$

Poisson Approximaitions

XBinomial(200, 0.01)

P(X= k)=e-kk!

P(X2) = 1- (0.135+0.27) = 0.595

5.4 The Hypergeometric Distribution

5.4.1 Definition

The following conditions characterize the hypergeometric distribution: The result of each draw (the elements of the population being sampled) can be classified into one of two mutually exclusive categories (e.g. Pass/Fail or Female/Male or Employed/Unemployed). The probability of a success changes on each draw, as each draw decreases the population (sampling without replacement from a finite population).

A random variable X follows the hypergeometric distribution if its probability mass function (pmf) is given by[1]

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}},$$

where

- N is the population size,
- K is the number of success states in the population,
- n is the number of draws,
- k is the number of observed successes,
- $\binom{a}{b}$ is a binomial coefficient.

When sampling is done without replacement of each sampled item taken from a finite population of items, the Bernoulli process does not apply because there is a systematic change in the probability of success as items are removed from the population.

- When sampling without replacement is used in a situation that would otherwise qualify as a Bernoulli process, the hypergeometric distribution is the appropriate discrete probability distribution.
- Given that X is the designated number of successes, N is the total number of items in the population, T is the total number of successes included in the population, and n is the number of items in the sample, the formula for determining hypergeometric probabilities is

Chapter 6

Continuous Probability Distributions

MathsCast 3 : The Uniform Distribution

The Uniform distribution is characterised by two parameters , the minimum and the maximum.

The expected value E(x) is given by

(i.e. the average of the maximum and minimum)

R Code for Graphics

y=c(20,20) x=c(20,100)

plot(x,y,xlim=c(0,120),ylim=c(0,30),pch=13,col='white',axes=FALSE) segments(20,20,100,20,col= 'red') segments(0,0,120,0)

segments(20,0,20,20,col='red') segments(100,0,100,20,col='red')

segments(0,0,0,40)

6.1 Continuous Uniform Distribution

A random variable X is called a continuous uniform random variable over the interval (a, b) if it's probability density function is given by

$$f_X(x) = \frac{1}{b-a} \qquad \text{when } a \leq x \leq b$$

The corresponding cumulative density function is

$$F_x(x) = \frac{x-a}{b-a} \qquad \text{when } a \leq x \leq b$$

The mean of the continuous uniform distribution is

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

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

The Exponential Distribution

The Memoryless property

The most interesting property of the exponential distribution is the ***memoryless*** property. By this , we mean that if the lifetime of a component is exponentially distributed, then an item which has been in use for some time is a good as a brand new item with regards to the likelihood of failure.

The exponential distribution is the only distribution that has this property.

Uniform Distribution: Exercise 24

Use the uniform distribution to simulate 100 throws of two dice. The outcome is the combined values of both dice. Use the appropriate R command to discretize values.

- What is the mean and standard deviation of the outcomes?
- Make a stem-and-leaf plot of the outcomes.
- Make a histogram of the outcomes. (hint: use breaks =seq(1.5,12.5))

6.1.1 Normal Distribution : Example 2

A machine produces components whose thicknesses are normally distributed with a mean of 0.40 cm and a standard deviation of 0.02 cm.

Components are rejected if they have a thickness outside the range 0.38 cm to 0.41 cm.

(i) What is the probability that a component will have a thickness exceeding 0.41 cm? (4 marks)

(ii) What is the probability that a component will have a thickness between 0.38 cm and 0.41 cm? (4 marks)

(iii) What is the thickness below which 25% of the components will be? (4 marks)

6.1.2 Normal Distribution : Example 3

A charity believes that when it puts out an appeal for charitable donations the donations it receives will normally distributed with a mean 50 and standard deviation 6, and it is assumed that donations will be independent of each other.

(i) Find the probability that the first donation it receives will be greater than 40.

(ii) Find the probability that it will be between 55 and 60.

(iii) Find the value x such that 5% of donations are more than x.

