Project Maths

SUPPLEMENT

Text & Tests 3

Leaving Certificate Ordinary Level Maths

Contains all the Deferred Material of Strand 1

O. D. Morris · Paul Cooke · Paul Behan

The Celtic Press

First published June 2014 by
The Celtic Press
Ground Floor – Block B
Liffey Valley Office Campus
Dublin 22

ISBN: 978-0-7144-2017-2

Printed in Ireland by
Turner Print Group
Earl Street
Longford

Preface

This booklet contains all the **Deferred Material** from Strand 1 of the Leaving Certificate Ordinary Level Course. This material was introduced in September 2013 for examination in June 2015 and onwards.

This chapter is printed as a supplement for those students who bought *Text & Tests 3* at the beginning of 5th Year in September 2013. All future editions of *Text & Tests 3* will contain this new chapter at the end of the book.

O. D. Morris
Paul Cooke
Paul Behan

June 2014

Inferential Statistics

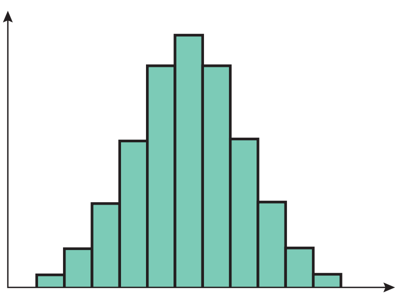
chapter 19

Key words

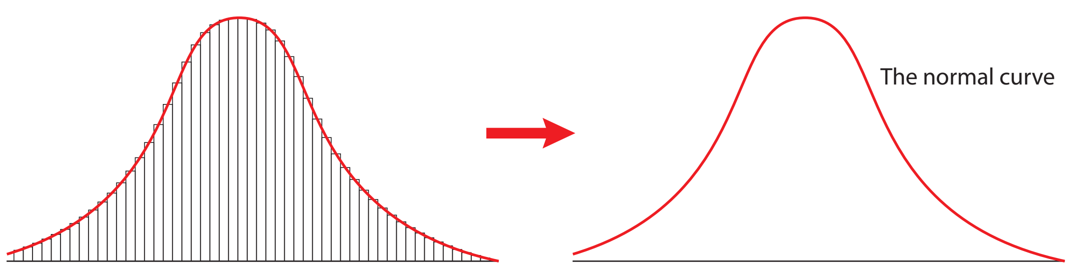
normal distribution **normal curve** **the Empirical Rule** **standard deviation** **margin of error** **sample** **population** **sample proportion** **population proportion** **confidence interval** **hypothesis test** **null hypothesis** **alternative hypothesis** **reject/accept**

Section 19.1 The Normal Distribution and The Empirical Rule

When the physical characteristics, such as height or weight, of a large number of individuals are arranged in order, from lowest to highest, in a frequency distribution, the same pattern shows up repeatedly. This pattern shows that a large number of values cluster near the middle of the distribution, as illustrated by the symmetrical histogram shown below.



If the distribution is very large and continuous, and the class intervals become sufficiently small, the distribution forms a symmetrical bell-shaped smooth curve called **the curve of normal distribution** or simply **the normal curve**, as shown.



In Section 8.6 of this book, we dealt with a measure of spread called **standard deviation**, σ, which gives an indication of the distance the data is from the mean.

There is a very important relationship between the normal curve and standard deviations. It is called **the Empirical Rule** and it is given below:

The Empirical Rule

For any large population with mean $\bar{x}$ and standard deviation σ

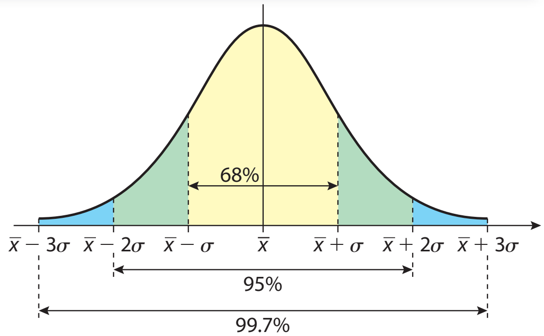
(i) about 68% of the values will lie within one standard deviation of the mean, that is, between $\bar{x} - \sigma$ and $\bar{x} + \sigma$

(ii) about 95% of the values will lie within two standard deviations of the mean, that is, between $\bar{x} - 2\sigma$ and $\bar{x} + 2\sigma$

(iii) almost all (99.7%) of the values will lie within three standard deviations of the mean, that is, $\bar{x} - 3\sigma$ and $\bar{x} + 3\sigma$

The normal curve on the right illustrates the Empirical Rule.

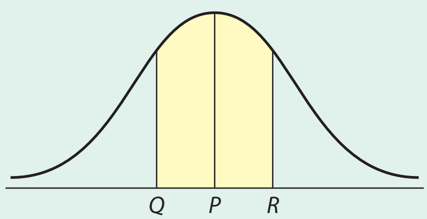
Based on the Empirical Rule, the probability that a score, selected at random, will be within one standard deviation of the mean is 68% or 0.68.



Example 1

The given curve represents a normal distribution with mean $\bar{x} = 80$ and standard deviation $\sigma = 10$.

P represents the mean, *Q* represents a value one standard deviation below *P* and *R* represents a value one standard deviation above *P*.



(i) Write down the values of *P*, *Q* and *R*.

