

**B126\_2**

**Data analysis: hypothesis testing**

**About this free course**

This free course is an adapted extract from the Open University course B126 Business data analytics and decision making - [www.open.ac.uk/courses/modules/b126](https://www.open.ac.uk/courses/modules/b126https://www.open.ac.uk/courses/modules/b126https://www.open.ac.uk/courses/modules/b126?LKCAMPAIGN=ebook_&amp;MEDIA=ou).

This version of the content may include video, images and interactive content that may not be optimised for your device.

You can experience this free course as it was originally designed on OpenLearn, the home of free learning from The Open University –

There you’ll also be able to track your progress via your activity record, which you can use to demonstrate your learning.

First published 2023.

Unless otherwise stated, copyright © 2023 The Open University, all rights reserved.

**Intellectual property**

Unless otherwise stated, this resource is released under the terms of the Creative Commons Licence v4.0 <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en>. Within that The Open University interprets this licence in the following way: [www.open.edu/openlearn/about-openlearn/frequently-asked-questions-on-openlearn](http://www.open.edu/openlearn/about-openlearn/frequently-asked-questions-on-openlearn). Copyright and rights falling outside the terms of the Creative Commons Licence are retained or controlled by The Open University. Please read the full text before using any of the content.

We believe the primary barrier to accessing high-quality educational experiences is cost, which is why we aim to publish as much free content as possible under an open licence. If it proves difficult to release content under our preferred Creative Commons licence (e.g. because we can’t afford or gain the clearances or find suitable alternatives), we will still release the materials for free under a personal end-user licence.

This is because the learning experience will always be the same high quality offering and that should always be seen as positive – even if at times the licensing is different to Creative Commons.

When using the content you must attribute us (The Open University) (the OU) and any identified author in accordance with the terms of the Creative Commons Licence.

The Acknowledgements section is used to list, amongst other things, third party (Proprietary), licensed content which is not subject to Creative Commons licensing. Proprietary content must be used (retained) intact and in context to the content at all times.

The Acknowledgements section is also used to bring to your attention any other Special Restrictions which may apply to the content. For example there may be times when the Creative Commons Non-Commercial Sharealike licence does not apply to any of the content even if owned by us (The Open University). In these instances, unless stated otherwise, the content may be used for personal and non-commercial use.

We have also identified as Proprietary other material included in the content which is not subject to Creative Commons Licence. These are OU logos, trading names and may extend to certain photographic and video images and sound recordings and any other material as may be brought to your attention.

Unauthorised use of any of the content may constitute a breach of the terms and conditions and/or intellectual property laws.

We reserve the right to alter, amend or bring to an end any terms and conditions provided here without notice.

All rights falling outside the terms of the Creative Commons licence are retained or controlled by The Open University.

Head of Intellectual Property, The Open University

# Contents

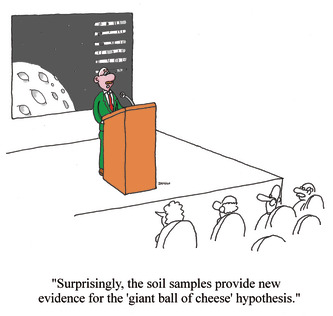
* [Introduction](#Session1)
* [Learning outcomes](#Session2)
* [1 Two types of hypotheses](#Session3)
  + [1.1 Formulating null and alternative hypotheses](#Session3_Section1)
  + [1.2 Population mean (µ)](#Session3_Section2)
* [2 Testing with data](#Session4)
* [3 Alpha (α) levels](#Session5)
* [4 One-tailed vs two-tailed test](#Session6)
  + [4.1 The normal distribution](#Session6_Section1)
  + [4.2 Two-tailed tests](#Session6_Section2)
  + [4.3 One-sided tests](#Session6_Section3)
  + [4.4 Check your understanding](#Session6_Section4)
* [5 Mean and z-score ranges](#Session7)
  + [5.1 Acceptance and rejection regions](#Session7_Section1)
  + [5.2 Test your understanding](#Session7_Section2)
  + [5.3 Using the z-score](#Session7_Section3)
* [6 P-value](#Session8)
  + [6.1 Defining the p-value](#Session8_Section1)
  + [6.2 Calculating the p-value](#Session8_Section2)
  + [6.3 Example: testing a hypothesis](#Session8_Section3)
  + [6.4 Test your understanding](#Session8_Section4)
* [7 Hypothesis testing for population proportions](#Session9)
  + [7.1 Example: testing a proportion](#Session9_Section1)
  + [7.2 Test your understanding](#Session9_Section2)
* [Conclusion](#Session10)
* [References](#Session11)
* [Acknowledgements](#Session12)
* [Solutions](#Solutions1)
* [Descriptions](#Descriptions1)

## Introduction

In this course, you will explore the process of testing hypotheses and making inferences about data. This will typically be a collection of individuals in a population in the statistical sense (people, companies, countries, etc.), and the quantity of interest will be a parameter describing the population.

You will learn about descriptive statistics, alpha (α) and confidence intervals as well as the distinction between one-tailed and two-tailed tests, and the concept of ‘p-value’. Furthermore, you will gain insight into how to test for differences in means and proportions. Ultimately, you will gain knowledge about the concept of ‘tests for statistical significance’.

Start of Figure



**Figure 1** Data sheds light on hypotheses

[View description - Figure 1 Data sheds light on hypotheses](" \l "Session1_Description1)

[View alternative description - Figure 1 Data sheds light on hypotheses](" \l "Session1_Alternative1)

End of Figure

This course requires Microsoft Excel in some activities therefore this course should be completed on a desktop or laptop rather than a mobile device. If you do not have access to Microsoft Excel there are various other free options – such as Google Sheets, Apple Numbers or LibreOffice.

This OpenLearn course is an adapted extract from the Open University course [B126 Business data analytics and decision making](https://www.open.ac.uk/courses/modules/b126).

## Learning outcomes

After studying this course, you should be able to:

* understand the principle of hypothesis testing
* understand the idea of alpha in hypothesis testing
* differentiate between one-tailed and two-tailed tests
* understand hypothesis testing of means and proportions
* report the exact p-value of a test.

## 1 Two types of hypotheses

During the process of data-driven decision making, managers typically follow four key steps:

1. Formulating a hypothesis
2. Identifying and obtaining the appropriate data to test the hypothesis
3. Executing the test
4. Making a decision based on the results.

In this course, your focus will be on gaining a comprehensive understanding of what a hypothesis is and how to effectively formulate one.

The term hypothesis refers to an explanation proposed for a phenomenon or idea (e.g. ‘the sky is blue’). In order for a hypothesis to be considered scientific, it must be tested scientifically. In simple terms, a hypothesis is a ‘guess’ that ***can be tested***.

Start of Figure



**Figure 2** iPhone

[View description - Figure 2 iPhone](" \l "Session3_Description1)

[View alternative description - Figure 2 iPhone](" \l "Session3_Alternative1)

End of Figure

For example, say when conducting market research, a respondent states that the selling price of Apple iPhones was too high (i.e. it was too expensive). This is a supposition but cannot be tested. This is because you do not know what is meant by the term ‘expensive’. However, if the respondent mentions that the Apple iPhone costs over £500, this can be considered as expensive. In this case, the respondent’s supposition becomes a hypothesis. In order to test the hypothesis, market researchers need to collect some data (typically by surveying a small group of people – a sample) to study the population. The concept of population refers to the complete collection of individuals or objects that are of interest in a given context. This can include, for example, all small- and medium-sized businesses within a particular country, or the entirety of online shoppers.

The concepts of null hypothesis and alternative hypothesis are fundamental in business decision making.

The null hypothesis is an algebraic statement that expresses the currently accepted value for a parameter in the population. It is commonly used as the default position when testing a hypothesis, and the researcher presumes it to be true unless there is sufficient evidence to the contrary. The null hypothesis is usually represented by ‘H0’. For instance, if it is believed that the average annual salary in the UK is approximately £26,000, it can be considered as the null hypothesis, which can be formally stated as:

H0: The average annual salary in the UK is £26,000.

On the other hand, the alternative hypothesis is a statement or assumption that challenges the null hypothesis. It is the opposite of the null hypothesis and is what the researcher aims to demonstrate to be true. The alternative hypothesis is denoted as ‘Ha’ or ‘H1’.

An alternative hypothesis is needed because it provides an alternative explanation or theory to the null hypothesis. You will look at more examples in the next section.

## 1.1 Formulating null and alternative hypotheses

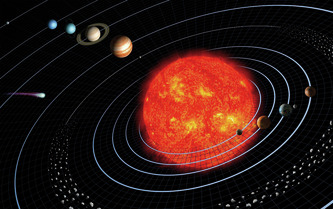
In the world of scientific inquiry, you often begin with a null hypothesis (H0), which expresses the currently accepted value for a parameter in the population. The alternative hypothesis (Ha), on the other hand, is the opposite of the null hypothesis and challenges the currently accepted value.

To illustrate this concept of null and alternative hypotheses, you will look at some well-known stories and examples.

In ancient and medieval times, the widely held belief was that all planets orbited around the Earth, as the Earth was considered the centre of the universe. This idea can be considered the null hypothesis, as it represents the currently accepted value for a parameter in the population. Thus, it can be written as:

H0: All planets orbit around the Earth.

Start of Figure



**Figure 3** Solar system

[View description - Figure 3 Solar system](" \l "Session3_Description2)

[View alternative description - Figure 3 Solar system](" \l "Session3_Alternative2)

End of Figure

In the world of business and finance, the idea that paper money must be backed by gold (the gold standard) was also a commonly held belief for a long time. This belief can be considered a null hypothesis. However, following the Great Depression, people began to question this belief and broke the link between banknotes and gold. This alternative hypothesis challenged the gold standard, and it eventually became widely accepted that the value of paper money is not necessarily equal to a fixed amount of gold. Thus, H0 and Ha statements can be written as:

H0: The value of paper money is equal to a fixed amount of gold.

Ha: The value of paper money is not equal to a fixed amount of gold.

In modern times, people generally place their trust in the value of banknotes issued by central banks or monetary authorities, which are backed by a strong government. This belief can be considered a null hypothesis. However, digital currency, such as Bitcoin, has emerged as an alternative to traditional paper money. Bitcoin is not backed by any central bank or monetary authority, and transactions involving Bitcoin are verified by network nodes using cryptography and recorded in a blockchain. This alternative hypothesis challenges the belief that the value of paper money is solely based on people's trust in central banks or monetary authorities. Thus, H0 and Ha statements can be written as:

H0: The value of paper money is equal to people’s trust in central banks or monetary authorities.

Ha: The value of paper money is not equal to people’s trust in central banks or monetary authorities.

In conclusion, the alternative hypothesis always challenges the idea expressed in the null hypothesis. By testing the null hypothesis against the alternative hypothesis, you can determine which idea is more supported by the available data. The alternative hypothesis is often referred to as a ‘research hypothesis’ because it initiates the motivation and opportunities for further research.

Let’s return to the first example given in Section 1. If you see that your friends and relatives make more or less than £26,000 annually on average, perhaps you should question the widely accepted proposition of £26,000 as the average annual salary in the UK. This will enable you to develop an alternative hypothesis:

Ha: Average annual salary in the UK is not equal to £26,000.

The following activity will test your knowledge of null and alternative hypotheses.

Start of Activity

**Activity 1 Null hypothesis versus alternative hypothesis**

Allow approximately 10 minutes to complete this activity

Start of Question

Read the following statements. Can you develop a null hypothesis and an alternative hypothesis?

‘It is believed that a high-end coffee machine produces a cup of caffè latte with an average of 1 cm of foam. The hotel employee claims that after the machine has been repaired, it is no longer able to produce a cup of caffè latte with 1cm foam.’

End of Question

*Provide your answer...*

[View discussion - Activity 1 Null hypothesis versus alternative hypothesis](" \l "Session3_Discussion1)

End of Activity

If you have developed the hypotheses H0 and Ha as mentioned in the discussion to Activity 1, you have shown that you are familiar with the structure of different types of hypotheses. However, in the next section you will explore the concept of hypothesis formulation further.

## 1.2 Population mean (µ)

As you have seen so far in Section 1, the formulation of H0 is based on widely accepted beliefs, which can also be interpreted as the population mean (µ), representing the average value in the population (where µ is the symbol for a population mean).

To illustrate this point, consider the example of the average UK salary, which is ‘widely believed’ to be £26,000 per year. In this case, µ is equivalent to £26,000, which can be expressed as:

H0: µ = £26,000

As H0 and Ha are always opposite (i.e. since the purpose of Ha is to challenge the belief of H0), you can also express Ha as:

Ha: µ ≠ £26,000.