Chapter 7

Normal Probability Distribution

MA4004 Revision Class B Please sign the attendance sheet

Today's class : Last years past paper

Normal Distribution Question (Dr David Ramsey's Equine Stats Class)

The mass of Arab horses is normally distributed with mean 900 lbs and standard deviation of 50lbs. Part i Calculate the probability that an Arab horse weighs more than 940 lbs.

Solution

Let X be mass of Arab horses.

We have to find $P(X \leq 940)$. (Remark "equality component" is included as a formality, but it is not important)

Find the Z value that corresponds to 940

$$Zo = \frac{Xo - \mu}{\sigma} = \frac{940 - 900}{50} = 0.8$$

$$P(X \leq 940) = P(Z \leq 0.8)$$

From Murdoch Barnes tables 3, we find that $P(Z \leq 0.8) = 0.2119$

Part ii Calculate the probability than an Arab horse weighs between 880 lbs and 960 lbs. Solution

$$P(880 \leq X \leq 960).$$

What proportion of horses are between 880 lbs and 960 lbs?

- Find out the probability of the complement event.
- The complement event is the combination of being too high or too low for this interval.
- Inside interval $P(880 \leq X \leq 960)$.
- Outside interval $P(X \leq 880) + P(X \geq 960)$
- Complement Rule $P(880 \leq X \leq 960) = 1 - [P(X \leq 880) + P(X \geq 960)]$

Find the probability of being too high?

$$Zo = \frac{Xo - \mu}{\sigma} = \frac{960 - 900}{50} = 1.2$$

$$P(X \leq 960) = P(Z \leq 1.2) = 0.1151$$

Find the probability of being too low? Zo=Xo-=880 -90050= -0.4 P(X880) = P(Z-0.4)

How to compute P(Z-0.4)

Symmetry: P(Z-0.4) = P(Z0.4) = 0.3446

- Outside Interval = 0.4596 (0.3446 + 0.1151)
- Inside Interval = 0.5404

What weight is exceeded by 97.5
Find Xo such that P(XXo) = 0.975
P(Z1.96) = 0.025 [From Tables]
P(Z-1.96) = 0.025 [Symmetry]
P(Z-1.96) = 0.975
-1.96 =Xo- 90050
Xo= 802 lbs [Answer]
Important:
Formulae at back of Exam Paper Murdoch Barnes Table 3 (The Z distribution) Murdoch Barnes Table 7 (The Student t distribution)
Important Considerations
Significance / confidence 1-
Number of tails one tailed procedure or two tailed procedure Confidence intervals are always two tailed
Sample size (degrees of freedom depends on sample size)
Question 3 - Paired T test
a) The weights of one group of Irish students were recorded both at the beginning of year 1 of their studies and at the end of year 4. The results (in kg) are given below:
Student 1 2 3 4 5 6 7 8 Year 1 72 58 68 81 65 69 75 84 Year 4 74 61 69 83 69 74 76 82

At a significance level of 5%, is there sufficient evidence to state that on average students gain weight over the four years of their university studies?

Solution

Before we start, we need to compute the average difference and the standard deviations of the differences.

Student 1 2 3 4 5 6 7 8 Diff 2 3 1 2 4 5 1 -2 Di-D 0 1 -1 0 2 3 -1 -4

Computing the average difference

Now we compute the standard deviation of the variances

From each difference value, subtract the mean, and square the resulting term.

Di-D 0 1 -1 0 2 3 -1 -4 (Di-D)2 0 1 1 0 4 9 1 16

= 4.571

Standard deviation is the square root of the variance

SD=4.571=2.137

hypotheses

H0:D0 Students have not gained weight through college

Ha:D¿ 0 Students have gained weight through college

N.B. This is a one-tailed test.