(ii) What percentage of the data lies in the shaded area?

(iii) If a value is chosen at random from all the data, what is the probability that it comes from the shaded area?

(i) *P* represents the mean; so $P = 80$

$Q = 80 - \sigma$ $\qquad$ $R = 80 + \sigma$

$= 80 - 10 \ldots \sigma = 10$ $\qquad$ $= 80 + 10$

$Q = 70$ $\qquad$ $R = 90$

(ii) The shaded area represents all the values that lie within one standard deviation of the mean. Accordingly, the shaded area contains 68% of the data.

(iii) The probability that a value comes from the shaded area $= 68\% = 0.68$.

Example 2

A normal distribution has mean $\bar{x} = 45$ and standard deviation $\sigma = 5$.

(i) Find the range within which 68% of the distribution lies.

(ii) Find the range within which 95% of the distribution lies.

(iii) What percentage of the distribution lies within 3 standard deviations of the mean?

(i) 68% of the distribution lies in the range $\bar{x} - \sigma,\ \bar{x} + \sigma$.

$\bar{x} - \sigma = 45 - 5 = 40$ and $\bar{x} + \sigma = 45 + 5 = 50$

$\therefore$ 68% lies in the range [40, 50]

(ii) 95% of the distribution lies in the range $\bar{x} - 2\sigma,\ \bar{x} + 2\sigma$.

$\bar{x} - 2\sigma = 45 - 2(5) = 45 - 10 = 35$

$\bar{x} + 2\sigma = 45 + 2(5) = 45 + 10 = 55$

$\therefore$ 95% lies in the range [35, 55]

(iii) 99.7% of the distribution lies in the range $[\bar{x} - 3\sigma, \bar{x} + 3\sigma]$.

Example 3

The mean $\bar{x}$ of a normal distribution is 84.
If 95% of the values are between 72 and 96, find the value of σ, the standard deviation.

By the Empirical Rule, 95% of the values lie in the range $[\bar{x} - 2\sigma, \bar{x} + 2\sigma]$

$\Rightarrow\ [\bar{x} - 2\sigma, \bar{x} + 2\sigma] = [72, 96]$

$\Rightarrow\ \bar{x} - 2\sigma = 72$

$84 - 2\sigma = 72$

$-2\sigma = -12 \quad \Rightarrow \quad 2\sigma = 12 \quad \Rightarrow \quad \sigma = 6$

The standard deviation is 6.

Exercise 19.1

1. Copy and complete these sentences:

(i) Approximately ___% of the data lies within one standard _____________ of the mean.

(ii) Approximately 95% of the data lies within _______ standard deviations of the ___________.

2. The mean $\overline{x}$ of a distribution is 54 and the standard deviation σ is 3.

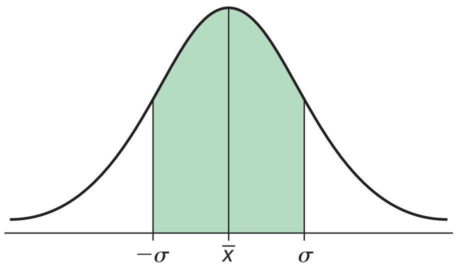
(i) Find the range $[\overline{x} - \sigma, \overline{x} + \sigma]$.

(ii) Find the range within which 95% of the data lies.

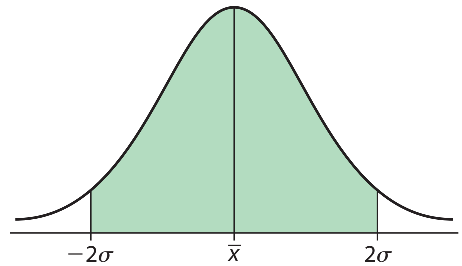
(iii) If a value is selected at random from the distribution, find the probability that it is in the range $[\overline{x} - \sigma, \overline{x} + \sigma]$.

3. For each of the following normal curves, find the percentage of all the values that are in the shaded area:

(i)



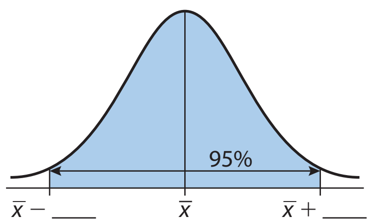
(ii)



4. (i) In the given normal curve, complete the labels on the horizontal axis.

(ii) If a value is selected at random, find the probability that it is in the shaded region.

(iii) If $\overline{x} = 84$ and $\sigma = 7$, what are the values of the incomplete labels in the diagram?



5. In a normal distribution, the mean $\overline{x} = 120$ and the standard deviation $\sigma = 15$.

(i) Find the range within which 68% of the values lie.

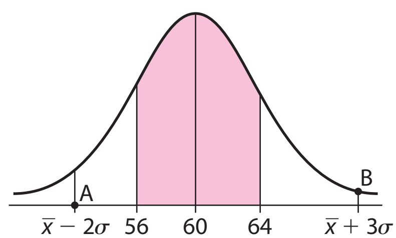
(ii) Find the range within which 95% of the values lie.

6. The standard deviation σ of the given normal curve is 4.

(i) Write down the value of the mean $\overline{x}$.

(ii) What percentage of the values lie in the shaded area?

(iii) Write down the values of A and of B.



