# The Normal Distribution

Sue Gordon

**Mathematics Learning Centre**
**University of Sydney**
**NSW 2006**

**Acknowledgements**

I would like to thank Jackie Nicholas for all her contributions including many ideas, examples and exercises as well as editing and suggestions for improvement. Jackie also did the LaTeX typesetting and drew the graphs.

Parts of this booklet are based on an earlier Mathematics Learning Centre booklet by Peter Petocz. I gratefully acknowledge Peter's ideas.

Sue Gordon
2006

# Contents

# 1 Introduction

## 1.1 The Normal curve

What is the Normal Curve? The normal curve is the beautiful bell shaped curve shown in Figure 1. It is a very useful curve in statistics because many attributes, when a large number of measurements are taken, are approximately distributed in this pattern. For example, the distribution of the wingspans of a large colony of butterflies, of the errors made in repeatedly measuring a 1 kilogram weight and of the amount of sleep you get per night are approximately normal. Many human characteristics, such as height, IQ or examination scores of a large number of people, follow the normal distribution.

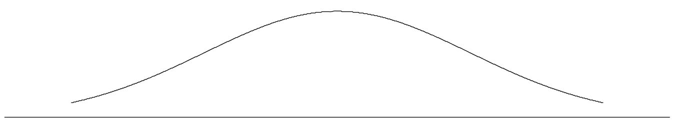


Figure 1: A normal curve.

You may be wondering what is "normal" about the normal distribution. The name arose from the historical derivation of this distribution as a model for the errors made in astronomical observations and other scientific observations. In this model the "average" represents the true or normal value of the measurement and deviations from this are errors. Small errors would occur more frequently than large errors.

The model probably originated in 1733 in the work of the mathematician Abraham Demoivre, who was interested in laws of chance governing gambling, and it was also independently derived in 1786 by Pierre Laplace, an astronomer and mathematician. However, the normal curve as a model for error distribution in scientific theory is most commonly associated with a German astronomer and mathematician, Karl Friedrich Gauss, who found a new derivation of the formula for the curve in 1809. For this reason, the normal curve is sometimes referred to as the "Gaussian" curve. In 1835 another mathematician and astronomer, Lambert Qutelet, used the model to describe human physiological and social traits. Qutelet believed that "normal" meant average and that deviations from the average were nature's mistakes.

When we draw a normal distribution for some variable, the values of the variable are represented on the horizontal axis called the $X$ axis. We will refer to these values as scores or observations. The area under the curve over any interval represents the proportion of scores in that interval. The height of the curve over an interval from $a$ to $b$, is the density or crowdedness of that interval; the higher the curve over an interval the more "crowded" that interval. This is illustrated in Figure 2.

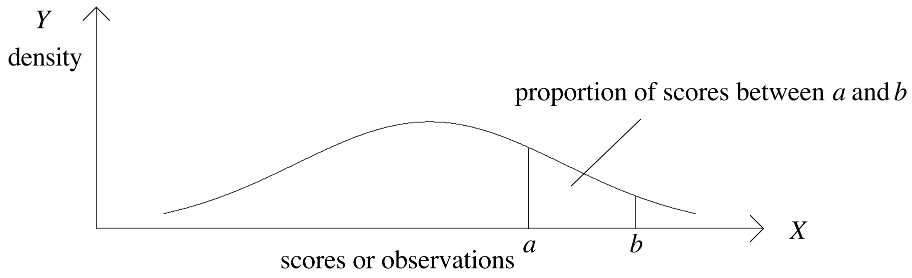



Figure 2: Representation of proportion of scores between two values of variable $X$.

Can you see where the normal distribution is most crowded or dense?

The scores or observations are most crowded (dense) in intervals around the mean, where the curve is highest. Towards the ends of the curve, the height is lower; the scores become less crowded the further from the mean we go. This tells us that observations around the mean are more likely to occur than observations further from the centre. In a random selection from the normal distribution, scores around the mean have a higher likelihood or probability of being selected than scores far away from the mean.

The normal distribution is not really the normal distribution but a family of distributions. Each of them has these properties:

**1.** the total area under the curve is 1;

**2.** the curve is symmetrical so that the mean, median and mode fall together;

**3.** the curve is bell shaped;

**4.** the greatest proportion of scores lies close to the mean. The further from the mean one goes (in either direction) the fewer the scores;

**5.** almost all the scores (0.997 of them) lie within 3 standard deviations of the mean.

The reason for these common properties is that all normal curves are based on the scary looking equation below. If we are measuring values ($x$) of a variable, such as height, then the distribution of these heights is given by $f(x)$ where

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

This equation does not need to concern us other than to note that it involves $\mu$, the mean of the population, and $\sigma$, the standard deviation of the population.

The value of the mean fixes the location of the normal curve, where it is centred. In all normal curves half the scores lie to the left of the mean and half to the right.

The value of the standard deviation determines the spread; the bigger $\sigma$, the more spread out or flat the curve.

If you would like to learn more about means and standard deviations, you can read the Mathematics Learning Centre booklet: Descriptive Statistics.

**Example**

In Figure 3 we have three normal curves.

In the first curve the mean is 0 and the standard deviation is 1.

The second curve has the same mean, 0, but a standard deviation of 2.

Can you see what the mean and standard deviation are for the third curve?

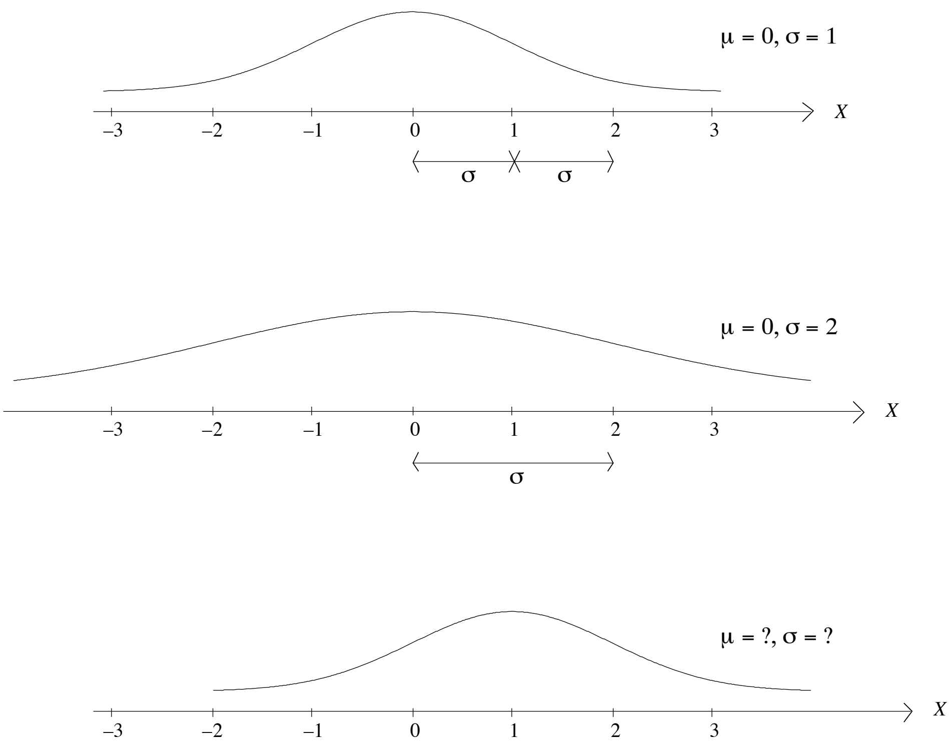



Figure 3: Normal curves with different means and standard deviations.

**Solution** $\mu = 1$ and $\sigma = 1$.

**Exercise**

A normal curve is given in Figure 4. Estimate the proportion of scores lying within one standard deviation of the mean. That is, estimate the proportion of scores between $\mu - \sigma$ and $\mu + \sigma$. Express your estimate as a decimal and as a percentage. This proportion is represented by the shaded area in Figure 4.

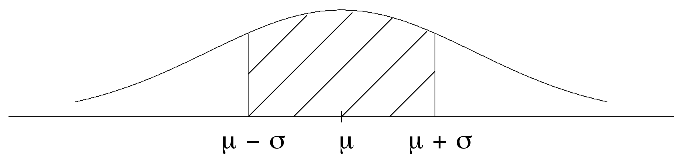



Figure 4: Normal curve showing proportion of scores within 1 standard deviation of mean.

**Solution**

The shaded area represents about 68 percent (0.68) of the scores. This proportion is the same for all normal curves. Check that this seems correct for the three curves in Figure 3.

**Notation**

We will adopt the convention of using capital $X$ when we are talking about the variable $X$, and little $x$ when we are talking about the values of the variable.

The notation for normal curves is as follows: if $X$ follows the normal distribution with mean $\mu_X$ and standard deviation $\sigma_X$ we write this as $X \sim N(\mu_X, \sigma_X^2)$. The symbol $\sigma_X^2$ is called the variance. It is equal to the square of the standard deviation.

The subscript $X$ in $\mu_X$ and $\sigma_X$ refers to the variable $X$. This is useful when we have more than one variable.

**Exercise**

Suppose $Y$ is a variable representing scores on a mathematics test. $Y$ is normally distributed with mean 75 and standard deviation 5.

Rewrite the following showing the values of $\mu_Y$ and $\sigma_Y^2$: $Y \sim N(\mu_Y, \sigma_Y^2)$.

**Solution** $Y \sim N(75, 25)$.

## 1.2 Shapes of distributions

Although many variables are approximately normal in distribution, many are not. For example, Figure 5 shows the hypothetical distribution of income for adults in Australia. As you can see this is not symmetrical in shape but has a "tail" of high earners. This is called skewed to the right.

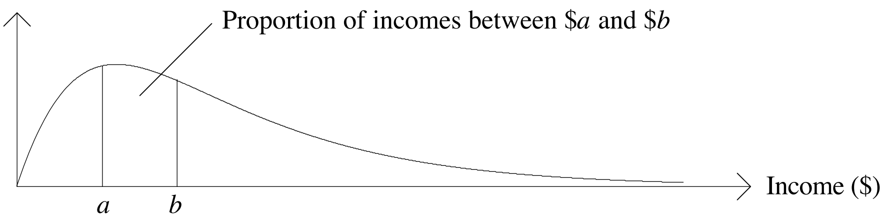



Figure 5: Example of a skewed distribution.

The outcomes of random events also do not necessarily follow the normal curve. For example, if you tossed a die over and over again, the long term pattern of outcomes would be **uniform**. That is, in theory, each number on the die from 1 to 6 would come up about one sixth of the time. The graph of the outcomes would look something like Figure 6.

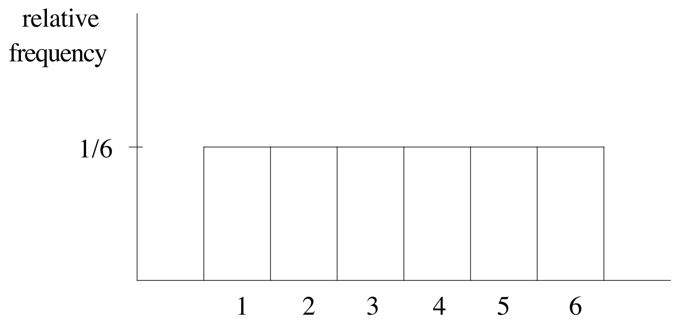



Figure 6: Uniform distribution.

Now here is an amazing fact which explains why the normal curve is so important in statistical investigations. If we take many, many random samples from some population of interest and calculate the sample mean in each case, then the distribution of these sample **means** will be approximately normal in shape provided the sample size is large.

Suppose, for example, we selected lots and lots of random samples of size 100,000 from the population of Australian adults and calculated the mean income for each sample. We would then have a big collection of different average incomes, one from each sample. The distribution of these average incomes (means) would be approximately normal, even though the distribution of individual incomes is not normal, as we have seen in Figure 5.

Similarly if you tossed a die 100 times, worked out the mean of the numbers that came up, and then repeated this experiment over and over again, the distribution of these **means** would be approximately normal.

This surprising result can be mathematically proved. It is a form of a profound and far reaching theorem called the Central Limit Theorem. It explains why many human characteristics follow the normal curve, as attributes such as height or weight can be thought of as a sort of "average". If we think of human weight or height as being a "sort of mean" of many factors (such as heredity, diet, race, sex, many others) then the Central Limit Theorem would lead us to expect that such human characteristics will follow the normal distribution.