Chapter 8

Inference Procedures

Test Statistic

Remember the general structure of a test statistic

$$TS = \frac{\text{Observed Value-Null Value}}{\text{Std. Error}}$$

Standard Error S.E.(D) =SDn=2.1378= 0.7555
Test statistis is a t random variable
Test Statistic x-0S.E.(D)=2 - 00.755= 2.649

Critical values

- The sample size (n=8) is small ($n \leq 30$). Use t distribution with n-1 degrees of Freedom.
- The test is one tailed. k= 1 (why? "¿" symbol in the alternative hypothesis).

- Murdoch Barnes table 7
- Column: $\alpha/k = 0.05/1 = 0.05$
- Row: df $= 7$
- Critical value $= 1.895$

Decision rule

Is the absolute value of the test statistic value greater than the critical value?

- If Yes: we reject the null hypothesis
- If No: We fail to reject the null hypothesis. (not enough evidence)

Here TS $= 2.64$ is greater than CV $= 1.895$.

Decision rule

We reject the null hypothesis. Students do put on weight during college.

Question 3 Part b : Confidence interval for the difference in means of two samples.

b) The mean and standard deviation of the weights of a sample of Irish students according to sex are given below

Number Mean Std. Dev. Male 100 75 10 Female 110 66 8

i) Calculate a 99students and the mean weight of all female students.(7 males)

General Structure of a Confidence Interval

Observed difference

- let X denote the weights of male students X= 75
- let Y denote the weights of female students Y= 66
- The difference in the mean of weights X-Y= 9

Quantile

Large sample (both groups are greater than 30).

Population variance is unknown. Use t distribution with degrees of Freedom.

Confidence level is 99Confidence intervals are always two tailed procedures.

Column = k =0.012= 0.005

Murdoch Barnes table 7

Row: df = Column = 0.005

Quantile = 2.576

Stardard Error

Confidence Interval is therefore

99

Part (ii) Based on this confidence interval, test the hypothesis that on average male students are 6kgs heavier than female students. State your hypotheses clearly. What is the significance level of this test? (3 marks)

H0:X-y= 6

Since we do not reject the null hypothesis at a significance level of 1

Question 4

a) The mean and standard deviation of the salaries of 16 Irish full-time workers are 5000 and 3000, respectively.

i) Test the hypothesis that the mean salary of all Irish full-time workers is 4000 at a significance level of 5

Step 1 : Formally state the null and alternative hypotheses

Step 2 : Determine the test statistic

Step 3 : Determine the critical value

Step 4 : Decision Rule

Given

Observed value : Sample mean x= 5000 Null Value (Expected mean under null hypothesis) o= 4000 Population standard deviation is unknown. Use sample standard deviation s = 3000 as an estimate. Significance = 0.05 Sample size n = 16 (small sample)

hypotheses

H0:=4000 True mean salary is 4000 Ha:4000 True mean salary is not 4000

Test statistic

Remember the general structure of a test statistic

TS =Observed Value-Null ValueStd. Error

Standard Error S.E.(x) =sn=300016= 750

Test Statistic x-0S.E.(X)=5000 -4 000750= 1.33

Critical values

The sample size (n=16) is small (n30). Use t distribution with n-1 degrees of Freedom.

The test is two tailed. k= 2 ("" symbol in the alternative hypothesis).

Column = k =0.052= 0.025

Murdoch Barnes table 7

Row: df = 15 (n-1)

Column = 0.025

Critical value = 2.131

Decision rule

Is the absolute value of the test statistic value greater than the critical value?

- If Yes: we reject the null hypothesis

- If No: We fail to reject the null hypothesis. (not enough evidence)

If No: We fail to reject the null hypothesis. (not enough evidence)

Here TS = 1.33 is not greater than CV = 2.131.