7. The mean speed of vehicles on a given road can be modelled by a normal distribution with mean 55 km/h and standard deviation 9 km/h.
What would be the speed of a vehicle that was travelling at

(i) one standard deviation below the mean

(ii) two standard deviations above the mean

(iii) three standard deviation above the mean?

8. A normal distribution has a mean $\bar{x} = 60$ and standard deviation $\sigma = 5$.

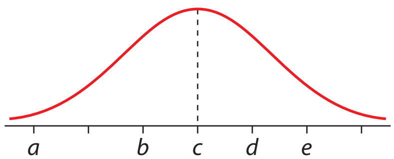
(i) Find the range within which 68% of the distribution lie.

(ii) Find the range within which 95% of the distribution lie.

9. The heights of students in a certain university are normally distributed with a mean of 180 cm and a standard deviation of 10 cm.

The given diagram represents this distribution and each mark on the horizontal line represents 1, 2 or 3 standard deviations from the mean.

Write down the values of a, b, c, d and e.



10. The heights of a large sample of adults are normally distributed with a mean of 170 cm and a standard deviation of 8 cm.

Within what limits do

(i) 68% of the heights lie

(ii) 99.7% of the heights lie?

11. The distribution of the times taken by factory workers to get to their place of work can be modelled by a normal distribution.

The mean time is 35 minutes and the standard deviation is 6 minutes.

(i) What percentage of the workers take longer than 35 minutes?

(ii) What percentage of the workers take between 29 minutes and 41 minutes?

(iii) What is the range of the times it takes 95% of the workers to get to their place of work?

12. The lengths of time taken to complete a crossword puzzle are normally distributed with a mean $\bar{x} = 18$ minutes.

If 68% take between 14 minutes and 22 minutes, find the standard deviation, σ, of the distribution.

13. The marks awarded in an examination are normally distributed with mean 74 marks.

If 95% of the marks are in the range [62 marks, 86 marks], work out the standard deviation, σ.

14. The weights of a group of 1000 schoolchildren were normally distributed with a mean of 42 kg and a standard deviation of σ.

950 of the children were in the range 30 kg to 54 kg.

(i) Express 950 as a percentage of 1000.

(ii) Complete the following sentence:

"95% of the children lie in the range $[\bar{x} - \_\_, \ \bar{x} + \_\_]$".

(iii) Use the information given to find the value of the standard deviation σ.

(iv) If a child is selected at random, find the probability that the child's weight is in the range 30 kg to 54 kg.

Section 19.2 Margin of error – Confidence intervals

When dealing with sampling in your earlier study of statistics, it was stated that the purpose of sampling is to gain information about the whole population by surveying a small part of the population. This small part is called a **sample**.

If data from a sample is collected in a proper way, then the sample survey can give a fairly accurate indication of the population characteristic that is being studied.

Before a General Election, a national newspaper generally requests a market research company to ask a sample of the electorate how they intend to vote in the election.
This survey is generally referred to as an opinion poll.
The number is usually about 1000.
If the number of people surveyed was increased to 2000, you would get a more accurate picture of voting intentions.

The result of the survey might appear in the daily newspaper as follows:

40% support for **The Democratic Right**.

The **40%** support is called the **sample proportion**, that is, the part or portion of the sample who indicated that they would vote for *The Democratic Right*.
The party could then expect to get somewhere around 40% of all voters in the general election.

The notation $\hat{p}$ is used to denote **sample proportion**.
The notation p is used to represent **population proportion**.
Since p is generally not known, $\hat{p}$ is used as an **estimator** for the true population proportion, p.

Of course everybody knows that sample surveys are rarely 100% accurate.
There is generally some 'element of chance' or **error** involved.

The newspaper might add to their headline the following sentence:

The margin of error is 3%.

The **margin of error** of 3% is a way of saying that the result of the survey is 40% $\pm$ 3%.
That means that the research company is quite 'confident' that the proportion of the whole electorate who intend to vote for *The Democratic Right* could be anywhere between 37% and 43%.

How does the research company calculate **the margin of error**?

The margin of error, E, in opinion polls is generally calculated using the formula,

$$E = \frac{1}{\sqrt{n}}, \text{ where } n \text{ is the sample size.}$$

Margin of error

$$E = \frac{1}{\sqrt{n}}$$

If the sample size is 400, then $E = \frac{1}{\sqrt{400}} = \frac{1}{20} = 0.05$.

If the sample size is 1000, then $E = \frac{1}{\sqrt{1000}} \approx 0.03$.

If the sample size is 4000, then $E = \frac{1}{\sqrt{4000}} \approx 0.016$.

Notice that the error decreases as the sample size increases.

Confidence interval

The result of the opinion poll above was given as 40% $\pm$ 3%.

$$40\% - 3\% = 37\% \quad \text{and} \quad 40\% + 3\% = 43\%.$$

So, the result of the opinion poll is anywhere in the interval 37% to 43%.
More formally, it can be written as $37\% < p < 43\%$, where p is the population proportion.
$37\% < p < 43\%$ is called a **confidence interval**.
But how confident are we that the result will lie in this interval?
There are many levels of confidence, but for our course the **confidence level** is pitched at **95%**.

The 95% confidence level implies that the interval was obtained by a method which 'works 95% of the time'.
The confidence interval, $37\% < p < 43\%$, is a way of stating that if you surveyed numerous samples of 1000 people on the same day, the results would be in the interval 37% to 43% on 95 occasions out of 100, that is, in 95% of the samples.