In the next chapter we will work through a demonstration of the Central Limit Theorem. The proof of this theorem is beyond the scope of this booklet.

## 1.3 Summary

Normal curves all have the same basic bell shape but different centres and spreads.

Values of the variable are represented continuously along the horizontal axis, the $X$ axis.

Areas under the curve represent proportions of scores. We can indicate these proportions as decimals, fractions or percentages.

The whole area under the curve is 1 or 100 percent.

Because normal distributions are well understood and tabulated, we can work out proportions of observations within intervals for normally distributed variables.

### 1.3.1 Exercises

**1.** Where is the median (middle score) of the normal distribution? Give a reason for your answer.

**2.** Where is the mode (most common score) of the normal distribution? Give a reason for your answer.

**3.** Figure 7 shows two normal distribution curves representing the time taken to prepare personal ("S") and business ("A") income tax returns:

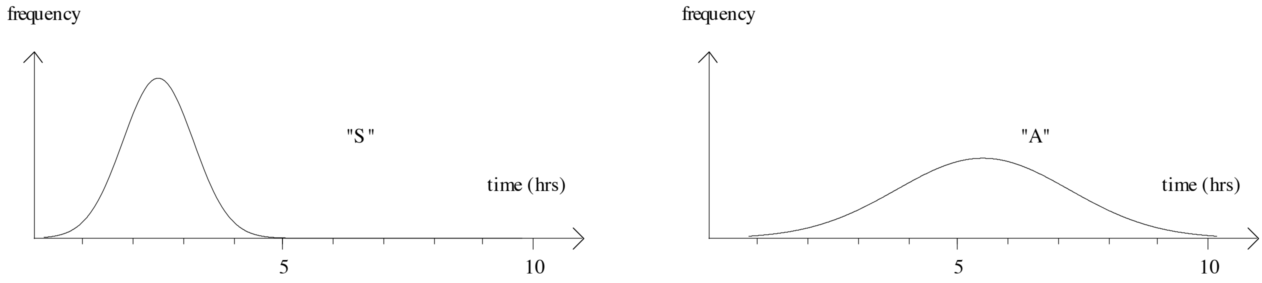



Figure 7: Two normal distribution curves.

(a) Which has the larger mean?

(b) Which has the larger standard deviation?

**4.** Select the correct alternative:

A normal distribution with a large standard deviation is
(more peaked / flatter) than one with a small standard deviation.

**5.** By running your finger along the curve in Figure 8, find the points where the concavity changes, that is where the curve changes from concave down to concave up. At these points the curve changes from steep to flatter. How many standard deviations away from the mean are these points?

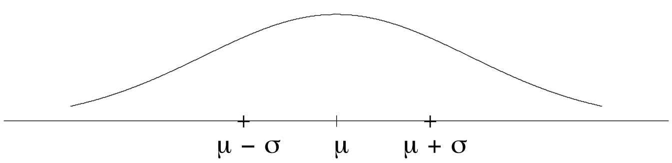
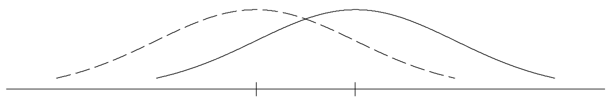



Figure 8: A normal distribution curve.

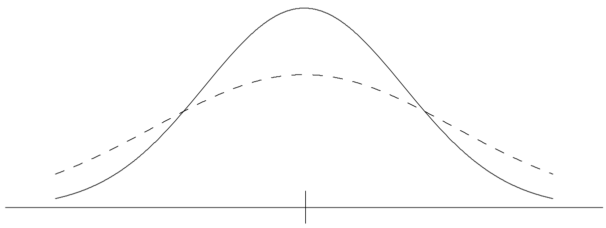
**6.** Three set of curves are given in Figure 9.

What could the solid curve and the dotted curve represent in each?

**a.**



**b.**



**c.**

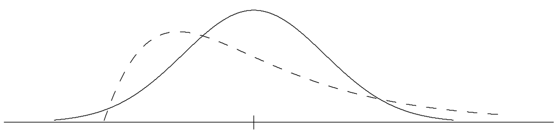


Figure 9: Three sets of probability distribution curves.

# 2 Why is the Normal Distribution Interesting?

A big part of statistical application concerns making inferences from a sample to a parent population. In this chapter we will explore why the normal distribution is useful in psychological research and other scientific applications.

Let us first revise some terminology.

A **population** is the whole group of interest. A population can be summarised by various parameters, or fixed numbers, such as the mean and variance.

For example, let $X$ = height of a student at Sydney University. The population consists of the heights of all students at the university. The mean height $\mu_X$, and the variance, $\sigma_X^2$, are two parameters or fixed values associated with this population. We could find $\mu_X$ and $\sigma_X^2$ by taking a census of the heights of students and calculating the mean and variance. The answers are constants, that is, numbers which do not fluctuate.

A **sample** is a selection from the parent population. Many statistical procedures make use of random samples. Samples can be of different sizes, where sample size is denoted by $n$. The mean of any one sample is likely to differ from the mean of a second sample from the same population. So the sample mean, $\overline{X}$, is a variable or statistic. It can take on many different values. For example, if we randomly select a sample of 25 students from the University we could calculate the sample mean of their heights. If we repeat the process over and over we are likely to get a range of different values of $\overline{X}$. So $\overline{X}$ is a variable and since it is a variable it has a distribution. This distribution is called the **sampling distribution of the mean**.

What do you think is the shape of the sampling distribution of the mean?

If you guessed the normal distribution you are sort of correct. Here is more of the story.

If $X$ is normally distributed $\overline{X}$ is normally distributed. If $X$ is not normally distributed, $\overline{X}$ is approximately normal if the sample size, $n$, is sufficiently big. This last, amazing and non-intuitive result explains why the normal distribution is useful to social scientists and others. It follows from a profound and far reaching theorem called the Central Limit Theorem.

## 2.1 Central Limit Theorem

Informally, the Central Limit Theorem expresses that if a random variable is the sum of $n$, independent, identically distributed, non-normal random variables, then its distribution approaches normal as $n$ approaches infinity.

As a consequence of the Central Limit Theorem we have the following corollary: The distribution of the sample mean $(\overline{X})$ approaches the normal distribution as the sample size $n$ increases, if the parent distribution from which the samples are drawn is not normal.

Let us look at a demonstration of this result. Suppose we have a box containing three tickets marked 1, 2, 3 as illustated in Figure 10.

| | | |
|---|---|---|
| 1 | 2 | 3 |

Figure 10: Box containing tickets marked 1, 2, 3.

If we draw out one ticket at random, record the number then replace the ticket and repeat this process over and over, there would be roughly an equal number of 1s 2s and 3s. Let $X$ = Number on the ticket drawn. This is our parent population. It has a **uniform distribution** which looks something like this.

| | | |
|---|---|---|
| ◇ | ◇ | ◇ |
| ◇ | ◇ | ◇ |
| 1 | 2 | 3 |
| 1 | 2 | 3 |
| 1 | 2 | 3 |
| 1 | 2 | 3 |

Figure 11: Distribution of $X$.

We can summarise some properties of this distribution as follows: Mean: $\mu_X = 2$, variance: $\sigma_X^2 = 0.\dot{6}$. Note $0.\dot{6}$ means $0.66666\cdots$.

Now suppose we draw out a ticket (at random), replace it and then draw out a second ticket. Our **sample size, $n$, is 2**. What are the possible samples we could draw? The following table lists all possible samples of size 2 and shows the value of the sample mean in each case.

| **Sample** | **Sample Means ($\overline{x}$)** |
|---|---|
| 1,1 | 1 |
| 1,2 | 1.5 |
| 1,3 | 2 |
| 2,1 | 1.5 |
| 2,2 | 2 |
| 2,3 | 2.5 |
| 3,1 | 2 |
| 3,2 | 2.5 |
| 3,3 | 3 |

Table 1: Samples of size 2 and corresponding sample means.

Notice that the variable $\overline{X}$ takes on the values 1 and 3 once each, the values 1.5 and 2.5 twice each and the value 2 three times.

| | | | | |
|---|---|---|---|---|
| | | 2 | | |
| | 1.5 | 2 | 2.5 | |
| 1 | 1.5 | 2 | 2.5 | 3 |

Figure 12: Sampling distribution of the mean for $n = 2$.

The distribution has parameters associated with it, such as mean, $\mu_{\overline{X}}$, and variance, $\sigma^2_{\overline{X}}$. Use your calculator to find the values of this mean and variance.

Solution: $\mu_{\overline{X}} = 2$, $\sigma^2_{\overline{X}} = 0.\dot{3}$, $n = 2$.

What do you notice about the shape of this distribution compared to that of the parent population? Compare the mean $\mu_{\overline{X}}$, above, with $\mu_X$, the mean of the parent distribution. What do you notice? Can you see a relationship between the variance $\sigma^2_{\overline{X}}$ above and the variance, $\sigma^2_X$ of the parent population?

Now suppose we select random samples of size 3, with replacement, and repeat the above process. This time, $n = 3$. The table below lists all 27 possible samples with $n = 3$ and their corresponding sample means.

| **Sample** | **Sample Means ($\overline{x}$)** | **Sample** | **Sample Means ($\overline{x}$)** |
|---|---|---|---|
| 1,1,1 | 1 | 2,2,3 | $2.\dot{3}$ |
| 1,1,2 | $1.\dot{3}$ | 2,3,1 | 2 |
| 1,1,3 | $1.\dot{6}$ | 2,3,2 | $2.\dot{3}$ |
| 1,2,1 | $1.\dot{3}$ | 2,3,3 | $2.\dot{6}$ |
| 1,2,2 | $1.\dot{6}$ | 3,1,1 | $1.\dot{6}$ |
| 1,2,3 | 2 | 3,1,2 | 2 |
| 1,3,1 | $1.\dot{6}$ | 3,1,3 | $2.\dot{3}$ |
| 1,3,2 | 2 | 3,2,1 | 2 |
| 1,3,3 | $2.\dot{3}$ | 3,2,2 | $2.\dot{3}$ |
| 2,1,1 | $1.\dot{3}$ | 3,2,3 | $2.\dot{6}$ |
| 2,1,2 | $1.\dot{6}$ | 3,3,1 | $2.\dot{3}$ |
| 2,1,3 | 2 | 3,3,2 | $2.\dot{6}$ |
| 2,2,1 | $1.\dot{6}$ | 3,3,3 | 3 |
| 2,2,2 | 2 | | |

Table 2: Samples of size 3 and corresponding sample means.

Now draw the distribution of $\overline{X}$ on Figure 13, below.

$1 \quad 1.\dot{3} \quad 1.\dot{6} \quad 2 \quad 2.\dot{3} \quad 2.\dot{6} \quad 3$

Figure 13: Sampling distribution of the mean for $n = 3$.

Compare your diagram with Figure 11 and Figure 12. Can you see what happened to the shape? See the end of the chapter for the solution.

We can calculate this mean and the variance of the 27 values of $\overline{X}$ above.

**Solution** $\mu_{\overline{X}} = 2$, variance $\sigma^2_{\overline{X}} = 0.\dot{2}$, $n = 3$.

So the mean is still the same as the mean of the parent population. The spread is decreasing as the sample size increases—the columns are closer together and the shape is becoming more peaked.

Now suppose we took samples of size 4. What do you think is the mean of this distribution ($\mu_{\overline{X}}$)? Can you guess the variance ($\sigma^2_{\overline{X}}$)? Look at the previous means and variances.

Recall that $\mu_X = 2$ and $\sigma^2_X = 0.\dot{6}$.

| $n$ | $\mu_{\overline{X}}$ | $\sigma^2_{\overline{X}}$ |
|---|---|---|
| 2 | 2 | $0.\dot{3}$ |
| 3 | 2 | $0.\dot{2}$ |
| 4 | | |

**Solution** For $n = 4$, $\mu_{\overline{X}} = 2$ and $\sigma^2_{\overline{X}} = \frac{0.\dot{6}}{4} = 0.1\dot{6}$.

What we have shown above is a demonstration, not a proof, of the Central Limit Theorem. The proof involves some fairly complex mathematics. There are some exceptions to the applications of the Central Limit Theorem but these are beyond the scope of this booklet.