If you revisit the caffè latte foam example from Activity 1, H0 and Ha can also be expressed as:

H0: µ = 1cm foam (representing the population mean for the height of foam in a caffè latte)

Ha: µ ≠ 1cm foam (indicating a departure from the widely accepted belief about the population mean).

In summary, the process of hypothesis formulation is a critical step in business decision making. It involves setting up the null hypothesis (H0), which represents the widely accepted belief or the status quo (or µ), and the alternative hypothesis (Ha), which challenges the belief and suggests the possibility of a significant difference. A clear understanding of the concept of hypothesis formulation is crucial for accurate business decision making.

Having discussed how hypotheses are formulated, you will now turn your attention to how to test them. In the following sections, you will look at test statistics and alpha levels, which play a critical role in determining the statistical significance of a hypothesis test.

## 2 Testing with data

Once hypotheses have been developed, the next step is to find existing data or design a study (to generate some primary data) to test them. The results of the testing will have two possible outcomes.

1. Reject null hypothesis (H0) – this also means that the alternative hypothesis (Ha) is accepted.
2. Fail to reject null hypothesis (H0) – this means that you accept H0 and reject alternative hypothesis (Ha).

You can test hypotheses empirically using test statistics that involve calculating sample data and using the results to decide whether you need to reject the null hypothesis or fail to reject the null hypothesis.

A ‘test statistic’ is a number calculated by a statistical test, which provides information about how much the relationship between variables in the test differs from the null hypothesis. There are different types of test statistics and the choice of test statistic depends on the type of hypothesis being tested and the nature of the data. For example, if the data is continuous and normally distributed, the t-test or z-test may be used as a test statistic. If the data is categorical, the chi-squared test may be used. Some of these will be discussed later in this course.

To begin, you will first obtain an overview of the general guidelines for using the test statistic to determine whether to reject or fail to reject the null hypothesis. Consider the coffee example from Activity 1. In order to test the null hypothesis (H0: µ = 1cm foam), you might sample 60 cups of caffè latte made by the coffee machine – i.e. select 60 cups of caffè latte at random (random sampling) from all the cups of caffè latte that the coffee machine made throughout the day. Then, you would measure the height of the foam to obtain the average height and calculate the test statistic.

You may be wondering why here the null hypothesis and not the alternative hypothesis (H0: µ ≠ 1cm foam) is being exclusively discussed. In the realm of hypothesis testing, you commence by assuming that the null hypothesis is true and then employ sample data to determine whether or not to reject it. If you reject the null hypothesis, you can infer that the alternative hypothesis is true. As you may recall, the null and alternative hypotheses are complementary and mutually exclusive. The null hypothesis denotes the prevailing belief, whereas the alternative hypothesis represents what you aim to demonstrate. Therefore, rejecting the null hypothesis implies that you have evidence to support the alternative hypothesis. It is noteworthy that you cannot directly prove the alternative hypothesis. Instead, you can only reject or fail to reject the null hypothesis. The decision to reject or fail to reject the null hypothesis is grounded in the test statistics.

Start of Figure



**Figure 4** Caffè latte

[View description - Figure 4 Caffè latte](" \l "Session4_Description1)

[View alternative description - Figure 4 Caffè latte](" \l "Session4_Alternative1)

End of Figure

For the study:

* Researcher 1 samples 60 cups of caffè latte and gets the average foam height equal to 1.1 cm.
* Researcher 2 samples 60 cups of caffè latte and gets the average foam height equal to 1.5 cm.
* Researcher 3 samples 60 cups of caffè latte and gets the average foam height equal to 2.6 cm.

The question now is how to make your decision. Using only the data collected by Researcher 1, the average height of the foam is 1.1 cm. It is not exactly equal to 1 cm, but it is very close. In this case, you may say that you cannot reject the null hypothesis. However, if you look at the data collected by Researcher 3 (the average height of the foam is equal to 2.6 cm), this is far beyond the commonly accepted idea of 1 cm foam for a caffè latte. Therefore, you can reject the null hypothesis and accept the alternative hypothesis (µ ≠ 1 cm foam). Considering the data collected by Researchers 1 and 3, it is easy to decide whether to reject the null hypothesis or not. However, based on the data collected by Researcher 2, it is extremely difficult to make a decision. Although the average foam height of 1.5 cm is not far away from 1 cm, does it go far enough to be considered sufficient?

In order to answer this question, you would need to introduce the concept of ‘statistical significance’, which you will look at in more detail in the next section.

## 3 Alpha (α) levels

The level of statistical significance is the threshold at which you decide whether to reject the null hypothesis (to make a decision). A result is statistically significant when the situation described in the null hypothesis is highly unlikely to have occurred.

By using the concept of statistical significance, you can have a concrete way of examining the claim concerning the null hypothesis, using the data collected, to make a clear decision on when to reject the null hypothesis and when to leave it. In this way, you do not have to guess whether the test statistic is too high or too low.

How can you determine the appropriate level of statistical significance to use? In order to answer this question, you must introduce the concept of ‘level of confidence’ (C) – how confident are you in your decision to reject the null hypothesis? C is related to the concept of ‘confidence intervals’ that you may have seen elsewhere. C is usually set at 95% when establishing a confidence interval. Having determined C, you can calculate the level of statistical significance that you want to use as 1 − C.

If you choose the level of confidence (C) at 95%, then the level of statistical significance that you set up for your decision is equal to 1 − 95% = 5%.

The level of statistical significance is referred to as ‘alpha’ or ‘α’.

So, returning back to the question on caffè latte foam, is 1.5 cm average foam height, which was only slightly different from the expected value of 1 cm, significant enough to reject or fail to reject the null hypothesis (H0)? The researchers had to consider the alpha level to determine the appropriate course of action. The alpha level is a crucial factor in hypothesis testing and is used to determine the threshold for statistical significance. In other words, it helps to determine whether the results obtained are due to chance or if they represent a genuine difference between the expected and observed values.

Calculating the alpha level can be tricky, but it is a necessary step in hypothesis testing (you will discuss how to calculate the alpha level in later sections). Once the alpha level has been determined, it can be used to decide whether to accept or reject the null hypothesis based on the findings of Researcher 2. The decision of whether to reject or fail to reject H0 based on the findings of Researcher 2 will ultimately depend on the alpha level chosen, so the researchers will have to carefully consider this factor before drawing any conclusions from the study.

You will be tested on your understanding of how to transfer levels of confidence to alpha levels in the following activity.

Start of Activity

**Activity 2  Level of alpha**

Allow approximately 10 minutes to complete this activity

Start of Question

Imagine you are working as a marketing analyst for a new product launch, and you want to make sure that your market research is accurate and reliable. You need to determine the appropriate alpha level for your survey results, based on the desired confidence level.

Can you determine α?

Start of Table

Table 1  Level of alpha

|  |  |
| --- | --- |
| **Level of Confidence (C)** | **α** |
| 90% | *Provide your answer...* |
| 95% | *Provide your answer...* |
| 99% | *Provide your answer...* |

End of Table

End of Question

[View discussion - Activity 2  Level of alpha](" \l "Session5_Discussion1)

End of Activity

There is an inverse relationship between confidence levels and alpha levels, as increasing the confidence level leads to a decrease in the alpha level, and vice versa. It is important to carefully consider the appropriate alpha level and confidence level when conducting hypothesis testing to ensure that the appropriate level of risk and certainty are balanced.

## 4 One-tailed vs two-tailed test

To gain a deeper understanding of how to conduct a hypothesis test, this section will delve into the concepts of one-tailed and two-tailed tests. These tests are vital tools in statistical hypothesis testing, and the decision of which test to employ depends on the research question and hypothesis under examination. It is crucial to give careful thought to the suitable type of test to ensure that the hypothesis is thoroughly tested and precise conclusions are derived from the data. This section will elaborate on this topic in greater detail.

To commence, complete the following activity pertaining to the formulation of null and alternative hypotheses. This exercise may be somewhat challenging, but it serves as an excellent introduction to upcoming discussions – don’t be concerned if you find it difficult!

Start of Activity

**Activity 3 Hypotheses setting**

Allow approximately 10 minutes to complete this activity

Start of Question

Read the following statements and then develop a null hypothesis and an alternative hypothesis.

‘It is believed that OU students need to set aside no longer than, on average, 15 hours to study an entire session of an OU course. However, a researcher believes that OU students spend longer studying an entire session of an OU course.’

End of Question

*Provide your answer...*

[View discussion - Activity 3 Hypotheses setting](" \l "Session6_Discussion1)

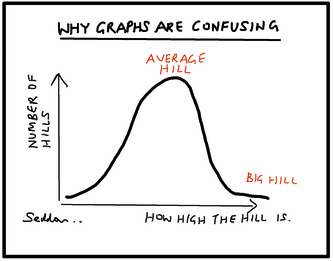
End of Activity

Did you identify any differences between the hypotheses you developed in Activity 1 and Activity 3? The set of hypotheses in Activity 1 has an equal (=) or not equal (≠) supposition (sign) in the statement. However, in Activity 3, the set of hypotheses has less than or equal to (≤) and greater than (>) supposition (sign) in the statement. This creates different conditions that lead to acceptance or rejection of the null hypothesis.

## 4.1 The normal distribution

Here, you will look at the concept of normal distribution and the bell-shaped curve. The peak point (the top of the bell) represents the most probable occurrences, while other possible occurrences are distributed symmetrically around the peak point, creating a downward-sloping curve on either side of the peak point.

Start of Figure



**Figure 5** The peak point is average

[View description - Figure 5 The peak point is average](" \l "Session6_Description1)

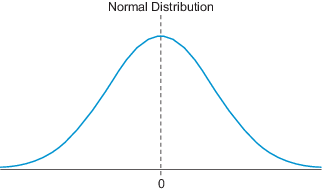
[View alternative description - Figure 5 The peak point is average](" \l "Session6_Alternative1)

End of Figure

In order to test hypotheses, you need to calculate the test statistic and compare it with the value in the bell curve. This will be done by using the concept of ‘normal distribution’.

A normal distribution is a probability distribution that is symmetric about the mean, indicating that data near the mean are more likely to occur than data far from it. In graph form, a normal distribution appears as a bell curve. The values in the x-axis of the normal distribution graph represent the z-scores. The test statistic that you wish to use to test the set of hypotheses is the ***z-score***. A z-score is used to measure how far the observation (sample mean) is from the 0 value of the bell curve (population mean). In statistics, this distance is measured by standard deviation. Therefore, when the z-score is equal to 2, the observation is 2 standard deviations away from the value 0 in the normal distribution curve.

Start of Figure



**Figure 6** Normal distribution – bell curve

[View description - Figure 6 Normal distribution – bell curve](" \l "Session6_Description2)

[View alternative description - Figure 6 Normal distribution – bell curve](" \l "Session6_Alternative2)

End of Figure

## 4.2 Two-tailed tests

Hypotheses that have an equal (=) or not equal (≠) supposition (sign) in the statement are called **non-directional hypotheses**. In non-directional hypotheses, the researcher is interested in whether there is a statistically significant difference or relationship between two or more variables, but does not have any specific expectation about which group or variable will be higher or lower. For example, a non-directional hypothesis might be: ‘There is a difference in the preference for brand X between male and female consumers.’ In this hypothesis, the researcher is interested in whether there is a statistically significant difference in the preference for brand X between male and female consumers, but does not have a specific prediction about which gender will have a higher preference. The researcher may conduct a survey or experiment to collect data on the brand preference of male and female consumers and then use statistical analysis to determine whether there is a significant difference between the two groups.

Non-directional hypotheses are also known as two-tailed hypotheses. The term ‘two-tailed’ comes from the fact that the statistical test used to evaluate the hypothesis is based on the assumption that the difference or relationship could occur in either direction, resulting in two ‘tails’ in the probability distribution. Using the coffee foam example (from Activity 1), you have the following set of hypotheses:

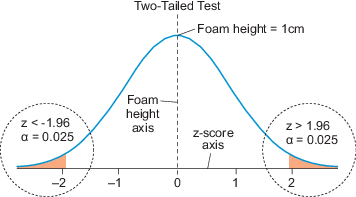
H0: µ = 1cm foam

Ha: µ ≠ 1cm foam

In this case, the researcher can reject the null hypothesis for the mean value that is either ‘much higher’ or ‘much lower’ than 1 cm foam. This is called a **two-tailed test** because the rejection region includes outcomes from both the upper and lower tails of the sample distribution when determining a decision rule. To give an illustration, if you set alpha level (α) equal to 0.05, that would give you a 95% confidence level. Then, you would reject the null hypothesis for obtained values of z < 1.96 and z > 1.96 (you will look at how to calculate z-scores later in the course).