The 95% confidence interval for a population proportion is given on the right.

The confidence interval on the right may also be expressed as $\boldsymbol{\hat{p} \pm \frac{1}{\sqrt{n}}}$.

Confidence interval is

$$\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$$

Example 1

A random sample of 400 persons are given a flu vaccine and 136 of them experienced some discomfort.

(i) Write down the sample size.

(ii) Calculate the margin of error using $\frac{1}{\sqrt{n}}$.

(iii) Find the sample proportion.

(iv) Find the confidence interval.

(i) The sample size is 400.

(ii) The margin of error is $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{400}} = \frac{1}{20} = 0.05$.

(iii) The sample proportion, $\hat{p} = \frac{136}{400} = 0.34$.

(iv) The confidence interval is $\hat{p} \pm \frac{1}{\sqrt{n}}$

$$\hat{p} \pm \frac{1}{\sqrt{n}} = 0.34 \pm 0.05 \ \ldots \text{ from (ii) and (iii) above}$$
$$= 0.34 - 0.05 \quad \text{to} \quad 0.34 + 0.05$$
$$= 0.29 \text{ to } 0.39$$

The confidence interval may be written as $0.29 < p < 0.39$.

Example 2

What sample size would be required to have a margin of error of

(i) 0.05 (ii) $2\frac{1}{2}\%$?

(i) $\frac{1}{\sqrt{n}} = 0.05$

$\therefore \frac{1}{n} = (0.05)^2 \ \ldots$ square both sides

$\therefore n = \frac{1}{(0.05)^2}$

$n = 400$

(ii) $\frac{1}{\sqrt{n}} = 2\frac{1}{2}\% = 0.025$

$\frac{1}{n} = (0.025)^2 \ \ldots$ square both sides

$n = \frac{1}{(0.025)^2}$

$n = 1600$

Exercise 19.2

1. Work out the margin of error for each of the following random samples at the 95% confidence level:

(i) 900 (ii) 1200 (iii) 2025 (iv) 800

Give your answer correct to 2 decimal places where necessary.

2. In a random sample of 500 households, 80 said that they had at least one pet.

(i) What is the sample size?

(ii) What is the margin of error?

(iii) What is the sample proportion?

3. In a random sample of 200 students, 48 said that they spend at least one hour each day watching television.
   (i) Write down the sample size.
   (ii) What is the margin of error?
   (iii) What is the sample proportion, $\hat{p}$?
   (iv) If you increase the sample size to 400, what effect would this have on the margin of error?

4. A manufacturer tests a random sample of 300 items and finds that 45 are defective.
   (i) Write down the sample size, n.
   (ii) Calculate the margin of error.
   (iii) Work out the sample proportion, $\hat{p}$.
   (iv) Using $\hat{p} \pm \frac{1}{\sqrt{n}}$, work out a confidence interval for the proportion of defective items produced.

5. A survey was undertaken to find the level of use of the internet by residents of a city. In a random sample of 150 residents, 45 said that they log onto the internet at least once a day.
   (i) Write down the sample size.
   (ii) Calculate the margin of error.
   (iii) Work out the sample proportion.
   (iv) Construct a confidence interval for p, the population proportion that log onto the internet at least once a day.

6. In a random sample of 400 computer shops, it was discovered that 168 of them sold computers at below the list price recommended by the manufacturer.
   (i) Write down the sample size.
   (ii) Calculate the margin of error.
   (iii) Work out the sample proportion.
   (iv) Construct an approximate 95% confidence interval for the proportion of shops selling below the list price.

7. Copy and complete this sentence:
"A 95% confidence level means that on ______ occasions out of ______, the proportion will be in this _____".

8. A college principal decides to consult the students about a proposed change to the times of lectures. She finds that, out of a random sample of 80 students, 36 of them are in favour of the change.
   (i) Calculate the margin of error.
   (ii) Work out the sample proportion, $\hat{p}$.
   (iii) Construct a confidence interval for the proportion of students who are in favour of change.

9. A survey was carried out in order to gauge the response to a new school "healthy eating" menu. A random sample of 200 schoolchildren was selected from different schools. It was found that 84 children approved of the new menu.

(i) Write down the sample size.
(ii) Calculate the margin of error, correct to 2 decimal places.
(iii) Work out the sample proportion, $\hat{p}$.
(iv) Establish a 95% confidence interval for the proportion of all students who approved of the new menu.

10. Write each of these percentages as a decimal:

(i) 10% (ii) 5% (iii) 3% (iv) 2.5% (v) 6.5%

11. What size sample, n, is required to have a margin of error of 6%.
The first two lines are done for you:

$$\frac{1}{\sqrt{n}} = \pm 0.06 \ldots \text{ 6\% = 0.06}$$

$$\left(\frac{1}{\sqrt{n}}\right)^2 = (\pm 0.06)^2$$

12. What size sample is required to have a margin of error of

(i) 5% (ii) 4% (iii) 12% (iv) 3.5% (v) 2.5%?

13. The results of a survey showed that 360 out of 1000 families regularly purchase the *Daily Bulletin* newspaper.

(i) Write down the sample size.
(ii) Calculate the margin of error, correct to two places of decimal.
(iii) For this survey, what is the sample proportion, $\hat{p}$?
(iv) Construct a confidence interval for the proportion of families that purchase the *Daily Bulletin*.