**Can you answer these?**

i What happens to the shape of the distribution of the sampling mean as $n$ increases?

ii What is the relation between the mean, $\mu_{\overline{X}}$, of the distribution of sampling mean, and the mean, $\mu_X$, of the parent population?

iii What is the relation between the variance, $\sigma^2_{\overline{X}}$, of the distribution of sampling mean and the variance, $\sigma^2_X$, of the parent population?

## 2.2 Summary

If the parent population is normally distributed then the distribution of the sampling mean is exactly normal. Otherwise the distribution of the sampling mean ($\overline{X}$) will become close to the normal distribution for $n$ (sample size) large.

The mean of this distribution of $\overline{X}$ is equal to the mean of the parent distribution: $\mu_{\overline{X}} = \mu_X$.

The variance of the distribution of the sampling mean is equal to the variance of the parent population divided by the sample size, $n$. That is, the variance gets smaller by a factor of $n$: $\sigma^2_{\overline{X}} = \dfrac{\sigma^2_X}{n}$.

The central limit theorem explains why the normal distribution is linked to so many measured phenomena in our world—roughly speaking, data which are influenced by many small and unrelated random effects are approximately normally distributed.

### 2.2.1 Exercises

**1.** Let $X$ = number of children in a household in Sydney. Suppose we take random samples of size 2 from the above parent population, that is we randomly select 2 households at a time and count the number of children in each household. The diagrams in Figure 14 represent the outcomes from six samples. So, for example, in sample 1, one household had no children the other had 2 children.

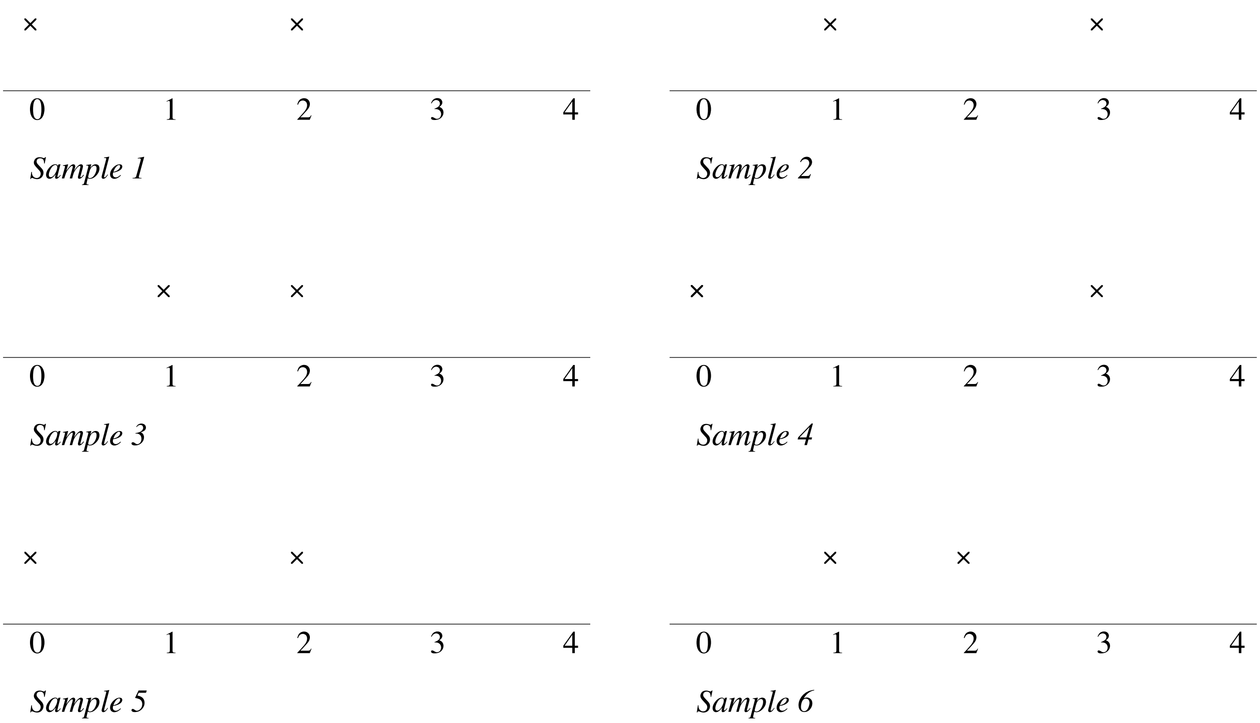



Figure 14: Samples showing number of children, $n = 2$.

**a.** Find the sample mean in each case and mark it on the diagram.

**b.** Draw the distribution of $\overline{X}$ for these six samples. Use the same scale on the axis as above.

**c.** Based on your data above what do you guess is the mean number of children in an Australian household? That is, estimate $\mu_X$ from the data. What could you do to improve your estimate?

**2.** If $X$ is distributed normally with $\mu_X = 10$ and $\sigma_X^2 = 25$, and we select samples of size 100, describe the distribution of $\overline{X}$ including its mean and variance.

**3.** Try this interactive demonstration of the Central Limit Theorem on the Web: Interactive Demonstrations for Statistics Education on the World Wide Web R. Webster West and R. Todd Ogden University of South Carolina Journal of Statistics Education v.6, n.3 (1998) http://www.amstat.org/publications/jse/v6n3/applets/CLT.html

In this demonstration we are simulating finding the distribution of $S$ where $S =$ total score showing on $n$ dice, for $n = 2, 3, 4, 5$. If $X =$ number showing on 1 die, see if you can estimate the mean $\mu_S$ in each case, in terms of $\mu_X$. Can you see a pattern?

**4.** How big must $n$ be for the distribution of the sample mean, $\overline{X}$, to be approximately normal?

### Solutions to diagram in text

The shape of the distribution of $\overline{X}$ for $n = 3$ is shown in Figure 15.

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | 2 | | | |
| | | $1.\dot{6}$ | 2 | $2.\dot{3}$ | | |
| | | $1.\dot{6}$ | 2 | $2.\dot{3}$ | | |
| | | $1.\dot{6}$ | 2 | $2.\dot{3}$ | | |
| | $1.\dot{3}$ | $1.\dot{6}$ | 2 | $2.\dot{3}$ | $2.\dot{6}$ | |
| | $1.\dot{3}$ | $1.\dot{6}$ | 2 | $2.\dot{3}$ | $2.\dot{6}$ | |
| 1 | $1.\dot{3}$ | $1.\dot{6}$ | 2 | $2.\dot{3}$ | $2.\dot{6}$ | 3 |

Figure 15: Sampling distribution of the mean for $n = 3$.

# 3 Areas Under the Standard Normal Curve

## 3.1 Finding areas under the standard normal curve

The standard normal distribution has a mean of 0 and a standard deviation and variance of 1. So if $Z$ is a standard normal variable, $\mu_Z = 0$, $\sigma_Z = 1$, $\sigma_Z^2 = 1$. The notation for this is $Z \sim N(0,1)$. Again, we distinguish between the variable, $Z$ (capital $Z$), and its values, called $z$ scores, for example $z = 1$, $z = 2$, written with a small $z$.

The following diagram, Figure 16, is a simplified representation of a standard normal distribution curve showing approximately the percentage of observations or scores in various regions.

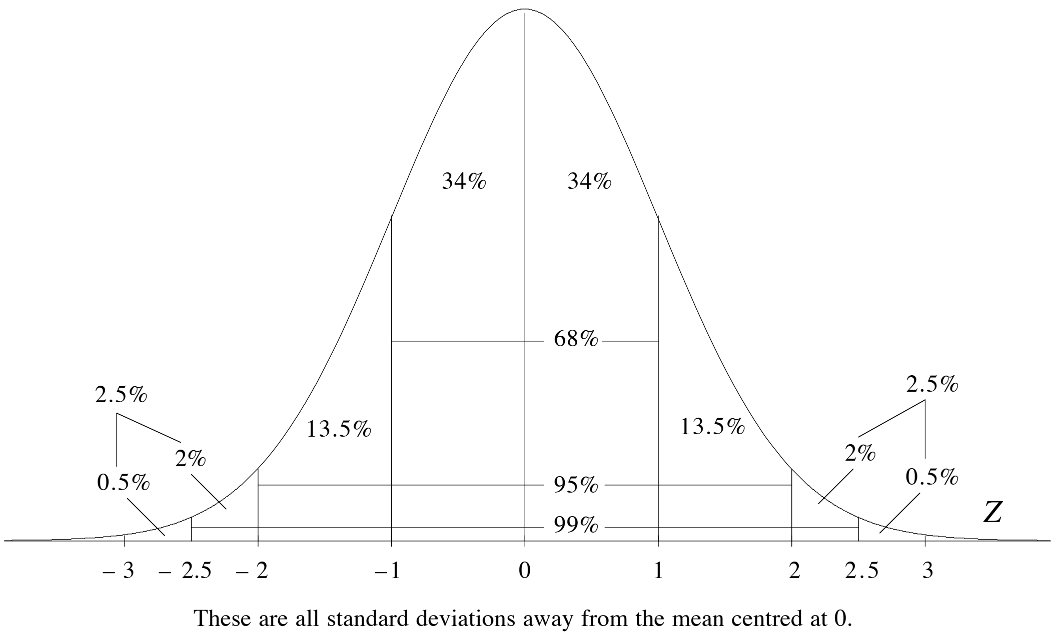



Figure 16: Standard normal curve showing approximate areas.

From Figure 16 we can see that:

i about 68% of the $z$ scores lie within 1 standard deviation of the mean, that is, between $-1$ and $+1$.

ii about 95% of the $z$ scores lie within 2 standard deviations of the mean, that is between $-2$ and $+2$.

iii almost all the $z$ scores lie between $-3$ and $+3$ standard deviations from the mean. (Our graph shows 100% of the observations lie between between $-3$ and $+3$ but more accurately this is 99.74%).

The $z$ **scores** are represented along the **horizontal axis**. The **area** under the curve corresponding to an interval of scores represents the percentage or **proportion** of scores in this interval.

The **probability** of selecting scores from a given interval is also represented by the **area** under the curve above that interval. For example, the probability of selecting a score greater than $z = 2$ is about 0.025 as the area above this interval is about 2.5%.

Notice the symmetry of the standard normal curve with respect to positive and negative $z$ scores and the corresponding areas.

### 3.1.1 Exercise

Study carefully the diagram of the normal curve given in Figure 16 and then complete the table using the percentages given.

| Interval | Percentage of area under the curve | Proportion of area under the curve expressed as a decimal | Probability of selecting a score in this interval |
|---|---|---|---|
| Between 0 and +1 | 34% | 0.34 | 0.34 |
| Between −3 and +3 | 100% | 1 | 1 |
| a) Less than 0 | | | |
| b) Greater than 0 | | | |
| c) Between 0 and +2 | | | |
| d) Between 0 and −2 | | | |
| e) Between −2 and +2 | | | |
| f) Outside (beyond) −2 and +2 | | | |
| g) Between +1 and $+2\frac{1}{2}$ | | | |
| h) Between −3 and −2 | | | |
| i) Greater than $2\frac{1}{2}$ | | | |
| j) Outside (beyond) $-2\frac{1}{2}$ and $+2\frac{1}{2}$ | | | |

The above exercise shows that if we randomly select a value of a normally distributed variable, then

i the probability of getting a value above the mean is 0.5. This is also the probability of getting a value below the mean

ii the approximate probability of getting a value beyond 2 standard deviations from the mean, that is, bigger than $z = 2$ or smaller than $z = -2$ is 0.05 ($2 \times 0.025$)

iii the approximate probability of getting a value beyond two and a half standard deviations from the mean is 0.01 ($2 \times 0.005$).

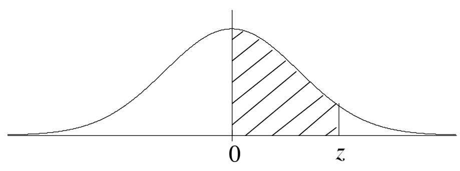
## 3.2 More about finding areas under the standard normal curve

Up to now we have only looked at areas under the normal curve corresponding to 1, 2 or 3 standard deviations above or below the mean. Now we will expand our understanding to a more comprehensive view of areas under the normal curve where the number of standard deviations from the mean may not be whole numbers, for example $z = 1.58$.