This can be plotted on a graph as shown in Figure 7.

Start of Figure



**Figure 7** Two-tailed test

[View description - Figure 7 Two-tailed test](" \l "Session6_Description3)

[View alternative description - Figure 7 Two-tailed test](" \l "Session6_Alternative3)

End of Figure

In a two-tailed hypothesis test, the null hypothesis assumes that there is no significant difference or relationship between the two groups or variables, and the alternative hypothesis suggests that there is a significant difference or relationship, but does not specify the direction of the difference or relationship.

When performing a two-tailed test, you need to determine the level of significance, which is denoted by alpha (α). The value of alpha, in this case, is 0.05. To perform a two-tailed test at a significance level of 0.05, you need to divide alpha by 2, giving a significance level of 0.025 for each distribution tail (0.05/2 = 0.025). This is done because the two-tailed test is looking for significance in either tail of the distribution. If the calculated test statistic falls in the rejection region of either tail of the distribution, then the null hypothesis is rejected and the alternative hypothesis is accepted. In this case, the researcher can conclude that there is a significant difference or relationship between the two groups or variables.

Assuming that the population follows a normal distribution, the tail located below the critical value of z = –1.96 (in a later section, you will discuss how this value was determined) and the tail above the critical value of z = +1.96 each represent a proportion of 0.025. These tails are referred to as the lower and upper tails, respectively, and they correspond to the extreme values of the distribution that are far from the central part of the bell curve. These critical values are used in a two-tailed hypothesis test to determine whether to reject or fail to reject the null hypothesis. The null hypothesis represents the default assumption that there is no significant difference between the observed data and what would be expected under a specific condition.

If the calculated test statistic falls within the critical values, then the null hypothesis cannot be rejected at the 0.05 level of significance. However, if the calculated test statistic falls outside the critical values (orange-coloured areas in Figure 7), then the null hypothesis can be rejected in favour of the alternative hypothesis, suggesting that there is evidence of a significant difference between the observed data and what would be expected under the specified condition.

## 4.3 One-sided tests

As well as non-directional hypotheses, you will also encounter hypotheses that have a less than or equal to (≤) and greater than (>) supposition (sign) in the statement (as you saw in Activity 3). This is called a **directional** hypothesis. A directional hypothesis is a type of research hypothesis that aims to predict the direction of the relationship or difference between two variables. Essentially, it specifies the anticipated outcome of a study prior to the collection of data.

For example, a directional hypothesis might propose that a marketing campaign will increase product sales, predicting the direction of the relationship (i.e. the marketing campaign will lead to an increase in product sales). In contrast, a non-directional hypothesis simply states that there is a relationship between two variables without specifying the direction of that relationship, such as: ‘There is a relationship between the marketing campaign and product sales.’

Directional hypotheses are often preferred in scientific research because they provide a more precise and focused prediction than non-directional hypotheses. In business management, a directional hypothesis can also be a useful tool. For example, a company may use a directional hypothesis to design a study that examines the effectiveness of a marketing campaign in enhancing sales. This approach provides a clearer understanding of the impact of the campaign and enables the company to make more informed decisions about future marketing strategies.

A one-tailed test is a statistical test employed to evaluate a directional hypothesis, which predicts the direction of the difference or association between two variables. Its objective is to ascertain if the data supports the anticipated direction.

To illustrate, consider the hypotheses from Activity 3:

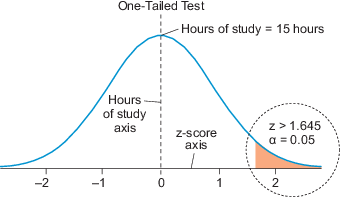
H0: µ ≤ 15 hours of studies

Ha: µ > 15 hours of studies

The null hypothesis (H0) posits that the population mean (µ) is less than or equal to 15 hours of studies, while the alternative hypothesis (Ha) predicts that the population mean is greater than 15 hours of studies.

To conduct a one-tailed test, a critical value must be established to determine whether the null hypothesis should be rejected or retained. Typically, a significance level (α) is set for this purpose. For instance, assuming α = 0.05, the z-score for a one-tailed test with α = 0.05 in a normal distribution is 1.645. Consequently, the null hypothesis would be rejected if the z-score exceeds 1.645. In other words, only the upper tail region of the distribution is rejected for a one-tailed test. Additionally, you employ distinct z-scores since, in contrast to a two-tailed test, the alpha level does not need to be divided by two. In a normal distribution, the area in the tail above z = +1.645 represents 0.5 of the distribution. This portion of the distribution is significantly remote from the centre of the bell curve at 0. Consequently, the null hypothesis would be rejected if the z-score exceeds 1.645 (as depicted in Figure 8).

Start of Figure



**Figure 8** One-tailed test – right (upper) tail

[View description - Figure 8 One-tailed test – right (upper) tail](" \l "Session6_Description4)

[View alternative description - Figure 8 One-tailed test – right (upper) tail](" \l "Session6_Alternative4)

End of Figure

In summary, a one-tailed test is used to assess a directional hypothesis in which the direction of the difference or association between two variables is predicted. The critical value for a one-tailed test is determined by the selected significance level (α), and the test is conducted to ascertain whether the data supports the predicted direction.

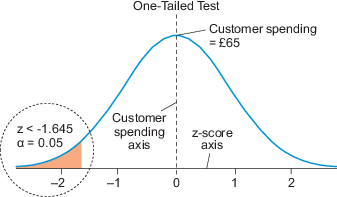
In addition, the one-tailed test is not limited to a single direction (greater than) but can also be employed in the opposite direction (less than). An example can be used to illustrate this type of hypothesis testing. Consider a situation where the management team believes that the average amount spent by customers during their visits to a department store is £65. However, the service manager observes that customers spend less than that amount during their visits. In this case, you can formulate the following set of hypotheses:

H0: µ ≥ £65

Ha: µ < £65

To test this directional hypothesis, a one-tailed test must be conducted. The alternative hypothesis states that the specific value of µ will be lower than the value specified in the hypothesis. Therefore, you must reject the region in the lower tail of the normal distribution. More specifically, the rejection region of the one-tailed test at alpha levels equals 0.05. The lower tail of the normal distribution has a z-score lower than -1.645. Any hypothesis in this region will be rejected. The graph in Figure 9 illustrates this.

Start of Figure



**Figure 9** One-tailed test – left (lower) tail

[View description - Figure 9 One-tailed test – left (lower) tail](" \l "Session6_Description5)

[View alternative description - Figure 9 One-tailed test – left (lower) tail](" \l "Session6_Alternative5)

End of Figure

In conclusion, the one-tailed test is not restricted to a specific direction and can be used in either direction, depending on the research question and the hypothesis being tested. The test is used to determine if the data supports a directional hypothesis, and a critical value is established based on the significance level chosen for the test.

## 4.4 Check your understanding

In the following activity, you will be tested on your understanding of one-tailed and two-tailed tests with respect to a bell-shaped curve.

Start of Activity

**Activity 4 One-tailed or two-tailed test**

Allow approximately 10 minutes to complete this activity

Start of Question

Can you match the statements with the bell curves that indicate either a one-tailed or two-tailed test?

Start of Media Content

Interactive content is not available in this format.

End of Media Content

End of Question

End of Activity

In conclusion, the choice between a one-tailed test and a two-tailed test (as well as between a directional and non-directional hypothesis) depends on the research question and the nature of the relationship or difference being studied. Two-tailed tests and non-directional hypotheses are useful when the research question is exploratory, or the direction of the relationship is unknown, while one-tailed tests and directional hypotheses are useful when a specific direction of the relationship or difference is predicted and is of interest. Careful consideration and planning are necessary when choosing between these options to ensure that the hypothesis is properly tested and the correct conclusions are drawn from the data. Ultimately, the choice of the type of test and hypothesis should be guided by the research question, the theory behind the study, and the available data.

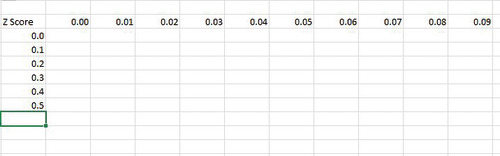
## 5 Mean and z-score ranges

In the previous section on hypothesis testing using the normal distribution, the z-score was frequently mentioned. This is because the z-score is the test statistic that is used to determine whether the null hypothesis should be accepted or rejected.

One way to gain an understanding of the calculated z-score and its alpha value is through the use of a z-score table. Using Excel functions, you can create a table that indicates z-scores and the corresponding area under the normal distribution curve. As you continue through the course, you will see how the z-score table can be used in hypothesis testing. For now, however, you will focus your attention on the use of Excel to create a table of z-scores, using the following steps.

**Step 1**: In an Excel spreadsheet, you can create a z-score table yourself that shows the z-score associated with any level of probability. To give an illustration, the screenshot below (Figure 10) shows a table with the values – 0.0; 0,1; 0.2; 0.3; 0.4; 0.5 – in a row and the values – 0.00; 0.01; 0.02; 0.03; 0.04; 0.05; 0.06; 0.07; 0.08; 0.09 – in columns. These are the first and second decimal digits of the probability whose z-score will be in the table. For example, the z-score in the first row (0.0) and second column (0.01) in the as-yet empty table refers to the probability p=0.01 (i.e., 0.0+0.01=0.01). The table is really just a long list for probabilities from p=0 to p=0.5 formatted as a table.

Start of Figure



**Figure 10** Creating a z-score table step 1

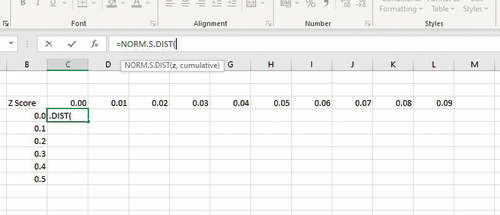
[View description - Figure 10 Creating a z-score table step 1](" \l "Session7_Description1)

[View alternative description - Figure 10 Creating a z-score table step 1](" \l "Session7_Alternative1)

End of Figure

**Step 2**: To use the ‘NORM.S.DIST’ Excel formula, begin by typing ‘=NORM.S.DIST(’ into the desired cell. The Excel software will prompt you to complete the formula’s value entry by entering the appropriate values for ‘z’ and ‘cumulative’.

Start of Figure



**Figure 11** Creating a z-score table step 2

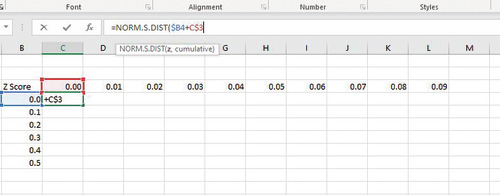
[View description - Figure 11 Creating a z-score table step 2](" \l "Session7_Description2)

[View alternative description - Figure 11 Creating a z-score table step 2](" \l "Session7_Alternative2)

End of Figure

**Step 3**: After initiating the NORM.S.DIST formula in the designated cell, you can assign a value to ‘z’ by adding the value in the row cell to the value in the column cell.

Start of Figure



**Figure 12** Creating a z-score table step 3

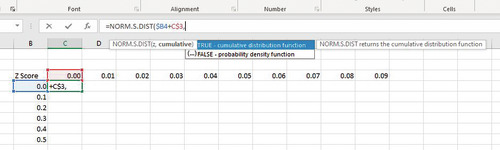
[View description - Figure 12 Creating a z-score table step 3](" \l "Session7_Description3)

[View alternative description - Figure 12 Creating a z-score table step 3](" \l "Session7_Alternative3)

End of Figure

**Step 4**: To indicate the cumulative distribution function, set the ‘cumulative’ argument to ‘TRUE’.

Start of Figure



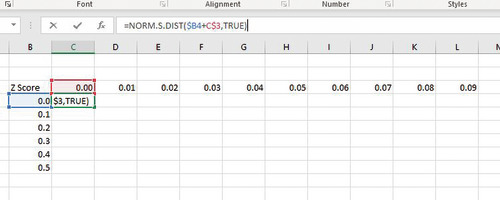
**Figure 13** Creating a z-score table step 4

[View description - Figure 13 Creating a z-score table step 4](" \l "Session7_Description4)

[View alternative description - Figure 13 Creating a z-score table step 4](" \l "Session7_Alternative4)

End of Figure

Start of Figure



**Figure 14** Creating a z-score table step 4

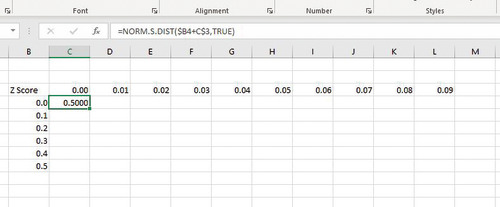
[View description - Figure 14 Creating a z-score table step 4](" \l "Session7_Description5)

[View alternative description - Figure 14 Creating a z-score table step 4](" \l "Session7_Alternative5)

End of Figure

**Step 5**: After completing the formula with the appropriate values, press ‘Enter’ to calculate the result for the selected cell (Figure 15). To apply the Excel function to all cells within the table, click and drag the green box to cover the desired area (Figure 16).