We fail to reject the null hypothesis. We can not rightly say that the mean salary is not 4000 per month.

ii) What assumption is made in this testing procedure? Is this assumption reasonable? (2 marks)

The assumption made in this testing procedure is that salaries are normally distributed. This is not a valid assumption as the distribution of salaries is known to be skewed (lots of low values, few high values).

b) A survey of 1000 Irish indicates that 750 have access to the Internet. A survey of 2000 Spaniards indicates that 1400 have access to the Internet.

i) By calculating the appropriate p-value, test the hypothesis that the proportion of all Irish having access to the Internet is equal to the proportion of all Spaniards having access to the internet at a significance level of 5

Step 1 : Formally state the null and alternative hypotheses

Step 2 : Determine the test statistic Step 3a : Determine the p.value Step 4a : Decision Rule for p-values.

Step 1 : Formally state the null and alternative hypotheses

Proportion of people having internet access is the same in both Ire-

land and Spain Proportion of people having internet access differs in Ireland and Spain

Alternatively we write the hypotheses as follows (the null value is more evident).

Step 2 Compute the Test Statistic

$$p_{Irl} = \frac{750}{1000} = 0.75$$

$$p_{Esp} = \frac{1400}{2000} = 0.70$$

Observed Difference = 0.75 - 0.70 = 0.05

Now lets compute the standard error (from Formulae)

ii) Calculate a 99Internet and the proportion of all Spaniards having access to the internet. (4 marks)]

Standard Error for confidence interval p1(1 -p1)n1+p2(1 -p2)n2 =0.750.25100 = 0.017103

Quantile for a 99

significance level =1number of tails = 2 degrees of freedom = quantile = 2.576

99

Useful pieces of information

Sample size n=100

Part i Calculate the equation of the least square regression line and interpret the value of the slope.

part ii

Using this regression model, estimate the mean weight of individuals who are 3 metres tall. part iii

Is such an estimate reliable? briefly explain why.

No it is not reliable. Consider the range of values of the x predcitro variable

8.1 Normal Distribution: Worked Examples

Q5. a) Assume that the amount of wine poured into a bottle has a normal distribution with a mean of 750ml and a variance of 144ml^2.

(i) Calculate the probability that a bottle contains more than 765ml. (2 marks)

(ii) Calculate the probability that a bottle contains between 744ml and 759ml. (3 marks)

A machine fills bags with animal feed. The nominal weight of a bag is 50kg. Because random variations the weight of a filled bag is normally distributed $N(\mu, \sigma^2)$. The variance (σ^2) is known to be 0.01kg^2 and μ is set by the operator to a particular value.

(i) If ? = 50kg calculate the probability of a bag containing less than 49.95kg?

(ii) Calculate the value of "?" such that only 2% of the output are under the nominal weight?

MA4004 Engineering Statistics SPRING 2008

The amount of beer in a bottle has a normal distribution with mean 500ml and variance 25ml2.

(i) Calculate the probability that the amount of beer in the bottle is between 498ml and 504ml.

(ii) What volume is exceeded by 20% of the bottles? (6 marks)

A certain farm produces two kinds of eggs on any given day; organic and non-organic. Let these two kinds of eggs be represented by the random variables X and Y respectively. Given that the joint probability density function of these variables is given by

a) Find the marginal PDF of X

b) Find the marginal PDF of Y

c) Find the $P(X = 1/2, Y = 1/2)$

Consider the AR(2) model

$$Y_t = \frac{1}{3}Y_{t-1} + \frac{1}{12}Y_{t-2} + \epsilon_t$$

for a process Y_t, where ϵ_t is a white noise process.$(-\infty \leq t \leq \infty)$

(i) Find the roots of the autoregressive characteristic equation and check that the stationarity condition is satisfied.

(ii) Find the Yule-Walker equations that are satisfied by the autocorrelation function t.

(iii) Obtain the value of ρ_1.

(iv) Show that a general expression for the autocorrelation function is given by

$$\rho_t = \frac{35}{44}\frac{1}{12}^{t} + \frac{9}{44}\frac{-1}{6}^{t}$$

where $\tau \geq 0$ (9)