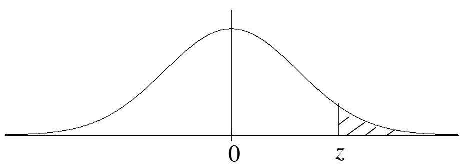
Turn to the end of this booklet to see the table giving areas under the standard normal curve for $z$ scores from 0 to 4.00. Remember that in a standard normal curve the mean is 0 and the standard deviation is 1. Since the normal curve is symmetric we can use the same table to find the areas below the mean corresponding to negative $z$ scores. The purpose of using this table is that we can find **probabilities** represented by these **areas**.

This is how the table works. The left hand column shows the $z$ **score**, that is, the number of standard deviations above the mean. These $z$ scores increase in jumps of 0.01. Notice that this column starts at $z = 0$ or $z = 0.00$, that is, the mean itself. The remaining three columns show **areas** under the normal curve. They are

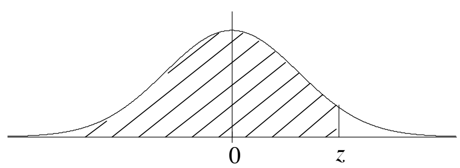
a.
the area between the mean and the $z$ score




b.
the area beyond the $z$ score, called the smaller portion




c.
the area up to the $z$ score, the larger portion.




Remember: The whole area under the curve is 1.

We will start with some examples of finding areas associated with positive and negative $z$ scores and the interpretations of these areas. It is useful to draw a diagram showing the $z$ score and required area.

Note: It is very important that you distinguish between $z$ scores which are represented as points on the horizontal axis and areas under the curve. These areas represent proportions or probabilities.

**Example**

**a.** If $z = 2.15$, what is the area beyond $z$? What does this tell us?

**b.** Find the area below (up to) $z$.

**c.** What is the sum of the above two areas? Why is this?

**d.** What is the area between the mean and 2.15 standard deviations?

**Solution**

**a.** We illustrate the $z$ score and the required area in Figure 17. The area beyond $z = 2.15$ is shaded.

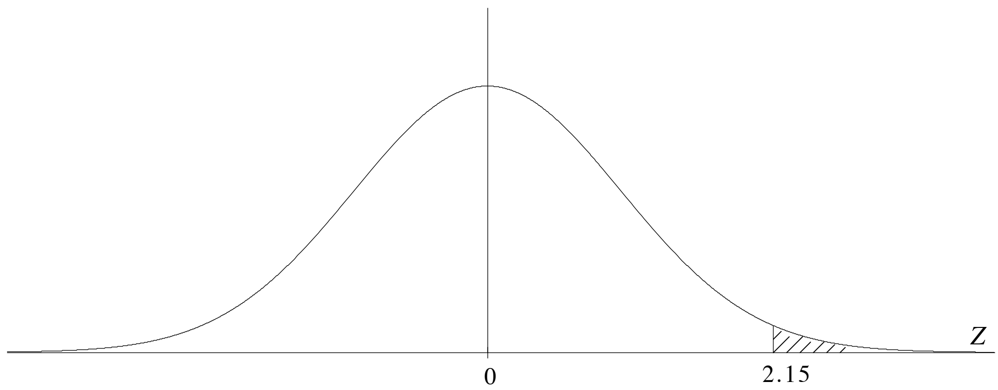



Figure 17: Shaded area represents proportion of scores beyond $z = 2.15$.

We now look down the left column of the table to find $z = 2.15$. Table 3 shows the areas between the mean and $z$, beyond $z$ (smaller portion) and up to $z$ (larger portion).

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 2.15 | 0.4842 | 0.0158 | 0.9842 |

Table 3: Areas corresponding to $z = 2.15$.

The area beyond $z = 2.15$ is 0.0158, the smaller portion. This means that the proportion of $z$ scores that exceed 2.15 is 0.0158 (ie less than 2% of the $z$ scores exceed 2.15). We can also interpret this as: the probability of selecting a $z$ score greater than 2.15 is 0.0158.

**b.** The area up to $z$ is 0.9842, the larger portion under the curve.

**c.** These two areas add up to 1, the total area under the normal curve.

**d.** The value of this area is shown in the table under the column Mean to $z$. It is 0.4842.

**Example**

What proportion of the $z$ scores are less than a $z$ score of 1.58?

**Solution**

The area representing this proportion is shaded in Figure 18.

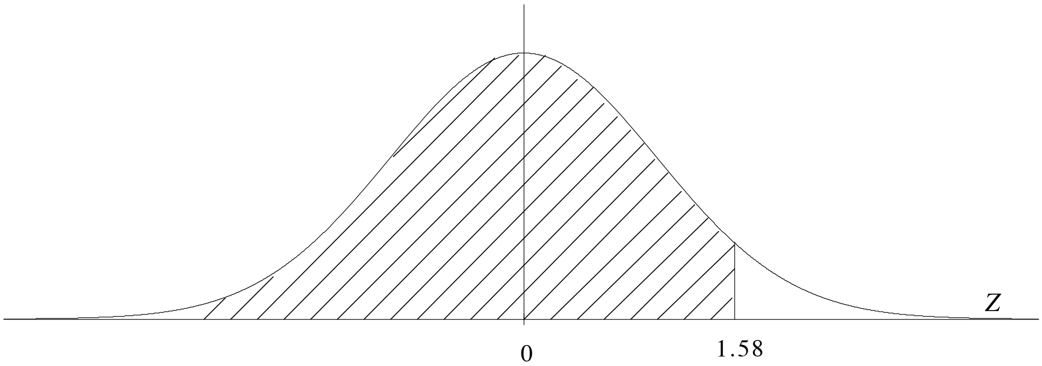



Figure 18: Shaded area represents proportion of scores up to $z = 1.58$.

The required part of the table is shown in Table 4.

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 1.58 | 0.4429 | 0.0571 | 0.9429 |

Table 4: Areas corresponding to $z = 1.58$.

The area below $z = 1.58$ is the larger portion: 0.9429. This means that 0.9429 of the $z$ scores are less than the $z$ score of 1.58. Alternatively we can say: 94.29% of the $z$ scores are less than $z = 1.58$.

**Example**

What is the area between the mean and 0.85 standard deviations below the mean (ie between $z$ scores of $-0.85$ and 0)?

**Solution**

The area is shaded in Figure 19.

Now, because the normal curve is symmetrical, the area we want is equal to the area under the curve between 0 and $+0.85$. We look up that area in our table.

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 0.85 | 0.3023 | 0.1977 | 0.8023 |

Table 5: Areas corresponding to $z = 0.85$.

The required area is 0.3023 or 30.23%.

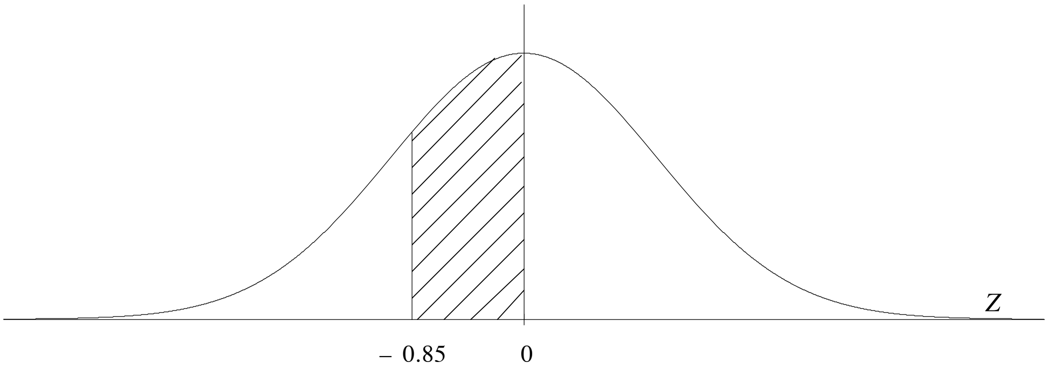



Figure 19: Shaded area represents proportion of scores between $z = -0.85$ and $z = 0$.

**Example**

What is the area between $z$ scores of 0.33 and 1.33?

**Solution**

The area is shaded in Figure 20.

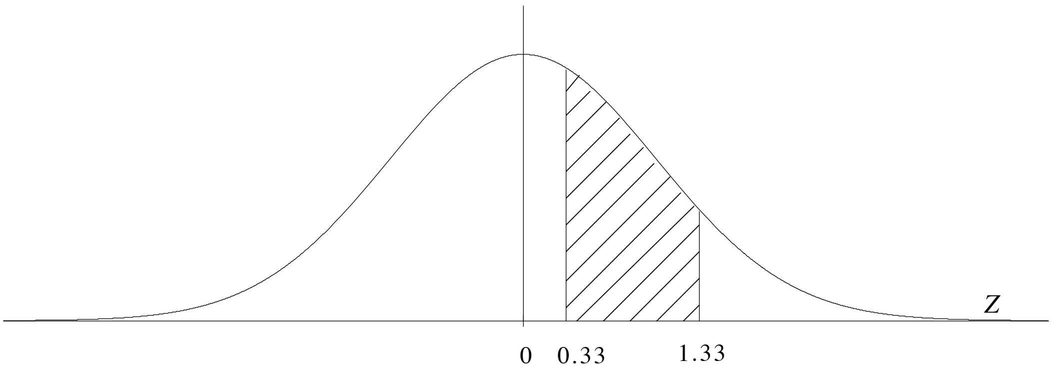



Figure 20: Shaded area represents proportion of scores between $z = 0.33$ and $z = 1.33$.

Looking up $z$=1.33 in the table gives an area of 0.9082 which is to the left of $z$=1.33 (larger portion). Similarly the area to the left of 0.33 can be seen as 0.6293. We find the required area by subtracting 0.6293 from 0.9082. So the shaded area is 0.2789 or 27.89%.

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 1.33 | 0.4082 | 0.0918 | 0.9082 |
| 0.33 | 0.1293 | 0.3707 | 0.6293 |

Table 6: Areas correponding to $z = 1.33$ and $z = 0.33$.

Can you find another way to get the same answer?

**Example**

What is the probability of obtaining a $z$ score between $-2.20$ and $0.25$ on the standard normal curve?

**Solution**

This probability is represented by the area under the curve between $z = -2.20$ and $z = 0.25$. This area is shaded in Figure 21.

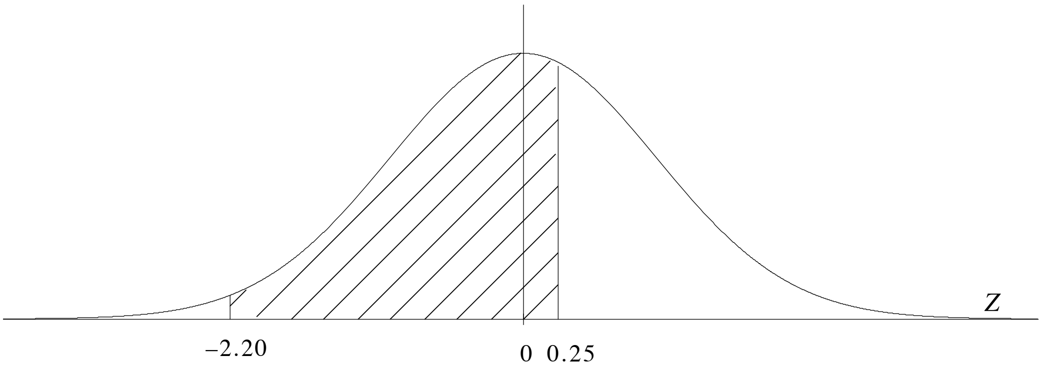



Figure 21: Shaded area represents proportion of scores between $z = -2.20$ and $z = 0.25$.

The area to the left of 0.25 can be found by looking up $z = 0.25$ in the table to get 0.5987 (larger portion). We need to subtract the area to the left of $z = -2.20$. Because of symmetry, this area is equal to the area to the right of $z = 2.2$ which is 0.0139 (smaller portion). The required area is $0.5987 - 0.0139 = 0.5848$. This means that the probability of obtaining a $z$ score in the stated interval is 0.5848.

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 0.25 | 0.0987 | 0.4013 | 0.5987 |
| 2.20 | 0.4861 | 0.0139 | 0.9861 |

Table 7: Areas corresponding to $z = -2.20$ and $z = 0.25$.

Can you find another way to get this answer?

**Example**

What $z$ score is exceeded by 10% of all scores under the normal curve?