Section 19.3 Hypothesis testing

Many people believe that the average or mean height of Irish people is greater than the mean height of Spanish people.
To illustrate this belief, we could make the following statement:

"The mean height of Irish people is greater than the mean height of Spanish people".

In mathematics, this statement is called an **hypothesis**.
Having made this statement, we now need some evidence to show the truth, or otherwise, of the statement.
The process of proving the truth of the statement is called **hypothesis testing**.

The assumption or statement made is called the **null hypothesis**.
The null hypothesis is denoted by $\mathbf{H_0}$.
Usually, the null hypothesis is a statement of "no change", "no difference", or "no effect".

Suppose a drug company claims that a "new improved drug" is more effective that the existing drug, then the null hypothesis could be worded as follows:

H_0: There is no change in the effectiveness of the new drug.

The drug company would not like this hypothesis and so would put forward an **alternative hypothesis**, $\mathbf{H_1}$.
This alternative hypothesis might read as follows:

H_1: There is a change in the effectiveness of the new drug.

We generally concentrate on the null hypothesis and carry out a **hypothesis test** to accept or reject the null hypothesis.

Here are some examples of null hypotheses and their corresponding alternative hypotheses.

H_0: The mean weight of adult males is 78 kg.
H_1: The mean weight of adults males is not 78 kg.

H_0: A football team is more likely to concede a goal just after it has scored a goal.
H_1: A football team is not more likely to concede a goal just after it has scored a goal.

H_0: The average marriage-age of men is higher than the average marriage-age of women.
H_1: The average marriage-age of men is not higher than the average marriage-age of women.

In Section 19.2 the results of a newspaper survey stated:

"40% support for *The Democratic Right*.

The margin of error was $\pm$3%.

40% $\pm$ 3% gives an **interval** 37% to 43%.

What this interval means is that in 95 samples out of 100, *The Democratic Right* would receive a percentage vote somewhere between 37% and 43%.
95 samples out of 100 is called the **95% confidence level**.

How do we prove the truth, or otherwise, of the newspaper's claim that the true percentage of people who intend to vote for *The Democratic Right* lies in the interval 37% to 43%.

To test the truth of the newspaper statement, or hypothesis, we carry out an **hypothesis test**.
To do this test, we select another sample of the same size and calculate the percentage of this sample who say that they will vote for *The Democratic Right*.

If the percentage of voters who intend to vote for the named party lies outside the interval 37% to 43%, we reject the null hypothesis; that is, we reject the claim made by the newspaper. If the percentage is within the interval 37% to 43%, we accept the newspaper's claim.
The steps involved in carrying out an hypothesis test are given on the next page.

Procedure for carrying out an hypothesis test

1. Write down $\mathbf{H_0}$, the **null hypothesis**, and $\mathbf{H_1}$, the **alternative hypothesis**
For example, to test whether a coin is biased if we get 7 heads in 10 tosses, we could formulate the following hypotheses:
H_0: The coin is not biased.
H_1: The coin is biased.

2. Write down or calculate the sample proportion, $\hat{p}$.

3. Find the margin of error, $\frac{1}{\sqrt{n}}$.

4. Write down the confidence interval for the population proportion, $\hat{p}$, using

$$\hat{p}-\frac{1}{\sqrt{n}} < p < \hat{p}+\frac{1}{\sqrt{n}} \qquad \left[\text{or } \hat{p} \pm \frac{1}{\sqrt{n}}\right]$$

5. (i) If the value of the population proportion stated is within the confidence interval, accept the null hypothesis H_0 and reject H_1.
(ii) If the value of the population proportion is outside the confidence interval, reject the null hypothesis H_0 and accept H_1.

Example 1

The *Kennell Club* claims that 30% of households keep a dog.
To test this claim, a group of statistics students carry out a survey of 400 households.
The results of the survey found that 112 of the 400 households kept a dog.
(i) Write down the null hypothesis.
(ii) Write down the alternative hypothesis.
(iii) Write down the sample size, n.
(iv) Work out the margin or error, $\frac{1}{\sqrt{n}}$.
(v) Calculate the sample proportion, $\hat{p}$.
(vi) Calculate the confidence interval using $\hat{p}-\frac{1}{\sqrt{n}} < p < \hat{p}+\frac{1}{\sqrt{n}}$.
(vii) At the 95% confidence level, can we accept the *Kennell Club's* claim?

(i) H_0: 30% of households keep a dog.
(ii) H_1: The percentage of households who keep a dog is not 30%.
(iii) The sample size, n, is 400.
(iv) Margin of error is $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{400}} = \frac{1}{20} = 0.05$.
(v) Sample proportion, $\hat{p} = \frac{112}{400} = 0.28$

(vi) Confidence interval is $\hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$= 0.28 - 0.05 < p < 0.28 + 0.05$$
$$= 0.23 < p < 0.33$$
$$= 23\% < p < 33\%$$

(vii) Since the claim, i.e. 30%, is within the confidence interval $23\% < p < 33\%$, we accept the null hypothesis.

Example 2

A school principal claimed that 45% of students were in favour of a change to the school uniform.
The students' council carried out a survey of 120 students to test the principal's claim.
The results of the survey found that 42 students were in favour of a change to the school uniform.