Start of Figure



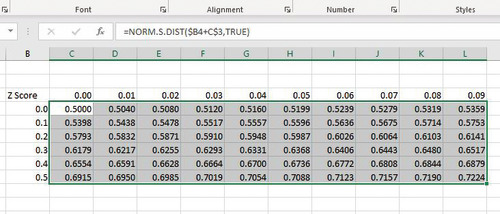
**Figure 15** Creating a z-score table step 5

[View description - Figure 15 Creating a z-score table step 5](" \l "Session7_Description6)

[View alternative description - Figure 15 Creating a z-score table step 5](" \l "Session7_Alternative6)

End of Figure

Start of Figure



**Figure 16** Creating a z-score table step 5

[View description - Figure 16 Creating a z-score table step 5](" \l "Session7_Description7)

[View alternative description - Figure 16 Creating a z-score table step 5](" \l "Session7_Alternative7)

End of Figure

In the following activity, you can practice using these Excel formulas to create a table of values of the normal distribution that correspond to a range of z-scores. These will be useful when testing hypotheses later.

Start of Activity

**Activity 5  Z-Score table**

Allow approximately 20 minutes to complete this activity

Start of Question

Using steps 1 to 5 described above, create a z-score table in Excel for z-scores ranging from -3 to 3.

Once you have completed these steps, reveal the discussion and compare your answers.

End of Question

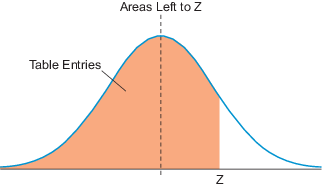
[View discussion - Activity 5  Z-Score table](" \l "Session7_Discussion1)

End of Activity

## 5.1 Acceptance and rejection regions

The z-score table you created in Activity 5 represents the area under the normal distribution bell curve left of z (as shown in Figure 17).

Start of Figure



**Figure 17** Area under the curve left of z

[View description - Figure 17 Area under the curve left of z](" \l "Session7_Description8)

[View alternative description - Figure 17 Area under the curve left of z](" \l "Session7_Alternative8)

End of Figure

The entries in this table can be used to determine whether to accept or reject the null hypothesis.

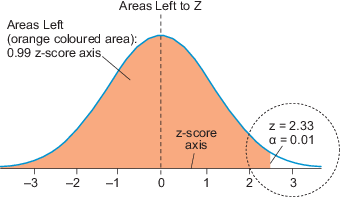
Suppose a marketing team at a company wishes to test the hypothesis that a new ad campaign will lead to a significant increase in sales. The team could use a one-tailed test with the reject region in the upper (right) tail and an alpha level of 1%.

Using the table created in Activity 5, the team can identify the range of z-scores that correspond to this test. They can then calculate the test statistic based on the data collected from the sales during and after the ad campaign. If the calculated test statistic falls within the rejection region identified by the table, the team can reject the null hypothesis and conclude that the ad campaign has had a significant impact on sales. This information can be used by the marketing team to justify the investment in the ad campaign and to make informed decisions about future marketing strategies.

In the context of the marketing team's hypothesis testing, the reject region for the one-tailed test with an alpha level of 1% corresponds to the range of z-scores that fall within the top 1% of the normal distribution. Conversely, the acceptable range refers to the range of z-scores that corresponds to the remaining 99% of the distribution to the left of z. Using the table created in Activity 5, the marketing team can identify the specific range of z-scores that correspond to the acceptable range and the reject region. Based on this table, the z-score of 2.33 corresponds to the upper limit of the acceptable range, as the area to the left of z = 2.33 represents approximately 99% of the area under the curve.

Therefore, if the team obtains a calculated z-score that is greater than 2.33, they can reject the null hypothesis and conclude that the new ad campaign has had a significant impact on sales. This information can help the marketing team make data-driven decisions about future campaigns and allocate resources effectively to maximise sales and profits. Figure 18 below illustrates this.

Start of Figure



**Figure 18** One-tailed test with Alpha level of 1%

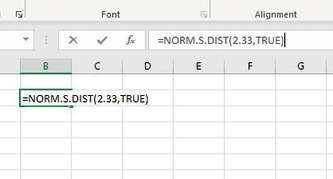
[View description - Figure 18 One-tailed test with Alpha level of 1%](" \l "Session7_Description9)

[View alternative description - Figure 18 One-tailed test with Alpha level of 1%](" \l "Session7_Alternative9)

End of Figure

Other than creating a z-score table, you calculate the region to the left of z by using the Excel formula **NORM.S.DIST(z, cumulative).** For example, you can calculate the region left of z when z = 2.33 by simply entering 2.33 as a z-score and setting the cumulative to be ‘TRUE’ in this Excel formula.

Start of Figure



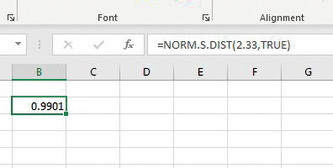
**Figure 19** Calculate the region left of z without using z-score table 1

[View description - Figure 19 Calculate the region left of z without using z-score table 1](" \l "Session7_Description10)

[View alternative description - Figure 19 Calculate the region left of z without using z-score table 1](" \l "Session7_Alternative10)

End of Figure

Start of Figure



**Figure 20** Calculate the region left of z without using z-score table 2

[View description - Figure 20 Calculate the region left of z without using z-score table 2](" \l "Session7_Description11)

[View alternative description - Figure 20 Calculate the region left of z without using z-score table 2](" \l "Session7_Alternative11)

End of Figure

You should get a value reading of 0.9901, which is exactly what you found in the z-score table in row 2.3 and column 0.03.

Here is another question. If you want to test hypotheses using the two-tailed test with the alpha level equal to 0.5%, how can you determine the z-scores region to reject the null hypothesis?

The two-tailed test requires you to divide the levels of alpha by 2.

Therefore, α for the two-tailed test = 0.05/2 = 0.0250

As the z-score table shows the area to the left of the value of z, a two-tailed test requires you to identify two entries. The area of one entry covers 0.975 (97.5%) of the area (where 0.025 of the area is outside the value of z on the right tail), and the area of another entry covers 0.025 of the area on the left tail.

Using the z-score table, you can determine the z-score = 1.96 or -1.96. Therefore, you will reject the null hypothesis for obtained z-score > 1.96 or z-score < 1.96.

## 5.2 Test your understanding

The following activity will test your understanding of z-scores and your ability to extract data from z-score tables.

Start of Activity

**Activity 6 Determine the range of z-scores used to reject a null hypothesis**

Allow approximately 45 minutes to complete this activity

This activity focuses on assessing your comprehension of z-scores and your ability to extract data from z-score tables. A z-score is a statistical measure that facilitates the comparison of data points from various data sets.

In this activity, you will receive a set of questions that require you to calculate or extract information from z-score tables. To complete this activity successfully, you will need to have a solid grasp of z-scores and their applications. By participating in this activity, you will not only assess your knowledge of z-scores but also refine your problem-solving abilities and capacity to interpret statistical data. Therefore, prepare yourself to challenge your understanding of z-scores in relation to hypotheses testing and enhance your statistical proficiency.

**Part A**

Start of Question

Using the z-score table that you created in Activity 5, can you determine the range of z-scores used to reject the null hypothesis?

Start of Table

Table 3  Calculating z-scores

|  |  |  |
| --- | --- | --- |
|  | **Alpha levels** | **z-scores** |
| cap h sub a not equals 60 | 0.02 | *Provide your answer...* |
| cap h sub a less than 12 | 0.04 | *Provide your answer...* |
| cap h sub a greater than 96 | 0.10 | *Provide your answer...* |
| cap h sub a not equals 42 | 0.06 | *Provide your answer...* |
| cap h sub a less than 35 | 0.07 | *Provide your answer...* |
| cap h sub a greater than 80 | 0.02 | *Provide your answer...* |

End of Table

Use the free response box below to justify your answers in this exercise.

End of Question

*Provide your answer...*

[View discussion - Part A](" \l "Session7_Discussion2)

**Part B**

Start of Question

Using the Excel formula **NORM.S.DIST(z, cumulative)**, can you determine the alpha levels by completing Table 4 below?

Start of Table

Table 4  Calculating alpha levels

|  |  |  |
| --- | --- | --- |
| **Test** | **z-scores** | **Alpha levels** |
| One-tailed test | 2.35 | *Provide your answer...* |
| Two-tailed test | 1.96 | *Provide your answer...* |
| One-tailed test | 1.65 | *Provide your answer...* |
| Two-tailed test | 1.29 | *Provide your answer...* |

End of Table

End of Question

[View discussion - Part B](" \l "Session7_Discussion3)

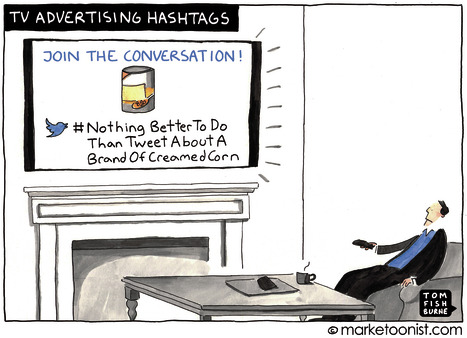
End of Activity

## 5.3 Using the z-score

So far, you have identified the range of z-scores that indicate the reject regions for the hypothesis tests.

For example, records indicate that customers need to be exposed to TV advertising commercials for an average of 10 seconds before being influenced by the advertisement, with a standard deviation of 1.6 seconds. A marketing manager suggests that it will take longer to influence customer behaviour. The z-score table can be used to identify the reject region (for the null hypothesis) in the right tail where the z-score is greater than 1.28. This will test the marketing manager’s claim at a 90% confidence level (α = 0.10) with a sample of 100 customers.

Start of Figure



**Figure 21** Advertisement influence

[View description - Figure 21 Advertisement influence](" \l "Session7_Description12)

[View alternative description - Figure 21 Advertisement influence](" \l "Session7_Alternative28)

End of Figure

However, what does this mean in practice? Marketing managers are not particularly concerned with the z-score indicating the hypothesis’ cut-off region. Rather, they want to know how long television advertisements should be on for on average in order to effectively influence customer behaviour with 90% confidence.

In order to answer this question, you need the formula for calculating the z-score.

Start of $1



[View alternative description - Uncaptioned Equation](" \l "Session7_Alternative29)

End of $1

x macron = sample mean

mu = population mean

sigma = population standard deviation

n = sample size

In the example used above, you know that µ equals 10 (TV advertising commercials are on for an average of 10 seconds before customers are influenced by the advertisement) and σ equals 1.6. You also know that the rejection region of the z-score is > 1.28, and the sample size equals 100 (a sample of 100 customers).

These values can be input into the z-score formula to solve the value of x.

**Step 1:**

Start of $1

1.28 equals x bar minus 10 divided by left parenthesis 1.6 divided by square root of 100 right parenthesis

[View alternative description - Uncaptioned Equation](" \l "Session7_Alternative33)

End of $1

**Step 2:**

Start of $1

 1.28 times 1.6 divided by square root of 100 equals x bar minus 10

[View alternative description - Uncaptioned Equation](" \l "Session7_Alternative34)

End of $1

**Step 3:**

Start of $1

left parenthesis 1.28 times 1.6 divided by square root of 100 right parenthesis plus 10 equals x bar

[View alternative description - Uncaptioned Equation](" \l "Session7_Alternative35)

End of $1

**Step 4:**

x bar equals 10.2

In the end, you can see that the sample mean (x macron) needs to be 10.2. In the decision rule, you can state that in terms of obtaining a z-score of 1.28, you will reject the null hypothesis for any sample mean above 10.2.