**Solution**

This question requires us to work backwards. The required $z$ score is shown on the horizontal axis of Figure 22.

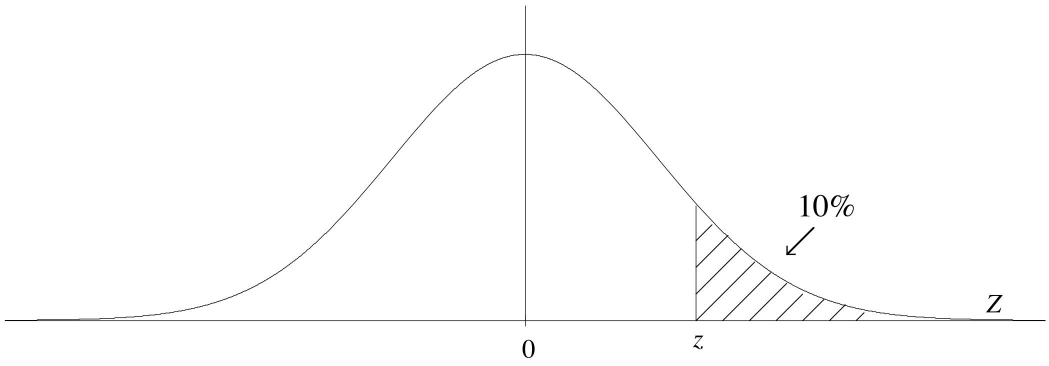



Figure 22: Shaded area represents 0.1 of scores.

To find $z$, we look in the "body" of the table under "smaller portion" column for 0.1. The closest we can get is 0.1003 which is the "smaller portion" corresponding to $z = 1.28$.

| $z$ | Mean to $z$ | smaller portion | larger portion |
|---|---|---|---|
| 1.28 | 0.3997 | 0.1003 | 0.8997 |

Table 8: Table showing smaller area approximately 0.1.

So, the required $z$ score is 1.28. This $z$ score is called the 90th percentile. That is, it is as high or higher than 90% of the $z$ scores.

You will find that your understanding of normal distributions is enhanced by being familiar with a few $z$ scores such as plus/minus 1, 2, 3 and their associated areas.

## 3.3 Summary

A standard normal distribution has a mean of 0 and a variance and standard deviation of 1.

Standardised scores are also called $z$ scores. The $z$ scores are most dense (most likely) around the mean of 0 and scores more extreme than $-3$ or $+3$ will be relatively rare.

The standard normal distribution or $Z$ distribution has been extensively tabulated and can be computer generated.

In these tables:

i $z$ scores are represented by points on the horizontal axis.

ii Areas under the curve represent the proportion of scores within an interval, or the percentage of scores within an interval, or the probability of selecting scores within an interval.

### 3.3.1 Exercises

Study the examples carefully and then try these exercises. The working is always easier to follow if you use a diagram.

**1.** Find the **areas** corresponding to the following intervals, expressing your answers as decimals and then percentages. Show each result on a diagram of the normal curve.

Area for $z$ scores:

**a.** below a $z$ score of $+0.85$;

**b.** above a $z$ score of $+2.75$;

**c.** below a $z$ score of $-1.03$;

**d.** between $z$ scores of $+1.58$ and $+2.35$;

**e.** between $z$ scores of $-2.80$ and $-2.50$;

**f.** between $z$ scores of $-1.55$ and $+1.55$;

**g.** between the mean and $z = +2.33$;

**h.** between the mean and 1.47 standard deviations above the mean;

**i.** between $z = -0.58$ and $z = 0$;

**j.** between the mean and 2.55 standard deviations below the mean.

**2.** Find the $z$ score in each case and show your answer on a sketch of the normal curve.

**a.** 50% of the $z$ scores exceed a $z$ score of ...?

**b.** 5% of the $z$ scores exceed a $z$ score of ...?

**c.** 99% of the $z$ scores exceed a $z$ score of ...?

**3.** What is the probability of selecting a $z$ score:

**a.** greater than $+1.96$?

**b.** smaller than $-1.96$?

**c.** greater that $+1.96$ or smaller than $-1.96$?

Show your answers on a sketch of the normal curve.

# 4 Transforming to Standard Scores

## 4.1 Transforming raw scores to $z$ scores

In the last chapter we saw that a standard normal curve is well understood and tabulated so that we can find areas associated with intervals of standard scores or $z$ scores.

Furthermore any normally distributed variable, $X$, can be transformed to a standard normal variable. To do this we shift the mean of the distribution to 0 and shrink or expand the standard deviation to 1. Suppose our population is normally distributed with mean $\mu_X$ and standard deviation $\sigma_X$. To transform raw scores to $z$ scores we must find out how many standard deviations the raw score is from the mean. To see how this is done consider this example.

Let $X$ = score on a nationwide English test. $X$ is normally distributed with $\mu_X = 80$ and $\sigma_X = 10$. For each student's raw score, termed $x$, we define the corresponding $z$ score as: $z$ = number of standard deviations from the mean.

Suppose Mike achieved 90 on the test. This is 10 marks above the mean and since the standard deviation is 10, Mike achieved a mark 1 standard deviation above the mean. So the $z$ score for Mike is 1. In short, for $x = 90$, $z = 1$.

**Exercise**

See if you can find the $z$ scores for the following students' marks on the test:

Mary achieved 70 on the test. $x = 70$ $z =?$

Jane achieved 100 on the test $x = 100$ $z =?$

Bob gained 80 on the test $x = 80$ $z =?$

**Solution**

| | | |
|---|---|---|
| Mary | $x = 70$ | $z = -1$ |
| Jane | $x = 100$ | $z = 2$ |
| Bob | $x = 80$ | $z = 0$ |

We can represent these transformations from raw scores to $z$ scores on a diagram like Figure 23.

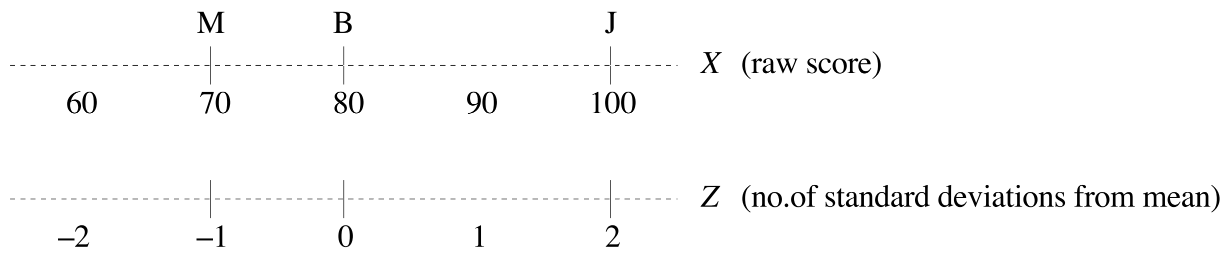



Figure 23: Raw scores and their equivalent $z$ scores.

So to transform from raw scores to $z$ scores, there are two steps:

Start with the raw score

1) subtract the mean

2) divide by the standard deviation.

In mathematics these two steps may be written as the following formula:

$$z = \frac{x - \mu_X}{\sigma_X}.$$

Using this formula allows us to convert any raw score to a $z$ score. For example, suppose Sam's mark on the test was 73. How did this compare with his classmates?

$$\begin{aligned} z &= \frac{x - \mu_X}{\sigma_X} \\ &= \frac{73 - 80}{10} \\ &= \frac{-7}{10} \\ &= -0.7 \end{aligned}$$

Sam's mark was 0.7 standard deviations below average.

Suppose Mei achieved 92 on the test. How many standard deviations was her mark above the mean?

$$\begin{aligned} z &= \frac{x - \mu_X}{\sigma_X} \\ &= \frac{92 - 80}{10} \\ &= \frac{12}{10} \\ &= 1.2 \end{aligned}$$

Mei achieved a grade 1.2 standard deviations above the mean.

We can now use our knowledge of the standard normal curve to find percentages, proportions or probabilities associated with intervals of scores for any normally distributed variable. First we must transform raw scores to $z$ scores, then we can use normal tables.

**Example**

Find the proportion of students who achieved a higher mark than Mei.

**Solution**

We represent the raw score, the $z$ score and the required area in Figure 24.

From our tables we see that the shaded area is 0.1151. Therefore about 12% of students achieved a higher mark than Mei.

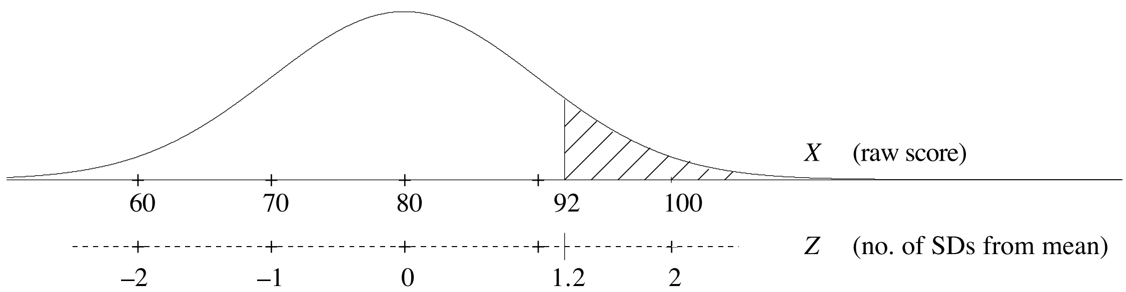



Figure 24: Shaded area represents proportion of students with a mark higher than Mei.

**Example**
Find the $z$ score corresponding to the mean in the English test.

**Solution**
In the above example $\mu_X = 80$. To find the corresponding $z$ score:

$$\begin{aligned} z &= \frac{x - \mu_X}{\sigma_X} \\ &= \frac{80 - 80}{10} \\ &= \frac{0}{10} \\ &= 0. \end{aligned}$$

Can you see why the $z$ score corresponding to the mean, $\mu_X$, will always be 0?

## 4.2 Transforming $z$ scores to raw scores

In this section we will reverse the process to get raw scores from z scores. Consider the English test above with $\mu_X = 80$ and $\sigma_X = 10$.
Suppose David achieved a grade 1.8 standard deviations above the mean ($z = 1.8$). What was his actual grade?

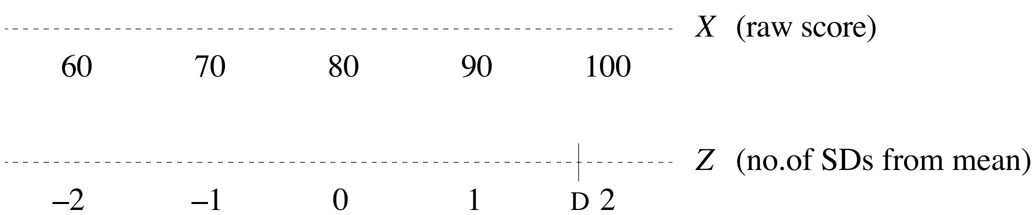



Figure 25: Marks on English test and the corresponding $z$ scores.

David's mark can be estimated from Figure 25 as close to, but below, 100. Since one standard deviation is 10 marks, 1.8 standard deviations above the mean is 18 marks above the mean. The mean is 80 so David's mark is $80 + 18 = 98$.

Using our two steps in section 4.1 in reverse:

Start with the $z$ score:

1) Multiply by the standard deviation

2) Add the mean.

The formula for this is

$$x = z\sigma_X + \mu_X.$$

It is more usual to write this as:

$$x = \mu_X + z\sigma_X.$$

**Example**

Use the above formula to convert the following $z$ scores to raw scores in the English test. Show all the results on a diagram.