(i) Write down the null hypothesis.
(ii) Write down the alternative hypothesis.
(iii) Calculate the margin of error, correct to two decimal places.
(iv) Work out the sample proportion, $\hat{p}$.
(v) Calculate the confidence interval for the population proportion.
(vi) At the 95% confidence level, is the principal's claim correct?

(i) H_0: 45% of students are in favour of a change to the school uniform.

(ii) H_1: The percentage of students in favour of a change is not 45%.

(iii) Margin of error is $\frac{1}{\sqrt{n}} = \frac{1}{\sqrt{120}} = 0.09$ … correct to 2 decimal places.

(iv) Sample proportion, $\hat{p} = \frac{42}{120} = 0.35$

(v) Confidence interval $= \hat{p} - \frac{1}{\sqrt{n}} < p < \hat{p} + \frac{1}{\sqrt{n}}$

$$= 0.35 - 0.09 < p < 0.35 + 0.09$$
$$= 0.26 < p < 0.44$$
$$= 26\% < p < 44\%$$

(vi) Since the principal's claim, i.e. 45%, lies **outside** the confidence interval $26\% < p < 44\%$, we **reject** the null hypothesis H_0 and accept the alternative hypothesis H_1.

Exercise 19.3

1. A school principal claims that 95% of students are on time for the first class at 9.00 a.m.
   (i) Write down H_0, the null hypothesis when carrying out an hypothesis test.
   (ii) Write down an alternative hypothesis, H_1.
   (iii) Describe the next steps you would take in testing the principal's claim.

2. The manufacturer of *Chummy Bits* claims that 80% of dog owners choose this product for their dogs.
   In a random sample of 200 dog owners, 155 chose *Chummy Bits*.
   (i) Write down the null hypothesis, H_0.
   (ii) Write down the alternative hypothesis, H_1.
   (iii) What is the sample size, n?
   (iv) Calculate the margin of error, $\frac{1}{\sqrt{n}}$.
   (v) Calculate the sample proportion.
   (vi) Work out the confidence interval.
   (vii) At the 95% confidence level, is the manufacturer's claim correct?

3. A company states that 20% of the visitors to its website purchase at least one of its products. A sample of 400 people who visited the site is checked and the number who purchased a product is found to be 64.
   (i) Write down the null hypothesis, H_0.
   (ii) What is the sample size?
   (iii) Calculate the margin of error.
   (iv) Work out the sample proportion.
   (v) Work out the confidence interval.
   (vi) At the 95% confidence level, is the company's statement correct?

4. In a pubic opinion poll, 1000 randomly-chosen voters were asked whether they would vote for the *Purple Party* at the next election and 350 replied "Yes".
   The leader of the *Purple Party* believes that the true proportion is 40%.
   The random sample was used to test the party leader's belief.
   (i) Write down the null hypothesis, H_0.
   (ii) Write down the sample size.
   (iii) Find the margin of error, correct to two decimal places.
   (iv) Calculate the sample proportion.
   (v) Work out the confidence interval for the percentage of voters who say that they will vote for the *Purple Party*.
   (vi) Is the leader's belief justified at the 95% level of confidence?

5. A college claims that it admits equal numbers of men and women.
   The Student's Union took a random sample of 500 students to test the college's claim.
   It found that 270 were men.

(i) What percentage of those admitted did the college claim were men?
(ii) Write down the null hypothesis.
(iii) Write down the sample size.
(iv) Calculate the margin of error, correct to 2 decimal places.
(v) Work out sample proportion, $\hat{p}$.
(vi) Work out the confidence interval for the percentage of men admitted by the college.
(vii) Is the college's claim justified at the 95% level of confidence?

6. A seed company sells pansy seeds in mixed packets and claims that 20% of the resulting plants will have red flowers. A packet of seeds is sown by a gardener who finds that only 16 out of 100 plants have red flowers.
(i) Write down the null hypothesis, H_0.
(ii) Calculate the margin of error.
(iii) What is the sample proportion of plants with red flowers?
(iv) Work out a confidence interval for the percentage of seeds that will go on to have red flowers.
(v) Is the percentage claimed by the company within this confidence limit?
(vi) Based on your answer to part (v), do we reject the null hypothesis?

7. A drugs company produced a new pain-relieving drug for migraine sufferers and claimed that the drug had a 90% success rate.
A group of doctors doubted the company's claim.
They prescribed the drug for a group of 150 patients.
After six months, 120 of these patients said that their migraine symptoms had been relieved.
(i) Write down the null hypothesis, H_0.
(ii) Write down H_1.
(iii) Write down the sample size.
(iv) Calculate the sample proportion.
(v) What is the margin of error, correct to 2 decimal places?
(vi) Work out a confidence interval for the percentage of patients who said their symptoms had been relieved.
(vii) At the 95% level of confidence, can the company's claim be upheld?

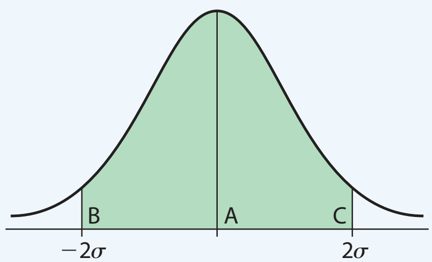
8. A university library claimed that 12% of returned books were overdue.
The college President wanted to test this claim.
A random sample of 200 returned books revealed that only 15 were overdue.
(i) Write down the null hypothesis, H_0.
(ii) Calculate the margin of error, correct to 2 decimal places.
(iii) Write down the sample proportion.
(iv) Work out a confidence interval for the percentage of returned books that were overdue.
(v) At the 95% confidence level, can the library's claim by justified?