In other words, if a marketing manager decides to conduct market research by surveying 100 customers about their responses to TV advertising commercials, they will reject the general belief – customers need to be exposed to TV advertising commercials on average 10 seconds before being influenced by the advertisement – when they find that the average time taken for TV advertising commercials to influence customers in this sample is greater than 10.2 seconds.

In the following activity, you will be tested on your understanding of how to calculate the mean of a sample, which is used to reject the null hypothesis.

Start of Activity

**Activity 7 Determine the range of mean used to reject null hypothesis**

Allow approximately 30 minutes to complete this activity

Start of Question

Calculating the mean is a key statistical analysis technique used to summarise data and draw conclusions. In hypothesis testing, the mean is crucial in determining whether to reject the null hypothesis, which assumes no significant difference between two variables. To reject this hypothesis, a statistical test is used to compare the mean of a sample to a known or expected value.

This activity aims to provide a comprehensive understanding of mean calculation and how to use it to reject the null hypothesis. By the end of the activity, you’ll be able to perform calculations, interpret statistical results and make informed decisions based on the mean.

According to records, customers are persuaded to purchase the products by a price discount of 50% in marketing promotions with a standard deviation of 5.3%. The market research team believes that the company should offer more price discounts (more than 50%) to motivate customers’ purchase intentions. To test this claim, the team will survey 1000 customers. At 99% confidence (significance) levels, can you state the decision rule concerning the value of mean to accept the claim made by the marketing team?

Use the free response box below to show your calculations.

End of Question

*Provide your answer...*

[View discussion - Activity 7 Determine the range of mean used to reject null hypothesis](" \l "Session7_Discussion4)

End of Activity

In conclusion, understanding the range of z-scores and mean calculation in hypothesis testing is essential for drawing accurate statistical conclusions. By calculating the z-score and mean of a sample, you can determine the level of significance and whether the null hypothesis should be rejected or not. Furthermore, understanding how to interpret and use the z-score and mean enables you to make informed decisions based on the results of statistical analysis. By applying these concepts, you can draw meaningful conclusions and make informed decisions in a variety of fields, from healthcare to economics.

## 6 P-value

In the previous sections, you learned about hypothesis testing and the significance level, or α, which determines whether to reject or fail to reject a null hypothesis based on a predetermined level of confidence. However, there is another important concept in hypothesis testing: the p-value.

The ***p-value*** is a statistical measure that helps determine the strength of evidence against the null hypothesis. It is the probability of obtaining a sample mean that is further away from the hypothesised value of µ (a population mean that represents the widely-held belief) specified in the null hypothesis than the value of the sample mean in the study, assuming that the null hypothesis is true.

In other words, the p-value provides a quantitative measure of the strength of evidence against the null hypothesis. This would allow companies to make an informed decision about the effectiveness of marketing campaigns, for instance; it would also help them to avoid making decisions based on chance or random variation in the data, and to make data-driven decisions with more confidence.

## 6.1 Defining the p-value

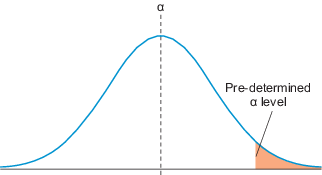
To conduct a hypothesis test using the p-value, you calculate the test statistic, such as a z-score or t-score, and then use it to determine the corresponding p-value from a probability distribution table or statistical software.

The decision rules for using p-values are:

* If p-value ≤ α, you reject the null hypothesis.
* If p-value > α, you fail to reject the null hypothesis (accept the null hypothesis).

You can use graphs to illustrate this decision rule. For example, in a one-tailed test, the orange region indicates the area outside the pre-determining α levels in the graph shown in Figure 22.

Start of Figure



**Figure 22** Alpha level and corresponding region to reject null hypothesis

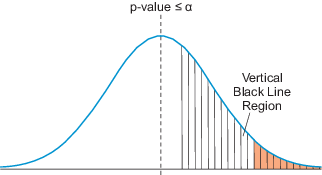
[View description - Figure 22 Alpha level and corresponding region to reject null hypothesis](" \l "Session8_Description1)

[View alternative description - Figure 22 Alpha level and corresponding region to reject null hypothesis](" \l "Session8_Alternative1)

End of Figure

If the p-value ≤ α (i.e. if the boundary of the shaded region falls inside the boundary of α), you will reject the null hypothesis.

Start of Figure



**Figure 23** P-value less than or equal to alpha level

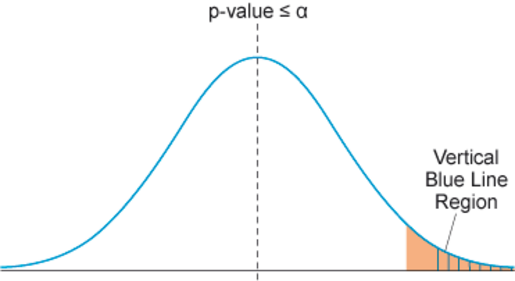
[View description - Figure 23 P-value less than or equal to alpha level](" \l "Session8_Description2)

[View alternative description - Figure 23 P-value less than or equal to alpha level](" \l "Session8_Alternative2)

End of Figure

If the p-value > α (i.e. if the vertical black line region is greater than the orange shaded region), you will not reject the null hypothesis.

Start of Figure



**Figure 24** P-value greater than alpha level

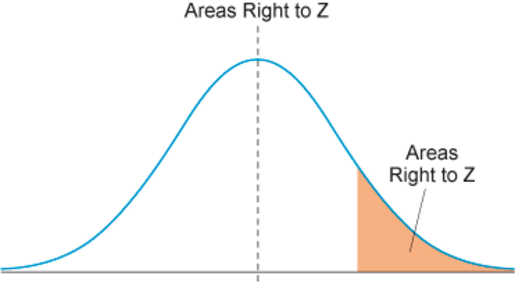
[View description - Figure 24 P-value greater than alpha level](" \l "Session8_Description3)

[View alternative description - Figure 24 P-value greater than alpha level](" \l "Session8_Alternative3)

End of Figure

As you can see from these illustrations, the p-value reflects the tail area of the normal distribution (the vertical blue line region is smaller than the orange-shaded region). Therefore, you can determine the exact p-value by calculating the z-score representing the boundary of the tail area of the normal distribution, and using the pre-determined z-score table to determine the area right of the z-score.

Start of Figure



**Figure 25** Area right of z

[View description - Figure 25 Area right of z](" \l "Session8_Description4)

[View alternative description - Figure 25 Area right of z](" \l "Session8_Alternative4)

End of Figure

## 6.2 Calculating the p-value

To calculate the p-value you must modify the z-score table to highlight the area right of z, instead of the area left of z. The area left of z and the area right of z add up to 100% (confidence level). In order to calculate the area right of z, you simply need to subtract the area left of z from 100%. You can perform this calculation using Excel, as described in the following steps.

**Step 1:** Set up a new table with the exact layout of the previous z-score table used in Activity 5.

Start of Figure



**Figure 26** Creating ‘Area right of z’ table step 1

[View description - Figure 26 Creating ‘Area right of z’ table step 1](" \l "Session8_Description5)

[View alternative description - Figure 26 Creating ‘Area right of z’ table step 1](" \l "Session8_Alternative5)

End of Figure

**Step 2:** Use the Excel function to calculate the values: the z-score in the ‘Area right of z’ table equals 1 minus the corresponding z-score in ‘Area left of z’ table.

Start of Figure



**Figure 27** Creating ‘Area right of z’ table step 2

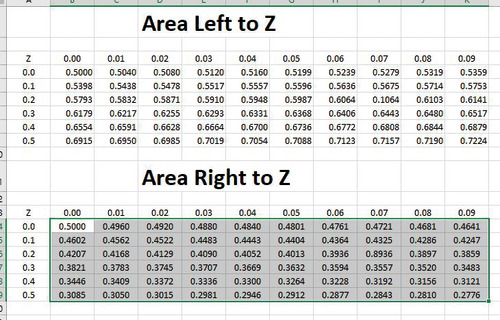
[View description - Figure 27 Creating ‘Area right of z’ table step 2](" \l "Session8_Description6)

[View alternative description - Figure 27 Creating ‘Area right of z’ table step 2](" \l "Session8_Alternative6)

End of Figure

**Step 3:** Apply this calculation to all the cells in the Excel sheet.

Start of Figure



**Figure 28** Creating ‘Area right of z’ table step 3

[View description - Figure 28 Creating ‘Area right of z’ table step 3](" \l "Session8_Description7)

[View alternative description - Figure 28 Creating ‘Area right of z’ table step 3](" \l "Session8_Alternative7)

End of Figure

Using these steps and the z-score table (Figure 28) that highlights the area left of z, you can create a new z-score table to highlight the area right of z..

Start of Table

Table 5   ‘Area right of z’ table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Z | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| -3 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| -2.9 | 0.9981 | 0.9982 | 0.9983 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| -2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9980 | 0.9980 | 0.9981 |
| -2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| -2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| -2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| -2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| -2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| -2.2 | 0.9861 | 0.9865 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| -2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| -2 | 0.9773 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| -1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9762 | 0.9767 |
| -1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9700 | 0.9706 |
| -1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| -1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9485 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| -1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9430 | 0.9441 |
| -1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| -1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| -1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| -1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| -1 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| -0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| -0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7996 | 0.8023 | 0.8051 | 0.8079 | 0.8106 | 0.8133 |
| -0.7 | 0.7580 | 0.7612 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| -0.6 | 0.7258 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7518 | 0.7549 |
| -0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| -0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| -0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| -0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| -0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5754 |
| 0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |

End of Table

Start of Table

Table 5 (continued)   ‘Area right of z’ table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Z | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| 0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |
| 0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3936 | 0.8936 | 0.3897 | 0.3859 |
| 0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| 0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| 0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| 0.6 | 0.2743 | 0.2709 | 0.2676 | 0.2643 | 0.2611 | 0.2578 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| 0.7 | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2296 | 0.2266 | 0.2236 | 0.2206 | 0.2177 | 0.2148 |
| 0.8 | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| 0.9 | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| 1 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| 1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| 1.2 | 0.1151 | 0.1131 | 0.1112 | 0.1093 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| 1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| 1.4 | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0721 | 0.0708 | 0.0694 | 0.0681 |
| 1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| 1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| 1.7 | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| 1.8 | 0.0359 | 0.0351 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| 1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| 2 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| 2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| 2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| 2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| 2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| 2.5 | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| 2.6 | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0041 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| 2.7 | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| 2.8 | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| 2.9 | 0.0019 | 0.0018 | 0.0018 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| 3 | 0.0013 | 0.0013 | 0.0013 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |

End of Table

Using Table 5, you can determine an exact p-value for the sample mean. In order to illustrate this point, try to solve the problems in the following section.

## 6.3 Example: testing a hypothesis

Records shows that customers are willing to pay less than or equal to £2200 for an 8-day premium Iceland winter travel tour, with a standard deviation of £500. A marketing manager believes that customers are willing to pay more for such a tour. To test this belief, the marketing manager asks 45 customers how much they are willing to pay. Below are the responses that the marketing manager received from 45 customers (in the value of £). You now need to determine what the p-values are.

Start of Table

Table 6  Customers’ willingness to pay

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| £2300 | £2200 | £2500 | £2600 | £2100 |
| £2350 | £2450 | £2100 | £2000 | £2150 |
| £2300 | £2400 | £2500 | £2650 | £2750 |
| £2150 | £2800 | £2100 | £2000 | £2600 |
| £2400 | £2000 | £2300 | £2200 | £2100 |
| £2600 | £2500 | £2500 | £2400 | £2300 |
| £2500 | £2200 | £2250 | £2350 | £2550 |
| £2400 | £2400 | £2500 | £2450 | £2350 |
| £2250 | £2450 | £2650 | £2500 | £2400 |

End of Table

**Step 1**: Based on the information statements above, you can formulate the hypotheses:

H0: Customers are willing to pay less than or equal to £2200 for an 8-day premium Iceland winter travel tour.

Ha: Customers are willing to pay more than £2200 for an 8-day premium Iceland winter travel tour.