**a.** $z = -2$

**b.** $z = 0.56$

**c.** $z = -1.4$

**d.** If Bob's mark was 0 standard deviations from the mean what was that mark?

**Solution**

**a.**
$$\begin{aligned} x &= \mu_X + z\sigma_X \\ &= 80 + (-2)(10) \\ &= 80 - 20 \\ &= 60 \end{aligned}$$

**b.**
$$\begin{aligned} x &= \mu_X + z\sigma_X \\ &= 80 + (0.56)(10) \\ &= 80 + 5.6 \\ &= 85.6 \end{aligned}$$

**c.**
$$\begin{aligned} x &= \mu_X + z\sigma_X \\ &= 80 + (-1.4)(10) \\ &= 80 - 14 \\ &= 66 \end{aligned}$$

**d.**
$$\begin{aligned} x &= \mu_X + z\sigma_X \\ &= 80 + (0)(10) \\ &= 80 + 0 \\ &= 80 \end{aligned}$$

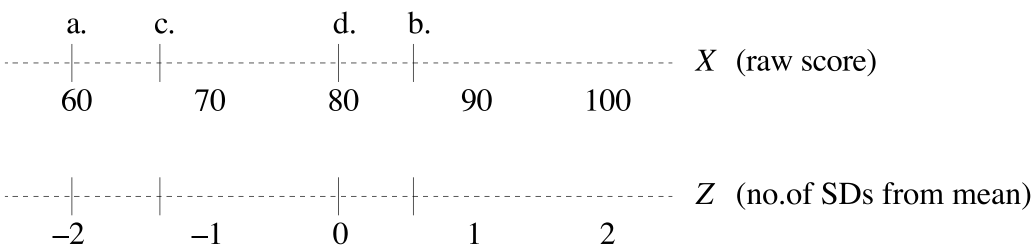



Figure 26: $z$ scores and the corresponding raw scores.

**Example**

Rob achieved a mark on the English test that exceeded 95% of all marks. Find Rob's English mark.

**Solution**

To first find Rob's English mark we need to find the $z$ score that exceeds 95% of all $z$ scores. This is marked in Figure 27.

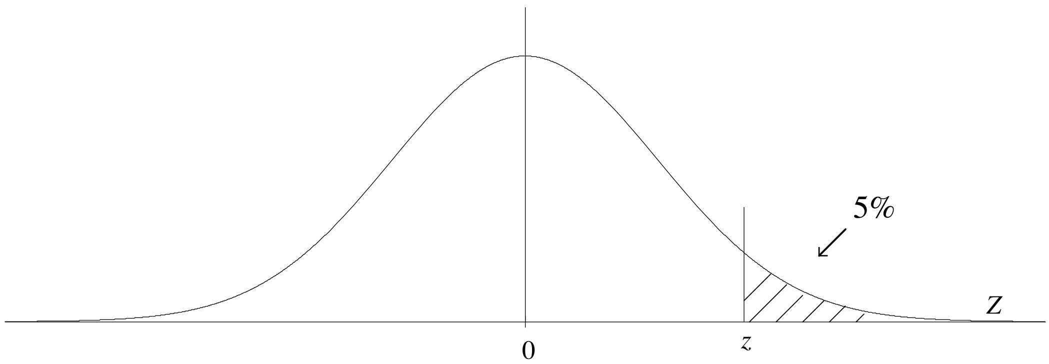



Figure 27: Normal curve with 95% of scores less than $z$.

From our tables, we find $z = 1.64$ (or $z = 1.65$).

Now we will convert this to a raw score.

$$\begin{aligned} x &= \mu_X + z\sigma_X \\ &= 80 + (1.64)(10) \\ &= 96.4 \end{aligned}$$

Therefore, Rob achieved a mark of 96% on the English test.

## 4.3 Summary

Any normally distributed variable, $X$, with mean $\mu_X$ and standard deviation $\sigma_X$ can be transformed to a standard normal variable, $Z$.

If $x$ is a raw score from this distribution, the formula $\dfrac{x - \mu_X}{\sigma_X}$ gives the corresponding $z$ score.

We can reverse the process to get a raw score, $x$, from a $z$ score using the formula $x = \mu_X + z\sigma_X$.

### 4.3.1 Exercises

**1.** Let $X$ be scores on a computer skills test with $\mu_X = 100$ and $\sigma_X = 10$. Assume the scores follow a normal distribution.

**a.** Find the number of standard deviations above or below the mean of each of the following scores on the computer test: 95, 110, 130.

**b.** Use a diagram to find the raw scores equivalent to the following $z$ scores: 0, $-1$, $-2$, 1, 2.

**c.** What is the $z$ score for a raw score of 118.4?

**2.** Assume the scores on the computer skills test follow the normal distribution in Question 1.

**a.** What proportion of the scores were greater than 118.4?

**b.** If a score is selected at random what is the probability that it is more than 1.96 standard deviations from the mean in either direction? This is $P(z < -1.96) + P(z > 1.96)$.

**c.** Find the 90th percentile for these scores, that is the score that exceeds 90% of the scores.

Hint: first use the tables at the back to find the $z$ score shown in Figure 28, then convert to a raw score.

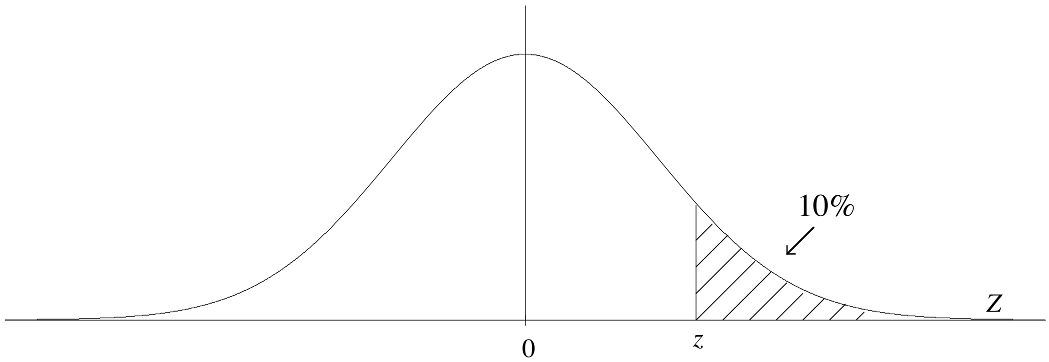



Figure 28: Normal curve with 90% of scores less than $z$.

**3.** Suppose scores on a mathematics test have a mean 60 and standard deviation 20 while scores on an English test have a mean 60 and standard deviation 10.

**a.** If Bob gets 80 on both tests, on which test did he do better relative to his class mates?

**b.** If the scores on the tests each follow a normal distribution how many students did better than Bob in each case?

# 5 Solutions to Exercises

**Solutions to exercises 1.3.1**

**1.** The median is the middle value of a distribution with 50% of the distribution less than the median and 50% greater than the median. As the normal distribution is symmetric, the median is equal to the mean, ie the centre of the distribution.

**2.** The mode of the normal distribution is equal to the mean. The highest point of the curve is above the mean.

**3.** **a.** The mean of distribution "S" is about 2.5 hours, while the mean of distribution "A" is about 5.5 hours. Therefore distribution "A" has the larger mean.

**b.** The normal distribution "A" is flatter or more spread out so has the larger standard deviation.

**4.** A normal distribution with a large standard deviation is flatter than one with a small standard deviation.

**5.** The normal distribution curve changes concavity one standard deviation above and below the mean. That is, in Figure 8 as you move along the curve from left to right, the concavity changes from shallower to steeper at $\mu - \sigma$ and from steeper to shallower at $\mu + \sigma$.

**6.** **a.** The dotted curve could represent the heights of all adult women, while the solid curve could represent the heights of all adult men.

**b.** The dotted curve could represent the distribution of heights of children aged 5-9, while the solid curve could represent the heights of children aged 6-8. The distributions have the same mean but the heights of the 5-9 years olds are more spread out.

**c.** The dotted curve could represent distribution of house prices in Sydney, while the solid curve could represent the distribution of house prices in a particular suburb.

**Solutions to exercises 2.2.1**

**1.** **a.** The sample means are 1, 2, 1.5, 1.5, 1, 1.5

**b.**

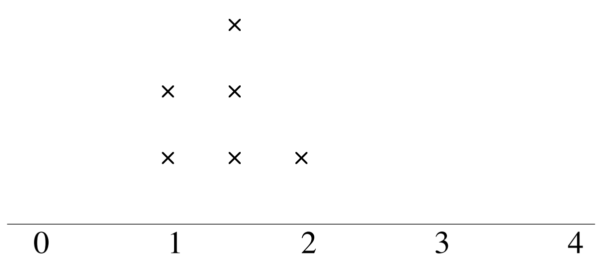



Figure 29: Distribution of $\overline{X}$.

**c.** We can see from a. that the values of $\overline{X}$ jump around from sample to sample. Our best estimate of $\mu_X$ is $\mu_{\overline{X}}$. We estimate $\mu_{\overline{X}}$ as 1.42.

To improve the estimate, increase the number of samples and the sample size.

**2.** $\overline{X}$ is distributed normally with mean $\mu_{\overline{X}} = 10$ and variance $\sigma^2_{\overline{X}} = 0.25$.

**3.** $\mu_S = n\mu_X$

**4.** There is no easy answer to this question. As stated, if the distribution of the parent population $X$ is normal, then the distribution of $\overline{X}$ is exactly normal. If the parent population $X$ is not normally distributed, then how big $n$ needs to be for the distribution of $\overline{X}$ to be approximately normal depends on the shape of the parent distribution. If the shape of the parent distribution is close to normal then $n$ could be quite small. In our demonstration example, we saw that for a uniform distribution, the distribution of $\overline{X}$ started moving towards an approximately normal shape quite quickly, even by $n = 3$. If, on the other hand, the parent distribution is very skewed, then $n$ would need to be quite large—how large is a difficult question to answer.

**Solution to exercise 3.1.1**

| Interval | Percentage of area under the curve | Proportion of area under the curve expressed as a decimal | Probability of selecting a score in this interval |
|---|---|---|---|
| Between 0 and +1 | 34% | 0.34 | 0.34 |
| Between −3 and +3 | 100% | 1 | 1 |
| a) Less than 0 | 50% | 0.5 | 0.5 |
| b) Greater than 0 | 50% | 0.5 | 0.5 |
| c) Between 0 and +2 | 47.5% | 0.475 | 0.475 |
| d) Between 0 and −2 | 47.5% | 0.475 | 0.475 |
| e) Between −2 and +2 | 95% | 0.95 | 0.95 |
| f) Outside −2 and +2 | 5% | 0.05 | 0.05 |
| g) Between +1 and $+2\frac{1}{2}$ | 15.5% | 0.155 | 0.155 |
| h) Between −3 and −2 | 2.5% | 0.025 | 0.025 |
| i) Greater than $2\frac{1}{2}$ | 0.5% | 0.005 | 0.005 |
| j) Outside $-2\frac{1}{2}$ and $+2\frac{1}{2}$ | 1% | 0.01 | 0.01 |

**Solutions to exercises 3.3.1**

**1.** **a.** Area = 0.8023 or 80.23%

**b.** Area = 0.0030 or 0.3%

**c.** Area = 0.1515 or 15.15%

**d.** Area = 0.0477 or 4.77%

**e.** Area = 0.0036 or 0.36%

**f.** Area = 0.8788 or 87.88%

**g.** Area = 0.4901 or 49.01%

**h.** Area = 0.4292 or 42.92%

**i.** Area = 0.2190 or 21.9%

**j.** Area = 0.4946 or 49.46%.

**2.** **a.** $z = 0$

**b.** $z = 1.645$ (value is between 1.64 and 1.65)

**c.** $z = -2.33$.

**3.** **a.** Probablity = 0.025

**b.** Probablity = 0.025

**c.** Probablity = 0.05 (adding the above two probabilities).

**Solution to exercises 4.3.1**

**1.** **a.** $x = 95 = 100 - 5$ and $\sigma_X = 10$, and so $x = 95$ is 0.5 standard deviations below the mean ($z = -0.5$). $x = 110$ is one standard deviation above the mean ($z = 1$). $x = 130$ is 3 standard deviations above the mean ($z = 3$).

**b.**

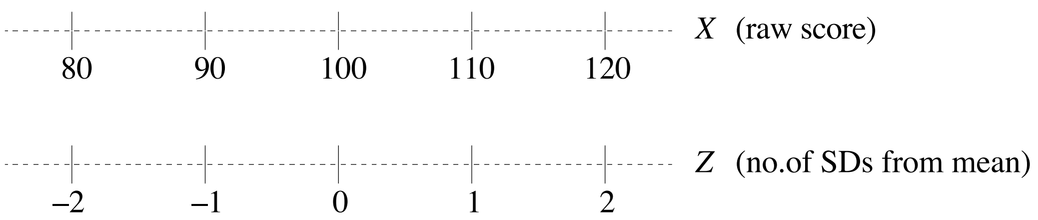



Figure 30: $z$ scores and the corresponding raw scores.