9. Jack rolled a dice 240 times and 52 sixes were recorded.
Jack suspected that the dice was biased.
He carried out a hypothesis test to see if the dice was biased.

(i) Write down the null hypothesis H_0.
(ii) Write down the alternative hypothesis H_1.
(iii) What is the sample size?
(iv) Calculate the margin of error, correct to 2 decimal places.
(v) Work out the sample proportion, correct to 2 decimal places.
(vi) If the dice was not biased, what percentage of sixes would Jack expect? Give your answer correct to the nearest whole number.
(vii) Work out a confidence interval for the proportion of sixes obtained.
(viii) Is Jack's suspicion justified at the 95% confidence level?

Test yourself 19

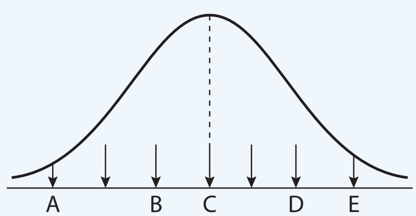
1. In the given normal curve, the mean $\bar{x}$ is 60 and the standard deviation $\sigma = 4$.



(i) Write down the percentage of all the values that are in the shaded area.
(ii) Write down the values represented by A, B and C.
(iii) Use the Empirical Rule to find the range within which 68% of the values lie.
(iv) If a value is selected at random, find the probability that it comes from the shaded area.
(v) If there are 1000 values in the full distribution, how many of them lie in the shaded area?
(vi) How many of the 1000 values are greater than 60?

2. (a) In a normal distribution, the mean $\bar{x} = 30$ and the standard deviation $\sigma = 3$.
(i) Find the range within which 95% of the values lie.
(ii) Find the range within which 99.7% of the values lie.
(iii) What percentage of the values lie in the range [27, 33]?

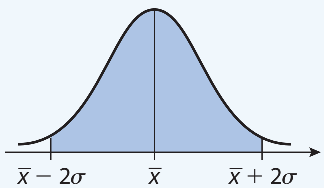
(b) In the given normal curve, the arrows indicate intervals of one standard deviation. If the mean $\bar{x} = 44$ and the standard deviation $\sigma = 6$, write down the values represented by A, B, C, D and E.



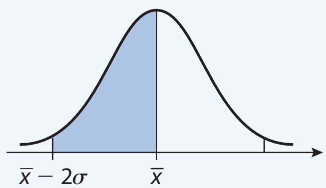
3. A political party claimed that it had the support of 23% of the electorate.
A newspaper carried out an opinion poll to test this claim.
In a random sample of 1000 voters, 250 stated that they support the party.
   (i) Write down the null hypothesis.
   (ii) Write down the alternative hypothesis.
   (iii) What is the sample size?
   (iv) Calculate the margin of error.
   (v) Work out the sample proportion.
   (vi) Work out a confidence interval.
   (vii) At the 95% confidence level, can the party's claim be accepted?

4. (a) The diagrams below show two normal curves.
Write down the percentage of values in each of the shaded areas.

(i)



(ii)



(b) The lengths of roofing nails are normally distributed with a mean of 20 mm and a standard deviation of 3 mm.
   (i) What percentage of nails lie between 17 mm and 23 mm?
   (ii) What percentage of nails lie between 14 mm and 26 mm?
   (iii) If 3000 nails were measured, how many of them would have a length between 17 mm and 23 mm?
   (iv) If a nail is selected at random, what is the probability that its length is greater than 20 mm?

5. The owner of a large apple-orchard claims that 10% of the apples on the trees in his orchard have been attacked by birds. A local wildlife organisation is looking to test this claim.
A random sample of 2500 apples is picked and 275 are found to have been attacked by birds.
   (i) Write down the null hypothesis.
   (ii) Write down the alternative hypothesis.
   (iii) What is the sample size?
   (iv) Work out the margin of error.
   (v) Calculate the sample proportion.
   (vi) Work out a confidence interval.
   (vii) At the 95% confidence level, should the orchard-owner's claim be accepted by the wildlife organisation?

Answers

Exercise 19.1

1. (i) 68% … deviation
(ii) two … mean

2. (i) [51, 57] (ii) [48, 60]
(iii) 68% or 0.68

3. (i) 68% (ii) 95%

4. (i) $\bar{x} - 2\sigma, \bar{x} + 2\sigma$
(ii) 95% or 0.95
(iii) 70 and 98

5. (i) [105, 135] (ii) [90, 150]

6. (i) 60
(ii) 68%
(iii) A = 52; B = 72

7. (i) 46 km/hr (ii) 73 km/hr
(iii) 82 km/hr

8. (i) [55, 65] (ii) [50, 70]

9. a = 150 cm; b = 170 cm; c = 180 cm;
d = 190 cm; e = 200 cm

10. (i) [162 cm, 178 cm]
(ii) [146 cm, 194 cm]

11. (i) 50%
(ii) 68%
(iii) [23 mins, 47 mins]

12. $\sigma = 4$

13. $\sigma = 6$

14. (i) 95% (ii) $[\bar{x} - 2\sigma, \bar{x} + 2\sigma]$
(iii) $\sigma = 6$ (iv) 95% or 0.95

Exercise 19.2

1. (i) 0.03 (ii) 0.03
(iii) 0.02 (iv) 0.04

2. (i) 500 (ii) 0.04
(iii) 0.16

3. (i) 200 (ii) 0.07
(iii) 0.24
(iv) It will reduce the margin of error.