**Step 2**: Identify all the factors in the z-score formula:

Start of $1



[View alternative description - Uncaptioned Equation](" \l "Session8_Alternative8)

End of $1

mu equals 2200

sigma equals 500

n equals 45

You can also calculate x macron for the data set.

x macron = 2366.67

**Step 3**: Use the z-score formula to calculate z-score.

cap z equals 2366.67 minus 2200 divided by left parenthesis 500 divided by Square root of 45 right parenthesis

cap z equals 2.24

**Step 4** Use the Excel formula **NORM.S.DIST(z, cumulative)**, to obtain a value that reflects the area right of z (p-value)

1. Use the Excel formula ‘NORM.S.DIST(z, cumulative)’ and enter z = 2.24
2. Set ‘cumulative’ to be TRUE
3. Press ‘Enter’, and you will get the value that represents the area left of z equal to 0.9875
4. The p-value = 1 − area left of z
5. p-value = 0.0125

**Step 5:** Interpret the findings and decide whether to reject H0.

So far, you do not have enough information from the problem statement to make this decision. This is because the confidence level (or α) is unknown.

So, more specifically:

* If the company is looking for a 95% confidence level (α = 0.050), as the p-value (0.0125) is less than or equal to α, you will reject H0. This means that the marketing manager’s claim is true that ‘customers are willing to pay more than £2200 for an 8-day premium Iceland winter travel tour’.
* If the company is looking for a 99% confidence level (α = 0.010), as the p-value (0.0125) is greater than α, you will not reject H0. This means that the marketing manager’s claim that ‘customers are willing to pay more than £2200 for an 8-day premium Iceland winter travel tour’ is false.

## 6.4 Test your understanding

The following activity will test your understanding of how to calculate the p-value used to conduct a hypothesis test.

Start of Activity

**Activity 8 Determine p-value and testing hypothesis**

Allow approximately 45 minutes to complete this activity

Start of Question

Records show that customers are willing to pay less than or equal to £20 for a premium box of chocolate with a standard deviation of 5.5. A marketing manager believes that customers are willing to pay more than this. The marketing manager asks 40 customers how much they are willing to pay to test this belief. Table 7 shows the responses the marketing manager received from the 40 customers (in the value of £).

At the 95% confidence level, can you determine whether the marketing manager’s claim is true or false? You can copy-paste the data below into Excel to calculate the required variable(s) or use a calculator of your choice.

Start of Table

Table 7  Customer responses

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 23 | 22 | 25 | 18 | 21 |
| 24 | 25 | 21 | 20 | 20 |
| 18 | 21 | 17 | 25 | 26 |
| 25 | 15 | 17 | 13 | 18 |
| 21 | 23 | 25 | 24 | 30 |
| 21 | 18 | 19 | 17 | 22 |
| 23 | 21 | 25 | 24 | 22 |
| 26 | 18 | 20 | 20 | 25 |

End of Table

Use the free response box below to show your calculations.

End of Question

*Provide your answer...*

[View discussion - Activity 8 Determine p-value and testing hypothesis](" \l "Session8_Discussion1)

End of Activity

In conclusion, p-values are used in hypothesis testing to determine whether results are likely to be as extreme as the observed ones. The null hypothesis should be rejected in favour of the alternative hypothesis if the p-value is less than the significance level. In light of the evidence gathered from their study, managers can make informed decisions based on the p-value. They are able to determine whether their results are statistically significant or merely a result of chance based on their confidence in the results.

## 7 Hypothesis testing for population proportions

In the previous sections, all the problems you encountered involved using population means and sample means. The z-score test can also be used to solve another type of problem – those related to population proportions. Proportions, or relative frequencies, can be determined by dividing the number of items in a specific group or category by the total number of items in the sample. This calculation yields a representation of the quantity of items belonging to a particular group or category, expressed as either a fraction or percentage. In statistical and data analysis, proportions are frequently employed to characterise variable distributions or to contrast distinct groups or categories within a sample.

Start of Figure



**Figure 29** Data made up

[View description - Figure 29 Data made up](" \l "Session9_Description1)

[View alternative description - Figure 29 Data made up](" \l "Session9_Alternative1)

End of Figure

Imagine, for example, that a marketing manager surveyed 140 customers and found that 75 customers have been using the company’s service for more than five years. In that case, you can calculate the proportion of loyal customers.

Start of $1

Proportion of loyal customers equals number of customers using the company apostrophe s service divided by number of customers in the sample

[View alternative description - Uncaptioned Equation](" \l "Session9_Alternative2)

End of $1

Start of $1

equals 75 divided by 140

[View alternative description - Uncaptioned Equation](" \l "Session9_Alternative3)

End of $1

Start of $1

equals 53.6 percent

[View alternative description - Uncaptioned Equation](" \l "Session9_Alternative4)

End of $1

By using the z-score test, you can address the proportion problem by determining the p-value and deciding whether to reject the null hypothesis. Nevertheless, you will need to use a slightly different formula to calculate the z-score for the problem related to population and the sample proportion.

In the next section, you will look at a problem to illustrate this.

## 7.1 Example: testing a proportion

‘A company believes that the percentage of people who subscribe to its mobile phone service is less than 25%. A sales manager disagrees with this because he conducted a study of 2000 people and found 540 of them subscribe to the company’s mobile phone services. At a 5% significance level, is there enough evidence to support this claim?’

**Step 1:** State the hypotheses.

The first step in addressing all the issues related to hypothesis tests is to formulate the null and alternative hypotheses. Unlike the example used in previous sections, which used the population mean (µ), here you need to use the population proportion (p).

H0: the percentage of people that subscribe to the company’s mobile phone service is 25% or less.

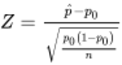
Ha: the percentage of people that subscribe to the company’s mobile phone service is more than 25%.

This can also be written as:

H0: p ≤ 25% (or 0.25)

Ha: p > 25%

**Step 2:** Use a z-score formula to calculate the z-score.



p hat = sample (observed) proportion = 0.27

p sub zero = population proportion = 0.25

n = sample size = 2000

Using these values, you can calculate the z-score:

equation sequence part 1 cap z equals part 2 0.27 minus 0.25 divided by Square root of 0.25 times left parenthesis one minus 0.25 right parenthesis divided by 2000 equals part 3 2.01

**Step 3:** Determine the p-value using the z-scores in Table 5 or NORM.S.DIST() in Excel.

Judging from the null and alternative hypotheses, this is a one-tailed test on the right (upper) tail.

In Table 5, you can find that the area to the right of z corresponding to a z-score of 2.01 is 0.0222.

Thus, the p-value = 2.22%

**Step 4:** Make a decision based on the p-value.

According to the problem statement, you want to test the sales manager’s claim at 5% level of significance.

The p-value from the calculation = 2.22% < 5%

Thus, you reject the null hypothesis and accept the sales manager’s claim that the percentage of people that subscribe to the company’s mobile phone service is more than 25%.

## 7.2 Test your understanding

In this section you will try a problem on your own. Using the following activity, you can test your understanding of how to calculate the z-score for a problem related to the population and the sample proportion.

Start of Activity

**Activity 9 Testing proportion hypothesis**

Allow approximately 45 minutes to complete this activity

Start of Question

According to records, 70% of customers are willing to spend £15 on a 3D movie ticket. A ticket agent believes that this value is different. He surveys 200 individuals and found that 128 stated their willingness to spend a different amount on a ticket to a 3D movie. At the 95% confidence level, is there enough evidence to reject the null hypothesis?

Use the free response box below to show your calculations.

End of Question

*Provide your answer...*

[View discussion - Activity 9 Testing proportion hypothesis](" \l "Session9_Discussion1)

End of Activity

This section aims to demonstrate how to use the z-score to make inferences about the population proportion based on sample data. When the population proportion is known, you can calculate the z-score to determine the likelihood of observing a particular sample proportion.

## Conclusion

The purpose of this course was to discuss hypotheses testing. Through the activities, you have gained a better understanding of the concept of alpha (α). You have learned the difference between a one-tailed test and a two-tailed test. Additionally, you have learned how to calculate z-scores and p-values as well as how to use them to determine whether null hypotheses should be accepted or rejected. Finally, the end of this course helped you gain an understanding of how to conduct hypothesis testing for population proportions.

A second OpenLearn course on data analysis, [Data analysis: visualisations in Excel](https://www.open.edu/openlearn/science-maths-technology/data-analysis-visualisations-excel/content-section-0), is now also available should you wish to take your studies further.

This OpenLearn course is an adapted extract from the Open University course [B126 Business data analytics and decision making.](https://www.open.ac.uk/courses/modules/b126)

## References

Warner, R. M. (2021) Applied statistics: From bivariate through multivariate techniques. 3rd edn. Thousand Oaks, CA: Sage Publications.

## Acknowledgements

This free course was written by Henry Lahr.

Except for third party materials and otherwise stated (see [terms and conditions](http://www.open.ac.uk/conditions)), this content is made available under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 Licence](http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en).

The material acknowledged below is Proprietary and used under licence (not subject to Creative Commons Licence). Grateful acknowledgement is made to the following sources for permission to reproduce material in this free course:

Course image: Image by congerdesign from Pixabay

Figure 1: Dominique Deckmyn / www.cartoonstock.com

Figure 2: Image by Jan Vaek from Pixabay

Figure 3: Image by WikiImages from Pixabay

Figure 4: Image by Tesa Robbins from Pixabay

Figure 5: Mike Seddon / www.cartoonstock.com

Figure 21: Tom Fishburne / Marketoonist.com

Figure 29: Tom Fishburne / Marketoonist.com

Every effort has been made to contact copyright owners. If any have been inadvertently overlooked, the publishers will be pleased to make the necessary arrangements at the first opportunity.

**Don't miss out**

If reading this text has inspired you to learn more, you may be interested in joining the millions of people who discover our free learning resources and qualifications by visiting The Open University – [www.open.edu/openlearn/free-courses](http://www.open.edu/openlearn/free-courses?LKCAMPAIGN=ebook_&MEDIA=ol).

## Solutions

## Activity 1 Null hypothesis versus alternative hypothesis

#### Discussion

H0: a coffee machine makes a cup of caffè latte with 1cm foam on average.

Ha: a coffee machine cannot make a cup of caffè latte with 1 cm foam on average.

[Back to - Activity 1 Null hypothesis versus alternative hypothesis](" \l "Session3_Activity1)

## Activity 2  Level of alpha

#### Discussion

Start of Table

Table 1  Level of alpha (completed)

|  |  |
| --- | --- |
| **Level of Confidence (C)** | **α** |
| 90% | 10% |
| 95% | 5% |
| 99% | 1% |

End of Table

α is calculated as 1 − C.

1 − 90% = 10%

1 − 95% = 5%

1 − 99% = 1%

[Back to - Activity 2  Level of alpha](" \l "Session5_Activity1)

## Activity 3 Hypotheses setting

#### Discussion

H0: OU students spend, on average, no more than 15 hours studying an entire session of OU course.

Ha: OU students spend, on average, more than 15 hours studying an entire session of OU course.

They can also be written as:

H0: µ ≤ 15 hours studies

Ha: µ > 15 hours studies

µ is a symbol for a population mean. Remember, H0 and Ha are always opposites.

[Back to - Activity 3 Hypotheses setting](" \l "Session6_Activity1)

## Activity 5  Z-Score table

#### Discussion

The following table illustrates a z-score that ranges from -3 to 3. The following is what you should have accomplished as a result of this activity.