**c.**

$$\begin{aligned} z &= \frac{x - \mu_X}{\sigma_X} \\ &= \frac{118.4 - 100}{10} \\ &= 1.84. \end{aligned}$$

**2.** **a.** The proportion of scores greater than 118.4 is equal to the proportion of $z$ scores greater than $z = 1.84$. From the tables, this is the smaller portion and is equal to 0.0329.

**b.** From the tables, $P(Z > 1.96) = 0.0250$. Since the normal distribution is symmetric, the required area is $2 \times 0.025 = 0.05$.

**c.** Using the tables, look up 0.1 in the smaller portion. This gives us $z = 1.28$. We find the raw score as follows:

$$x = \mu_X + z\sigma_X = 100 + 1.28(10) = 112.8.$$

So, the 90th percentile for these scores is 112.8.

**3.** **a.** On the English test, Bob's raw score of 80 corresponds to

$$z = \frac{x - \mu_X}{\sigma_X} = \frac{80 - 60}{10} = 2.$$

On the mathematics test, Bob's raw score of 80 corresponds to

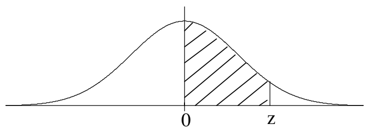
$$z = \frac{x - \mu_X}{\sigma_X} = \frac{80 - 60}{20} = 1.$$

So, relative to his class mates, Bob did better on the English test.

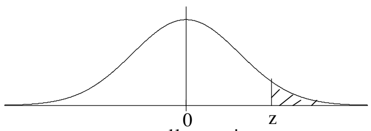
**b.** For the English test, from the tables $P(Z > 2) = 0.0228$. So about 2% of students did better than Bob on the English test.

For the mathematics test, from the tables, $P(Z > 1) = 0.1587$. So about 16% of students did better than Bob on the mathematics test.

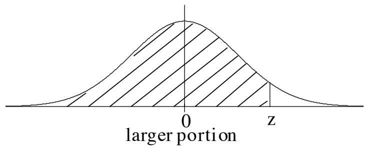
# 6 The Standard Normal Distribution Tables



mean to z

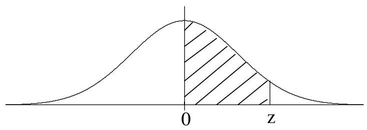


smaller portion

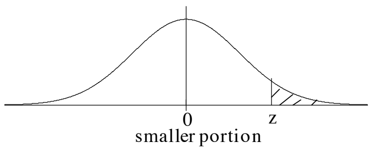


larger portion

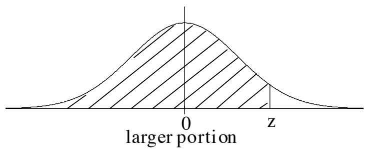
| $z$ score | mean to $z$ | smaller portion | larger portion | $z$ score | mean to $z$ | smaller portion | larger portion |
|---|---|---|---|---|---|---|---|
| 0.00 | 0.0000 | 0.5000 | 0.5000 | 0.40 | 0.1554 | 0.3446 | 0.6554 |
| 0.01 | 0.0040 | 0.4960 | 0.5040 | 0.41 | 0.1591 | 0.3409 | 0.6591 |
| 0.02 | 0.0080 | 0.4920 | 0.5080 | 0.42 | 0.1628 | 0.3372 | 0.6628 |
| 0.03 | 0.0120 | 0.4880 | 0.5120 | 0.43 | 0.1664 | 0.3336 | 0.6664 |
| 0.04 | 0.0160 | 0.4840 | 0.5160 | 0.44 | 0.1700 | 0.3300 | 0.6700 |
| 0.05 | 0.0199 | 0.4801 | 0.5199 | 0.45 | 0.1736 | 0.3264 | 0.6736 |
| 0.06 | 0.0239 | 0.4761 | 0.5239 | 0.46 | 0.1772 | 0.3228 | 0.6772 |
| 0.07 | 0.0279 | 0.4721 | 0.5279 | 0.47 | 0.1808 | 0.3192 | 0.6808 |
| 0.08 | 0.0319 | 0.4681 | 0.5319 | 0.48 | 0.1844 | 0.3156 | 0.6844 |
| 0.09 | 0.0359 | 0.4641 | 0.5359 | 0.49 | 0.1879 | 0.3121 | 0.6879 |
| 0.10 | 0.0398 | 0.4602 | 0.5398 | 0.50 | 0.1915 | 0.3085 | 0.6915 |
| 0.11 | 0.0438 | 0.4562 | 0.5438 | 0.51 | 0.1950 | 0.3050 | 0.6950 |
| 0.12 | 0.0478 | 0.4522 | 0.5478 | 0.52 | 0.1985 | 0.3015 | 0.6985 |
| 0.13 | 0.0517 | 0.4483 | 0.5517 | 0.53 | 0.2019 | 0.2981 | 0.7019 |
| 0.14 | 0.0557 | 0.4443 | 0.5557 | 0.54 | 0.2054 | 0.2946 | 0.7054 |
| 0.15 | 0.0596 | 0.4404 | 0.5596 | 0.55 | 0.2088 | 0.2912 | 0.7088 |
| 0.16 | 0.0636 | 0.4364 | 0.5636 | 0.56 | 0.2123 | 0.2877 | 0.7123 |
| 0.17 | 0.0675 | 0.4325 | 0.5675 | 0.57 | 0.2157 | 0.2843 | 0.7157 |
| 0.18 | 0.0714 | 0.4286 | 0.5714 | 0.58 | 0.2190 | 0.2810 | 0.7190 |
| 0.19 | 0.0753 | 0.4247 | 0.5753 | 0.59 | 0.2224 | 0.2776 | 0.7224 |
| 0.20 | 0.0793 | 0.4207 | 0.5793 | 0.60 | 0.2257 | 0.2743 | 0.7257 |
| 0.21 | 0.0832 | 0.4168 | 0.5832 | 0.61 | 0.2291 | 0.2709 | 0.7291 |
| 0.22 | 0.0871 | 0.4129 | 0.5871 | 0.62 | 0.2324 | 0.2676 | 0.7324 |
| 0.23 | 0.0910 | 0.4090 | 0.5910 | 0.63 | 0.2357 | 0.2643 | 0.7357 |
| 0.24 | 0.0948 | 0.4052 | 0.5948 | 0.64 | 0.2389 | 0.2611 | 0.7389 |
| 0.25 | 0.0987 | 0.4013 | 0.5987 | 0.65 | 0.2422 | 0.2578 | 0.7422 |
| 0.26 | 0.1026 | 0.3974 | 0.6026 | 0.66 | 0.2454 | 0.2546 | 0.7454 |
| 0.27 | 0.1064 | 0.3936 | 0.6064 | 0.67 | 0.2486 | 0.2514 | 0.7486 |
| 0.28 | 0.1103 | 0.3897 | 0.6103 | 0.68 | 0.2517 | 0.2483 | 0.7517 |
| 0.29 | 0.1141 | 0.3859 | 0.6141 | 0.69 | 0.2549 | 0.2451 | 0.7549 |
| 0.30 | 0.1179 | 0.3821 | 0.6179 | 0.70 | 0.2580 | 0.2420 | 0.7580 |
| 0.31 | 0.1217 | 0.3783 | 0.6217 | 0.71 | 0.2611 | 0.2389 | 0.7611 |
| 0.32 | 0.1255 | 0.3745 | 0.6255 | 0.72 | 0.2642 | 0.2358 | 0.7642 |
| 0.33 | 0.1293 | 0.3707 | 0.6293 | 0.73 | 0.2673 | 0.2327 | 0.7673 |
| 0.34 | 0.1331 | 0.3669 | 0.6331 | 0.74 | 0.2704 | 0.2296 | 0.7704 |
| 0.35 | 0.1368 | 0.3632 | 0.6368 | 0.75 | 0.2734 | 0.2266 | 0.7734 |
| 0.36 | 0.1406 | 0.3594 | 0.6406 | 0.76 | 0.2764 | 0.2236 | 0.7764 |
| 0.37 | 0.1443 | 0.3557 | 0.6443 | 0.77 | 0.2794 | 0.2206 | 0.7794 |
| 0.38 | 0.1480 | 0.3520 | 0.6480 | 0.78 | 0.2823 | 0.2177 | 0.7823 |
| 0.39 | 0.1517 | 0.3483 | 0.6517 | 0.79 | 0.2852 | 0.2148 | 0.7852 |