4. (i) 300 (ii) 0.06
(iii) 0.15 (iv) [0.09, 0.21]

5. (i) 150 (ii) 0.08
(iii) 0.3 (iv) [0.22, 0.38]

6. (i) 400 (ii) 0.05
(iii) 0.42 (iv) [0.37, 0.47]

7. 95 … 100 … interval

8. (i) 0.11 (ii) 0.45
(iii) [0.34, 0.56]

9. (i) 200 (ii) 0.07
(iii) 0.42 (iv) [0.35, 0.49]

10. (i) 0.1 (ii) 0.05
(iii) 0.03 (iv) 0.025
(v) 0.065

11. 278

12. (i) 400 (ii) 625
(iii) 70 (iv) 817
(v) 1600

13. (i) 1000 (ii) 0.03
(iii) 0.36 (iv) [0.33, 0.39]

Exercise 19.3

1. (i) H_0: 95% of students are on time for the first class at 9.00 a.m.
(ii) H_1: 95% of students are not on time for the first class at 9.00 a.m.
(iii) I would select a random sample of 30 to 50 students to find out how many were late for the first class. This would give the sample proportion. We could then get the sample error.

2. (i) H_0: 80% of dog owners choose "Chummy Bits".
(ii) H_1: 80% of dog owners do not choose "Chummy Bits".
(iii) 200 (iv) 0.07
(v) 0.775 (vi) [0.705, 0.845]
(vii) Yes, as the claim of 80% (i.e. 0.8) is within the confidence limit [0.705, 0.845]

3. (i) H_0: 20% of visitors to the website purchase at least one of the products.
(ii) 400 (iii) 0.05
(iv) 0.16 (v) [0.11, 0.21]
(vi) Yes, as 20% (0.2) is within the confidence limit [0.11, 0.21]

4. (i) H_0: 40% of voters will vote for the Purple Party.
(ii) 1000 (iii) 0.03
(iv) 0.35 (v) [0.32, 0.38]
(vi) No, as 40% (0.4) is not within the confidence limit [0.32, 0.38]

5. (i) 50%
(ii) H_0: 50% of students admitted are men (or women)
(iii) 500 (iv) 0.04
(v) 0.54 (vi) [0.5, 0.58]
(vii) Yes, as 50% (0.5) is within the confidence limit [0.5, 0.58].

6. (i) H_0: 20% of the plants will have red flowers.
(ii) 0.1 (iii) 0.16
(iv) [0.06, 0.26] (v) Yes
(vi) No, because 20% (0.2) is within the confidence limit [0.06, 0.26]

7. (i) H_0: 90% of migraine sufferers will get relief.
(ii) H_1: 90% of migraine sufferers will not get relief.
(iii) 150 (iv) 0.8
(v) 0.08 (vi) [0.72, 0.88]
(vii) No, as 0.9 (90%) is not within the confidence interval [0.72, 0.88].

8. (i) No: 12% of returned books are overdue.
(ii) 0.07 (iii) 0.075
(iv) [0.005, 0.145]
(v) Yes, as 12% (0.12) is within the confidence limit [0.005, 0.145].

9. (i) H_0: The dice is not biased.
(ii) H_1: The dice is biased.
(iii) 240 (iv) 0.06
(v) 0.22 (vi) 17%
(vii) [0.16, 0.28]
(viii) No, as we accept the null hypothesis because 17% (0.17) is within the confidence limit [0.16, 0.28]

Test yourself 19

1. (i) 95%
(ii) A = 60; B = 52; C = 68
(iii) [56, 64] (iv) 95%
(v) 950 (vi) 500

2. (a) (i) [24, 36]
(ii) [21, 39]
(iii) 68%
(b) A = 26; B = 38; C = 44; D = 56; E = 62

3. (i) H_0: The party has the support of 23% of the electorate.
(ii) H_1: The party does not have the support of 23% of the electorate.
(iii) 1000 (iv) 0.03
(v) 0.25 (vi) [0.22, 0.28]
(vii) Yes, as 23% is within the confidence limit [0.22, 0.28].

4. (a) (i) 95% (ii) 47.5%
(b) (i) 68% (ii) 95%
(iii) 2040 (iv) 50% (or 0.5)

5. (i) H_0: 10% of the apples on the trees in the orchard have been attacked by birds.
(ii) H_1: 10% of the apples have not been attacked by birds.
(iii) 2500 (iv) 0.02
(v) 0.11 (vi) [0.09, 0.13]
(vii) Yes, as 10% (0.1) is within the confidence limit [0.09, 0.13].

SUPPLEMENT

Project Maths

Text & Tests 3

Leaving Certificate Ordinary Level Maths

Contains all the **Deferred Material** of Strand 1
for examination in 2015 and onwards

Statistics: Chapter 19 – Inferential Statistics

The Celtic Press

Ground Floor – Block B,
Liffey Valley Office Campus,
Dublin 22

www.celticpress.ie

ISBN 978-0-7144-2017-2