Start of Table

Table 2  A z-score table with ranges −3 to 3

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Z** | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| **−3.0** | 0.0014 | 0.0013 | 0.0013 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |
| **−2.9** | 0.0019 | 0.0018 | 0.0018 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| **−2.8** | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| **−2.7** | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| **−2.6** | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0042 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| **−2.5** | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| **−2.4** | 0.0082 | 0.0080 | 0.0078 | 0.0076 | 0.0073 | 0.0071 | 0.0070 | 0.0068 | 0.0066 | 0.0064 |
| **−2.3** | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| **−2.2** | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0126 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| **−2.1** | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| **−2.0** | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| **−1.9** | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| **−1.8** | 0.0359 | 0.0352 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| **−1.7** | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| **−1.6** | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| **−1.5** | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| **−1.4** | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0722 | 0.0708 | 0.0694 | 0.0681 |
| **−1.3** | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| **−1.2** | 0.1151 | 0.1131 | 0.1112 | 0.1094 | 0.1075 | 0.1057 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| **−1.1** | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| **−1.0** | 0.1587 | 0.1563 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| **−0.9** | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| **−0.8** | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| **−0.7** | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2297 | 0.2266 | 0.2236 | 0.2207 | 0.2177 | 0.2148 |
| **−0.6** | 0.2743 | 0.2709 | 0.2676 | 0.2644 | 0.2611 | 0.2579 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| **−0.5** | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| **−0.4** | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| **−0.3** | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| **−0.2** | 0.4207 | 0.4168 | 0.4129 | 0.4091 | 0.4052 | 0.4013 | 0.3974 | 0.3936 | 0.3897 | 0.3859 |
| **−0.1** | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| **−0.0** | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4841 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |
| **−0.0** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.1** | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| **0.2** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6064 | 0.1064 | 0.6103 | 0.6141 |
| **0.3** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| **0.4** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| **0.5** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| **0.6** | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| **0.7** | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| **0.8** | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| **0.9** | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| **1.0** | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| **1.1** | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| **1.2** | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| **1.3** | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| **1.4** | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| **1.5** | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| **1.6** | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| **1.7** | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| **1.8** | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| **1.9** | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| **2.0** | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| **2.1** | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| **2.2** | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| **2.3** | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| **2.4** | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| **2.5** | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| **2.6** | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| **2.7** | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| **2.8** | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| **2.9** | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| **3.0** | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |

End of Table

[Back to - Activity 5  Z-Score table](" \l "Session7_Activity1)

## Activity 6 Determine the range of z-scores used to reject a null hypothesis

### Part A

#### Discussion

Here are the answers:

Start of Table

Table 3  Calculating z-scores (completed)

|  |  |  |
| --- | --- | --- |
|  | **Alpha levels** | **z-scores** |
| cap h sub a not equals 60 left parenthesis cap h sub zero equals 60 right parenthesis | 0.02 (Area left of z: 0.01 and 0.99) | ± 2.33 |
| cap h sub a less than 12 left parenthesis cap h sub zero greater than or equals 12 right parenthesis | 0.04 (Area left of z: 0.04) | −1.75 |
| cap h sub a greater than 96 left parenthesis cap h sub zero less than or equals 96 right parenthesis | 0.10 (Area left of z: 0.90) | 1.28 |
| cap h sub a not equals 42 left parenthesis cap h sub zero equals 42 right parenthesis | 0.06 (Area left of z: 0.03 and 0.97) | ±1.88 |
| cap h sub a less than 35 left parenthesis cap h sub zero greater than or equals 35 right parenthesis | 0.07 (Area left of z: 0.07) | −1.48 |
| cap h sub a greater than 80 left parenthesis cap h sub zero less than or equals 80 right parenthesis | 0.02 (Area left of z: 0.98) | 2.06 |

End of Table

[Back to - Part A](" \l "Session7_Part1)

### Part B

#### Discussion

There are several things that you need to pay attention to when answering these questions.

1. The values indicated in the hypothesis statement will not affect the determination of the z-score.
2. The table states the alternative hypothesis, therefore you must translate it into the null hypothesis before answering the question.
3. For the two-tailed test, you need to divide the alpha level by 2.
4. Alpha levels are stated in the table, so you must translate them to the area left of z before using the z-score table.

Here are the answers

Start of Table

Table 4  Calculating alpha levels (completed)

|  |  |  |
| --- | --- | --- |
| **Test** | **z-scores** | **Alpha levels** |
| One-tailed test | 2.35 | tilde operator 0.01 |
| Two-tailed test | 1.96 | tilde operator 0.01 |
| One-tailed test | 1.65 | tilde operator 0.05 |
| Two-tailed test | 1.29 | tilde operator 0.05 |

End of Table

[Back to - Part B](" \l "Session7_Part2)

## Activity 7 Determine the range of mean used to reject null hypothesis

#### Discussion

**Step 1**: State the hypotheses.

H0: Customers will be persuaded to purchase the products by a price discount of less than or equal to 50% in a marketing promotion (µ ≤ 50% price discount).

Ha: Customers will be persuaded to purchase the products with a price discount of more than 50% in a marketing promotion (µ > 50% price discount).

**Step 2**: Identify the z-score and rejection region.

Using the z-score table, you can identify the z-score for a one-tailed test with 99% significant levels. In this case, the z-score equals 2.33.

**Step 3**: Identify all the values.

mu equals 0.5

sigma equals 0.053

n equals 1000

z equals 2.33

**Step 4**: Solve the x macron from the formula 

If 2.33 equals x bar minus 0.5 divided by left parenthesis 0.053 divided by square root of 1000 right parenthesis

Then multirelation x macron almost equals 0.504 equals 50.4 percent

**Step 5**: Write the decision rule concerning the value of mean.

The market research team will reject the general belief – customers will be persuaded to purchase the products by less than or equal to 50% price discount in marketing promotion – when survey results (from 1000 customers) show that customers are persuaded to purchase the products when price discount promotion is greater than 50%.

[Back to - Activity 7 Determine the range of mean used to reject null hypothesis](" \l "Session7_Activity3)

## Activity 8 Determine p-value and testing hypothesis

#### Discussion

**Step 1:** Based on the information statements above, you can formulate the hypotheses:

H0: customers are willing to pay less than or equal to £20 for a premium box of chocolate.

Ha: customers are willing to pay more than £20 for a premium box of chocolate.

**Step 2:** Identify all the factors in the z-score formula.

mu equals 20

sigma equals 5.5

n equals 40

You can also calculate x macron for the data set.

x macron equals 21.45

**Step 3**: Use the z-score formula to calculate the z-score.

cap z equals 21.45 minus 20 divided by left parenthesis 5.5 divided by Square root of 40 right parenthesis

cap z equals 1.67

**Step 4** Use the Excel formula ‘NORM.S.DIST(z, cumulative)’ to obtain a value that reflects the area right of z (p-value).

When z = 1.67, p-value = 0.0475

**Step 5:** Interpret the findings and decide whether to reject H0.

You have been told the company is looking for a 95% confidence level (α = 0.050).

As the p-value (0.0475) is less than or equal to α, you will reject H0. This means that the marketing manager’s claim is true that ‘customers are willing to pay more than £20 for a premium box of chocolate’.

[Back to - Activity 8 Determine p-value and testing hypothesis](" \l "Session8_Activity1)

## Activity 9 Testing proportion hypothesis

#### Discussion

**Step 1:** State the hypotheses.

The first step to tackle all the problems concerning hypothesis tests is to state the null and alternative hypotheses.

H0: 70% of customers are willing to spend £15 on a ticket to a 3D movie.

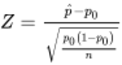
Ha: 70% of customers are not willing to spend £15 on a ticket to a 3D movie.

Or it can be written as

H0: p = 70% (or 0.7)

Ha: p ± 70%

**Step 2:** Use a z-score formula to calculate the z-score.



p hat = sample (observed) proportion = 128 divided by 200 = 0.64

p0 = population proportion = 0.70

n = sample size = 200

Using these values, you can calculate the z-score.

equation sequence part 1 cap z equals part 2 0.64 minus 0.70 divided by Square root of 0.70 times left parenthesis one minus 0.70 right parenthesis divided by 200 equals part 3 negative 1.85

**Step 3:** Determine p-value using the Excel formula ‘NORM.S.DIST(z, cumulative)’

Judging from the null and alternative hypotheses, this is a two-tailed test.

Therefore, you need to divide the levels of significance by 2.

At 95% confidence level, the significance level equals 1 − 95%. Thus, α = 0.05

0.05 divided by two equals 0.025.

In order to reject the null hypothesis at 95% confidence levels in a two-tailed test, the p-value needs to be less than 0.025.

Given the z-score is negative, you can use Table  5 to determine the p-value. You can find the area left of z corresponding to a z-score of 1.85 is 0.0322.

**Step 4:** Make a decision based on the p-value.

The p-value from the calculation = 0.0322, which is ≥ 0.025.

Thus, you will not reject the null hypothesis that claims 70% of customers are willing to spend £15 on a ticket to a 3D movie.

Just a quick note: if this is not a two-tailed test (i.e. if it is a one tailed-test, associated with the left tail), you will not divide the significance levels by 2. In this situation, you will still use α = 0.05. So, as 0.0322 (p-value) < 0.025 (α), you would then reject H0.

[Back to - Activity 9 Testing proportion hypothesis](" \l "Session9_Activity1)

## Descriptions

### Figure 1 Data sheds light on hypotheses

Cartoon of a lecture with a tag line

[Back to - Figure 1 Data sheds light on hypotheses](" \l "Session1_Figure1)

### Figure 1 Data sheds light on hypotheses

The picture portrays a professor conducting an open lecture on the moon, with a tag line displayed at the bottom which states, ‘Surprisingly, the soil samples provide new evidence for the “giant ball of cheese” hypothesis.’

[Back to - Figure 1 Data sheds light on hypotheses](#Session1_Figure1)

### Figure 2 iPhone

A hand holding an Apple iPhone.

[Back to - Figure 2 iPhone](" \l "Session3_Figure1)

### Figure 2 iPhone

A hand holding an Apple iPhone.

[Back to - Figure 2 iPhone](#Session3_Figure1)

### Figure 3 Solar system

Planets revolving around the Sun

[Back to - Figure 3 Solar system](" \l "Session3_Figure2)

### Figure 3 Solar system

Planets revolving around the Sun

[Back to - Figure 3 Solar system](#Session3_Figure2)

### Figure 4 Caffè latte

A cup of caffè latte with milk foam.

[Back to - Figure 4 Caffè latte](" \l "Session4_Figure1)

### Figure 4 Caffè latte

A cup of caffè latte with milk foam.

[Back to - Figure 4 Caffè latte](#Session4_Figure1)

### Figure 5 The peak point is average

Cartoon showing a bell-shaped curve.

[Back to - Figure 5 The peak point is average](" \l "Session6_Figure1)

### Figure 5 The peak point is average

The cartoon shows a bell-shaped curve. The x-axis is titled ‘How high the hill is’ and the y-axis is titled ‘Number of hills’. The top of the bell-shaped curve is labelled ‘Average hill’, but on the lower right tail of the bell-shaped curve is labelled ‘Big hill’.

[Back to - Figure 5 The peak point is average](#Session6_Figure1)

### Figure 6 Normal distribution – bell curve

A symmetrical graph reminiscent of a bell showing normal distribution.

[Back to - Figure 6 Normal distribution – bell curve](" \l "Session6_Figure2)

### Figure 6 Normal distribution – bell curve

A symmetrical graph reminiscent of a bell. The top of the bell-shaped curve appears where the x-axis is at 0. This is labelled as Normal distribution.

[Back to - Figure 6 Normal distribution – bell curve](#Session6_Figure2)

### Figure 7 Two-tailed test

A two-tailed test shown in a symmetrical graph reminiscent of a bell

[Back to - Figure 7 Two-tailed test](" \l "Session6_Figure3)

### Figure 7 Two-tailed test

A symmetrical graph reminiscent of a bell. The x-axis is labelled ‘z-score’ and the y-axis is labelled ‘probability density’. The x-axis increases in increments of 1 from -2 to 2.

The top of the bell-shaped curve is labelled ‘Foam height = 1cm’. The graph circles the rejection regions of the null hypothesis on both sides of the bell curve. Within these circles are two areas shaded orange: beneath the curve from -2 downwards which is labelled z < -1.96 and α = 0.025; and beneath the curve from 2 upwards which is labelled z > 1.96 and α = 0.025.

[Back to - Figure 7 Two-tailed test](#Session6_Figure3)

### Figure 8 One-tailed test – right (upper) tail

A one tailed test shown in a symmetrical graph reminiscent of a bell

[Back to - Figure 8 One-tailed test – right (upper) tail](" \l "Session6_Figure4)

### Figure 8 One-tailed test – right (upper) tail

A symmetrical graph reminiscent of a bell. The x-axis is labelled ‘z-score’ and the y-axis is labelled ‘probability density’. The x-axis increases in increments of 1 from -2 to 2.

The top of the bell-shaped curve is labelled ‘Hours of study = 15 hours’. The graph circles the rejection regions of the null hypothesis on the right hand side of the bell curve. Within this circle is an area shaded orange which is labelled z > 1.645 and α = 0.05.

[Back to - Figure 8 One-tailed test – right (upper) tail](#Session6_Figure4)

### Figure 9 One-tailed test – left (lower) tail

A one tailed test shown in a symmetrical graph reminiscent of a bell

[Back to - Figure 9 One-tailed test – left (lower) tail](" \l "Session6_Figure5)

### Figure 9 One-tailed test – left (lower) tail

A symmetrical graph reminiscent of a bell. The x-axis is labelled ‘z-score axis’ and the y-axis is labelled ‘customer spending axis’. The x-axis increases in increments of 1 from -2 to 2.

The top of the bell shaped curve is labelled ‘Customer spending = £65’. The graph circles the rejection regions of the null hypothesis on left hand side of the bell curve. Within this circle is an area shaded orange which is labelled z < -1.645 and α= 0.05.

[Back to - Figure 9 One-tailed test – left (lower) tail](#Session6_Figure5)

### Figure 10 Creating a z-score table step 1

A z-score table with no entries

[Back to - Figure 10 Creating a z-score table step 1](" \l "Session7_Figure1)

### Figure 10 Creating a z-score table step 1

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09).

[Back to - Figure 10 Creating a z-score table step 1](#Session7_Figure1)

### Figure 11 Creating a z-score table step 2

A z-score table showing the entry of Excel formula ‘NORM.S.DIST(z, cumulative)’

[Back to - Figure 11 Creating a z-score table step 2](" \l "Session7_Figure2)

### Figure 11 Creating a z-score table step 2

A picture of a z-table with rows with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). It shows the entry of Excel formula ‘NORM.S.DIST(z, cumulative)’.

[Back to - Figure 11 Creating a z-score table step 2](#Session7_Figure2)

### Figure 12 Creating a z-score table step 3

A z-score table showing the entry of Excel formula ‘NORM.S.DIST($B4+C$3’

[Back to - Figure 12 Creating a z-score table step 3](" \l "Session7_Figure3)

### Figure 12 Creating a z-score table step 3

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). It shows the entry of value in the Excel formula ‘NORM.S.DIST($B4+C$3’.

[Back to - Figure 12 Creating a z-score table step 3](#Session7_Figure3)

### Figure 13 Creating a z-score table step 4

A z-score table showing the entry of Excel formula ‘NORM.S.DIST($B4+C$3’ and selecting ‘TRUE - cumulative distribution function’

[Back to - Figure 13 Creating a z-score table step 4](" \l "Session7_Figure4)

### Figure 13 Creating a z-score table step 4

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). It shows the entry of value in the Excel formula ‘NORM.S.DIST($B4+C$3’ and selecting ‘TRUE - cumulative distribution function’.

[Back to - Figure 13 Creating a z-score table step 4](#Session7_Figure4)

### Figure 14 Creating a z-score table step 4

A z-score table showing the entry of value in the Excel formula ‘NORM.S.DIST($B4+C$3", TRUE)’

[Back to - Figure 14 Creating a z-score table step 4](" \l "Session7_Figure5)

### Figure 14 Creating a z-score table step 4

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). It shows the entry of value in the Excel formula ‘NORM.S.DIST($B4+C$3", TRUE)’.

[Back to - Figure 14 Creating a z-score table step 4](#Session7_Figure5)

### Figure 15 Creating a z-score table step 5

A z-score table displaying the result 0.5000

[Back to - Figure 15 Creating a z-score table step 5](" \l "Session7_Figure6)

### Figure 15 Creating a z-score table step 5

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). After the calculation, it displays the result (0.5000).

[Back to - Figure 15 Creating a z-score table step 5](#Session7_Figure6)

### Figure 16 Creating a z-score table step 5

A z-score table displaying all the results

[Back to - Figure 16 Creating a z-score table step 5](" \l "Session7_Figure7)

### Figure 16 Creating a z-score table step 5

A picture of a z-table with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09). All of the results are displayed after the calculation has been completed as follows:

Start of Table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Z Score** | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| **0.0** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.1** | 0.5398 | 0.5438 | 0.5478 | 0.05517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| **0.2** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| **0.3** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| **0.4** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| **0.5** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

End of Table

[Back to - Figure 16 Creating a z-score table step 5](#Session7_Figure7)

### Figure 17 Area under the curve left of z

A symmetrical graph resembling a bell. Areas left of z are coloured orange in the graph.

[Back to - Figure 17 Area under the curve left of z](" \l "Session7_Figure8)

### Figure 17 Area under the curve left of z

A symmetrical graph resembling a bell. Areas left of z are coloured orange in the graph.

[Back to - Figure 17 Area under the curve left of z](#Session7_Figure8)

### Figure 18 One-tailed test with Alpha level of 1%

A symmetrical graph reminiscent of a bell showing the z-score azis and the rejection regions of null hypothesis

[Back to - Figure 18 One-tailed test with Alpha level of 1%](" \l "Session7_Figure9)

### Figure 18 One-tailed test with Alpha level of 1%

A symmetrical graph reminiscent of a bell. The graph points out z-score axis. Areas left of z are coloured orange in the graph. It also circles the rejection regions of null hypothesis when z = 2.33 and alpha = 0.01.

[Back to - Figure 18 One-tailed test with Alpha level of 1%](#Session7_Figure9)

### Figure 19 Calculate the region left of z without using z-score table 1

A table showing the entry of Excel formula and value ‘NORM.S.DIST(2.33, TRUE)’

[Back to - Figure 19 Calculate the region left of z without using z-score table 1](" \l "Session7_Figure10)

### Figure 19 Calculate the region left of z without using z-score table 1

A picture of a table made in Excel. It shows the entry of Excel formula and value ‘NORM.S.DIST(2.33, TRUE)’.

[Back to - Figure 19 Calculate the region left of z without using z-score table 1](#Session7_Figure10)

### Figure 20 Calculate the region left of z without using z-score table 2

A table displaying the result 0.9901

[Back to - Figure 20 Calculate the region left of z without using z-score table 2](" \l "Session7_Figure11)

### Figure 20 Calculate the region left of z without using z-score table 2

A picture of a table made in Excel. After the calculation, it displays the result (0.9901).

[Back to - Figure 20 Calculate the region left of z without using z-score table 2](#Session7_Figure11)

### Figure 21 Advertisement influence

Cartoon showing a TV advert for a can of corn

[Back to - Figure 21 Advertisement influence](" \l "Session7_Figure12)

### Figure 21 Advertisement influence

This figure shows a TV ad for a can of corn, inviting viewers to ‘Join in the conversation’; the tag line reads ‘#Nothing Better TO DO Than Tweet About A Brand Of Creamed Corn’.

[Back to - Figure 21 Advertisement influence](#Session7_Figure12)

### Uncaptioned Equation

cap z equals x bar minus mu divided by left parenthesis sigma divided by square root of n right parenthesis

[Back to - Uncaptioned Equation](" \l "Session7_Equation1)

### Uncaptioned Equation

1.28 equals x bar minus 10 divided by left parenthesis 1.6 divided by square root of 100 right parenthesis

[Back to - Uncaptioned Equation](" \l "Session7_Equation2)

### Uncaptioned Equation

1.28 times 1.6 divided by square root of 100 equals x bar minus 10

[Back to - Uncaptioned Equation](" \l "Session7_Equation3)

### Uncaptioned Equation

left parenthesis 1.28 times 1.6 divided by square root of 100 right parenthesis plus 10 equals x bar

[Back to - Uncaptioned Equation](" \l "Session7_Equation4)

### Figure 22 Alpha level and corresponding region to reject null hypothesis

A symmetrical graph reminiscent of a bell. The graph points out a pre-determined α level.

[Back to - Figure 22 Alpha level and corresponding region to reject null hypothesis](" \l "Session8_Figure1)

### Figure 22 Alpha level and corresponding region to reject null hypothesis

A symmetrical graph reminiscent of a bell. The graph points out a pre-determined α level.

[Back to - Figure 22 Alpha level and corresponding region to reject null hypothesis](#Session8_Figure1)

### Figure 23 P-value less than or equal to alpha level

A symmetrical graph reminiscent of a bell showing the vertical black line region is greater than the orange shaded region.

[Back to - Figure 23 P-value less than or equal to alpha level](" \l "Session8_Figure2)

### Figure 23 P-value less than or equal to alpha level

A symmetrical graph reminiscent of a bell. The graph points out that the vertical black line region is greater than the orange shaded region.

[Back to - Figure 23 P-value less than or equal to alpha level](#Session8_Figure2)

### Figure 24 P-value greater than alpha level

A symmetrical graph reminiscent of a bell showing the vertical blue line region is smaller than the orange-shaded region.

[Back to - Figure 24 P-value greater than alpha level](" \l "Session8_Figure3)

### Figure 24 P-value greater than alpha level

A symmetrical graph reminiscent of a bell. The graph points out that the vertical blue line region is smaller than the orange-shaded region.

[Back to - Figure 24 P-value greater than alpha level](#Session8_Figure3)

### Figure 25 Area right of z

A symmetrical graph reminiscent of a bell. The graph points out areas right of z in orange

[Back to - Figure 25 Area right of z](" \l "Session8_Figure4)

### Figure 25 Area right of z

A symmetrical graph reminiscent of a bell. The graph points out areas right of z in orange colour.

[Back to - Figure 25 Area right of z](#Session8_Figure4)

### Figure 26 Creating ‘Area right of z’ table step 1

A table showing ‘Area right of z’

[Back to - Figure 26 Creating ‘Area right of z’ table step 1](" \l "Session8_Figure5)

### Figure 26 Creating ‘Area right of z’ table step 1

A picture of an ‘Area left to z’ table on the top and an ‘Area right to z’ table below with rows (labelled: 0.0; 0.1; 0.2; 0.3; 0.4; 0.5) and columns (labelled: 0.00; 0.01; 0.02; 0.03; 0.04, 0.05; 0.06; 0.07; 0.08; 0.09).

The ‘Area left to z’ table is filled out as follows:

Start of Table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Z Score** | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| **0.0** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.1** | 0.5398 | 0.5438 | 0.5478 | 0.05517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| **0.2** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| **0.3** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| **0.4** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| **0.5** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

End of Table

[Back to - Figure 26 Creating ‘Area right of z’ table step 1](#Session8_Figure5)

### Figure 27 Creating ‘Area right of z’ table step 2

A table showing ‘Area left to z’ and ‘Area right to z’

[Back to - Figure 27 Creating ‘Area right of z’ table step 2](" \l "Session8_Figure6)

### Figure 27 Creating ‘Area right of z’ table step 2

In the image above, we have an ‘Area left to z’ table on the top and an ‘Area right to z’ table below with the equation 1 minus values from the area left to z table.

The ‘Area left to z’ table is filled out as follows:

Start of Table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Z Score** | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| **0.0** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.1** | 0.5398 | 0.5438 | 0.5478 | 0.05517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| **0.2** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| **0.3** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| **0.4** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| **0.5** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

End of Table

The formula ‘=1-B4’ appears in the ‘Area right to z’ table in the cell for row 0.0, column 0.00.

[Back to - Figure 27 Creating ‘Area right of z’ table step 2](#Session8_Figure6)

### Figure 28 Creating ‘Area right of z’ table step 3

A table showing ‘Area left to z’ and ‘Area right to z’

[Back to - Figure 28 Creating ‘Area right of z’ table step 3](" \l "Session8_Figure7)

### Figure 28 Creating ‘Area right of z’ table step 3

In the image above, we have an ‘Area left to z’ table on the top and an ‘Area right to z’ table below.

The ‘Area left to z’ table is filled out as follows:

Start of Table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Z Score** | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| **0.0** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.1** | 0.5398 | 0.5438 | 0.5478 | 0.05517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| **0.2** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| **0.3** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| **0.4** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| **0.5** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

End of Table

The ‘Area right to z’ is completed as follows:

Start of Table

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Z | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| 0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |
| 0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3936 | 0.8936 | 0.3897 | 0.3859 |
| 0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| 0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| 0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |

End of Table

[Back to - Figure 28 Creating ‘Area right of z’ table step 3](#Session8_Figure7)

### Uncaptioned Equation

cap z equals x bar minus mu divided by left parenthesis sigma divided by square root of n right parenthesis

[Back to - Uncaptioned Equation](" \l "Session8_Equation1)

### Figure 29 Data made up

Cartoon showing a pie chart in a meeeting room with a tag line

[Back to - Figure 29 Data made up](" \l "Session9_Figure1)

### Figure 29 Data made up

A pie chart in front of the meeting room. The tag line says ‘86.4% of people will believe any data you put in a PowerPoint side, even it you just totally made it up to prove your point’.

[Back to - Figure 29 Data made up](#Session9_Figure1)

### Uncaptioned Equation

Proportion of loyal customers equals number of customers using the company apostrophe s service divided by number of customers in the sample

[Back to - Uncaptioned Equation](" \l "Session9_Equation1)

### Uncaptioned Equation

equals 75 divided by 140

[Back to - Uncaptioned Equation](" \l "Session9_Equation2)

### Uncaptioned Equation

equals 53.6 percent

[Back to - Uncaptioned Equation](" \l "Session9_Equation3)