mean to z

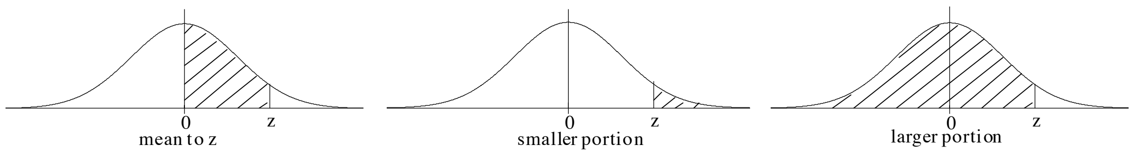



smaller portion

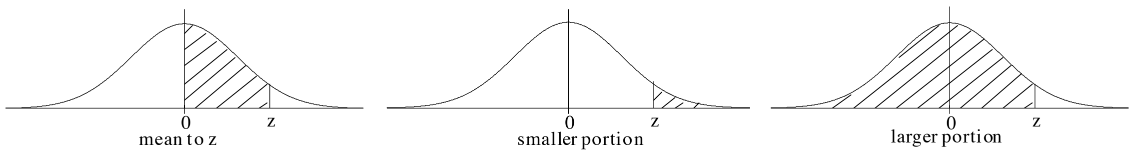



larger portion

| $z$ score | mean to $z$ | smaller portion | larger portion | $z$ score | mean to $z$ | smaller portion | larger portion |
|---|---|---|---|---|---|---|---|
| 0.80 | 0.2881 | 0.2119 | 0.7881 | 1.20 | 0.3849 | 0.1151 | 0.8849 |
| 0.81 | 0.2910 | 0.2090 | 0.7910 | 1.21 | 0.3869 | 0.1131 | 0.8869 |
| 0.82 | 0.2939 | 0.2061 | 0.7939 | 1.22 | 0.3888 | 0.1112 | 0.8888 |
| 0.83 | 0.2967 | 0.2033 | 0.7967 | 1.23 | 0.3907 | 0.1093 | 0.8907 |
| 0.84 | 0.2995 | 0.2005 | 0.7995 | 1.24 | 0.3925 | 0.1075 | 0.8925 |
| 0.85 | 0.3023 | 0.1977 | 0.8023 | 1.25 | 0.3944 | 0.1056 | 0.8944 |
| 0.86 | 0.3051 | 0.1949 | 0.8051 | 1.26 | 0.3962 | 0.1038 | 0.8962 |
| 0.87 | 0.3078 | 0.1922 | 0.8078 | 1.27 | 0.3980 | 0.1020 | 0.8980 |
| 0.88 | 0.3106 | 0.1894 | 0.8106 | 1.28 | 0.3997 | 0.1003 | 0.8997 |
| 0.89 | 0.3133 | 0.1867 | 0.8133 | 1.29 | 0.4015 | 0.0985 | 0.9015 |
| 0.90 | 0.3159 | 0.1841 | 0.8159 | 1.30 | 0.4032 | 0.0968 | 0.9032 |
| 0.91 | 0.3186 | 0.1814 | 0.8186 | 1.31 | 0.4049 | 0.0951 | 0.9049 |
| 0.92 | 0.3212 | 0.1788 | 0.8212 | 1.32 | 0.4066 | 0.0934 | 0.9066 |
| 0.93 | 0.3238 | 0.1762 | 0.8238 | 1.33 | 0.4082 | 0.0918 | 0.9082 |
| 0.94 | 0.3264 | 0.1736 | 0.8264 | 1.34 | 0.4099 | 0.0901 | 0.9099 |
| 0.95 | 0.3289 | 0.1711 | 0.8289 | 1.35 | 0.4115 | 0.0885 | 0.9115 |
| 0.96 | 0.3315 | 0.1685 | 0.8315 | 1.36 | 0.4131 | 0.0869 | 0.9131 |
| 0.97 | 0.3340 | 0.1660 | 0.8340 | 1.37 | 0.4147 | 0.0853 | 0.9147 |
| 0.98 | 0.3365 | 0.1635 | 0.8365 | 1.38 | 0.4162 | 0.0838 | 0.9162 |
| 0.99 | 0.3389 | 0.1611 | 0.8389 | 1.39 | 0.4177 | 0.0823 | 0.9177 |
| 1.00 | 0.3413 | 0.1587 | 0.8413 | 1.40 | 0.4192 | 0.0808 | 0.9192 |
| 1.01 | 0.3438 | 0.1562 | 0.8438 | 1.41 | 0.4207 | 0.0793 | 0.9207 |
| 1.02 | 0.3461 | 0.1539 | 0.8461 | 1.42 | 0.4222 | 0.0778 | 0.9222 |
| 1.03 | 0.3485 | 0.1515 | 0.8485 | 1.43 | 0.4236 | 0.0764 | 0.9236 |
| 1.04 | 0.3508 | 0.1492 | 0.8508 | 1.44 | 0.4251 | 0.0749 | 0.9251 |
| 1.05 | 0.3531 | 0.1469 | 0.8531 | 1.45 | 0.4265 | 0.0735 | 0.9265 |
| 1.06 | 0.3554 | 0.1446 | 0.8554 | 1.46 | 0.4279 | 0.0721 | 0.9279 |
| 1.07 | 0.3577 | 0.1423 | 0.8577 | 1.47 | 0.4292 | 0.0708 | 0.9292 |
| 1.08 | 0.3599 | 0.1401 | 0.8599 | 1.48 | 0.4306 | 0.0694 | 0.9306 |
| 1.09 | 0.3621 | 0.1379 | 0.8621 | 1.49 | 0.4319 | 0.0681 | 0.9319 |
| 1.10 | 0.3643 | 0.1357 | 0.8643 | 1.50 | 0.4332 | 0.0668 | 0.9332 |
| 1.11 | 0.3665 | 0.1335 | 0.8665 | 1.51 | 0.4345 | 0.0655 | 0.9345 |
| 1.12 | 0.3686 | 0.1314 | 0.8686 | 1.52 | 0.4357 | 0.0643 | 0.9357 |
| 1.13 | 0.3708 | 0.1292 | 0.8708 | 1.53 | 0.4370 | 0.0630 | 0.9370 |
| 1.14 | 0.3729 | 0.1271 | 0.8729 | 1.54 | 0.4382 | 0.0618 | 0.9382 |
| 1.15 | 0.3749 | 0.1251 | 0.8749 | 1.55 | 0.4394 | 0.0606 | 0.9394 |
| 1.16 | 0.3770 | 0.1230 | 0.8770 | 1.56 | 0.4406 | 0.0594 | 0.9406 |
| 1.17 | 0.3790 | 0.1210 | 0.8790 | 1.57 | 0.4418 | 0.0582 | 0.9418 |
| 1.18 | 0.3810 | 0.1190 | 0.8810 | 1.58 | 0.4429 | 0.0571 | 0.9429 |
| 1.19 | 0.3830 | 0.1170 | 0.8830 | 1.59 | 0.4441 | 0.0559 | 0.9441 |




| $z$ score | mean to $z$ | smaller portion | larger portion | $z$ score | mean to $z$ | smaller portion | larger portion |
|---|---|---|---|---|---|---|---|
| 1.60 | 0.4452 | 0.0548 | 0.9452 | 2.00 | 0.4772 | 0.0228 | 0.9772 |
| 1.61 | 0.4463 | 0.0537 | 0.9463 | 2.01 | 0.4778 | 0.0222 | 0.9778 |
| 1.62 | 0.4474 | 0.0526 | 0.9474 | 2.02 | 0.4783 | 0.0217 | 0.9783 |
| 1.63 | 0.4484 | 0.0516 | 0.9484 | 2.03 | 0.4788 | 0.0212 | 0.9788 |
| 1.64 | 0.4495 | 0.0505 | 0.9495 | 2.04 | 0.4793 | 0.0207 | 0.9793 |
| 1.65 | 0.4505 | 0.0495 | 0.9505 | 2.05 | 0.4798 | 0.0202 | 0.9798 |
| 1.66 | 0.4515 | 0.0485 | 0.9515 | 2.06 | 0.4803 | 0.0197 | 0.9803 |
| 1.67 | 0.4525 | 0.0475 | 0.9525 | 2.07 | 0.4808 | 0.0192 | 0.9808 |
| 1.68 | 0.4535 | 0.0465 | 0.9535 | 2.08 | 0.4812 | 0.0188 | 0.9812 |
| 1.69 | 0.4545 | 0.0455 | 0.9545 | 2.09 | 0.4817 | 0.0183 | 0.9817 |
| 1.70 | 0.4554 | 0.0446 | 0.9554 | 2.10 | 0.4821 | 0.0179 | 0.9821 |
| 1.71 | 0.4564 | 0.0436 | 0.9564 | 2.11 | 0.4826 | 0.0174 | 0.9826 |
| 1.72 | 0.4573 | 0.0427 | 0.9573 | 2.12 | 0.4830 | 0.0170 | 0.9830 |
| 1.73 | 0.4582 | 0.0418 | 0.9582 | 2.13 | 0.4834 | 0.0166 | 0.9834 |
| 1.74 | 0.4591 | 0.0409 | 0.9591 | 2.14 | 0.4838 | 0.0162 | 0.9838 |
| 1.75 | 0.4599 | 0.0401 | 0.9599 | 2.15 | 0.4842 | 0.0158 | 0.9842 |
| 1.76 | 0.4608 | 0.0392 | 0.9608 | 2.16 | 0.4846 | 0.0154 | 0.9846 |
| 1.77 | 0.4616 | 0.0384 | 0.9616 | 2.17 | 0.4850 | 0.0150 | 0.9850 |
| 1.78 | 0.4625 | 0.0375 | 0.9625 | 2.18 | 0.4854 | 0.0146 | 0.9854 |
| 1.79 | 0.4633 | 0.0367 | 0.9633 | 2.19 | 0.4857 | 0.0143 | 0.9857 |
| 1.80 | 0.4641 | 0.0359 | 0.9641 | 2.20 | 0.4861 | 0.0139 | 0.9861 |
| 1.81 | 0.4649 | 0.0351 | 0.9649 | 2.21 | 0.4864 | 0.0136 | 0.9864 |
| 1.82 | 0.4656 | 0.0344 | 0.9656 | 2.22 | 0.4868 | 0.0132 | 0.9868 |
| 1.83 | 0.4664 | 0.0336 | 0.9664 | 2.23 | 0.4871 | 0.0129 | 0.9871 |
| 1.84 | 0.4671 | 0.0329 | 0.9671 | 2.24 | 0.4875 | 0.0125 | 0.9875 |
| 1.85 | 0.4678 | 0.0322 | 0.9678 | 2.25 | 0.4878 | 0.0122 | 0.9878 |
| 1.86 | 0.4686 | 0.0314 | 0.9686 | 2.26 | 0.4881 | 0.0119 | 0.9881 |
| 1.87 | 0.4693 | 0.0307 | 0.9693 | 2.27 | 0.4884 | 0.0116 | 0.9884 |
| 1.88 | 0.4699 | 0.0301 | 0.9699 | 2.28 | 0.4887 | 0.0113 | 0.9887 |
| 1.89 | 0.4706 | 0.0294 | 0.9706 | 2.29 | 0.4890 | 0.0110 | 0.9890 |
| 1.90 | 0.4713 | 0.0287 | 0.9713 | 2.30 | 0.4893 | 0.0107 | 0.9893 |
| 1.91 | 0.4719 | 0.0281 | 0.9719 | 2.31 | 0.4896 | 0.0104 | 0.9896 |
| 1.92 | 0.4726 | 0.0274 | 0.9726 | 2.32 | 0.4898 | 0.0102 | 0.9898 |
| 1.93 | 0.4732 | 0.0268 | 0.9732 | 2.33 | 0.4901 | 0.0099 | 0.9901 |
| 1.94 | 0.4738 | 0.0262 | 0.9738 | 2.34 | 0.4904 | 0.0096 | 0.9904 |
| 1.95 | 0.4744 | 0.0256 | 0.9744 | 2.35 | 0.4906 | 0.0094 | 0.9906 |
| 1.96 | 0.4750 | 0.0250 | 0.9750 | 2.36 | 0.4909 | 0.0091 | 0.9909 |
| 1.97 | 0.4756 | 0.0244 | 0.9756 | 2.37 | 0.4911 | 0.0089 | 0.9911 |
| 1.98 | 0.4761 | 0.0239 | 0.9761 | 2.38 | 0.4913 | 0.0087 | 0.9913 |
| 1.99 | 0.4767 | 0.0233 | 0.9767 | 2.39 | 0.4916 | 0.0084 | 0.9916 |




| $z$ score | mean to $z$ | smaller portion | larger portion | $z$ score | mean to $z$ | smaller portion | larger portion |
|---|---|---|---|---|---|---|---|
| 2.40 | 0.4918 | 0.0082 | 0.9918 | 2.80 | 0.4974 | 0.0026 | 0.9974 |
| 2.41 | 0.4920 | 0.0080 | 0.9920 | 2.81 | 0.4975 | 0.0025 | 0.9975 |
| 2.42 | 0.4922 | 0.0078 | 0.9922 | 2.82 | 0.4976 | 0.0024 | 0.9976 |
| 2.43 | 0.4925 | 0.0075 | 0.9925 | 2.83 | 0.4977 | 0.0023 | 0.9977 |
| 2.44 | 0.4927 | 0.0073 | 0.9927 | 2.84 | 0.4977 | 0.0023 | 0.9977 |
| 2.45 | 0.4929 | 0.0071 | 0.9929 | 2.85 | 0.4978 | 0.0022 | 0.9978 |
| 2.46 | 0.4931 | 0.0069 | 0.9931 | 2.86 | 0.4979 | 0.0021 | 0.9979 |
| 2.47 | 0.4932 | 0.0068 | 0.9932 | 2.87 | 0.4979 | 0.0021 | 0.9979 |
| 2.48 | 0.4934 | 0.0066 | 0.9934 | 2.88 | 0.4980 | 0.0020 | 0.9980 |
| 2.49 | 0.4936 | 0.0064 | 0.9936 | 2.89 | 0.4981 | 0.0019 | 0.9981 |
| 2.50 | 0.4938 | 0.0062 | 0.9938 | 2.90 | 0.4981 | 0.0019 | 0.9981 |
| 2.51 | 0.4940 | 0.0060 | 0.9940 | 2.91 | 0.4982 | 0.0018 | 0.9982 |
| 2.52 | 0.4941 | 0.0059 | 0.9941 | 2.92 | 0.4982 | 0.0018 | 0.9982 |
| 2.53 | 0.4943 | 0.0057 | 0.9943 | 2.93 | 0.4983 | 0.0017 | 0.9983 |
| 2.54 | 0.4945 | 0.0055 | 0.9945 | 2.94 | 0.4984 | 0.0016 | 0.9984 |
| 2.55 | 0.4946 | 0.0054 | 0.9946 | 2.95 | 0.4984 | 0.0016 | 0.9984 |
| 2.56 | 0.4948 | 0.0052 | 0.9948 | 2.96 | 0.4985 | 0.0015 | 0.9985 |
| 2.57 | 0.4949 | 0.0051 | 0.9949 | 2.97 | 0.4985 | 0.0015 | 0.9985 |
| 2.58 | 0.4951 | 0.0049 | 0.9951 | 2.98 | 0.4986 | 0.0014 | 0.9986 |
| 2.59 | 0.4952 | 0.0048 | 0.9952 | 2.99 | 0.4986 | 0.0014 | 0.9986 |
| 2.60 | 0.4953 | 0.0047 | 0.9953 | 3.00 | 0.4987 | 0.0013 | 0.9987 |
| 2.61 | 0.4955 | 0.0045 | 0.9955 | | | | |
| 2.62 | 0.4956 | 0.0044 | 0.9956 | 3.25 | 0.4994 | 0.0006 | 0.9994 |
| 2.63 | 0.4957 | 0.0043 | 0.9957 | | | | |
| 2.64 | 0.4959 | 0.0041 | 0.9959 | 3.50 | 0.4998 | 0.0002 | 0.9998 |
| 2.65 | 0.4960 | 0.0040 | 0.9960 | | | | |
| 2.66 | 0.4961 | 0.0039 | 0.9961 | 3.75 | 0.4999 | 0.0001 | 0.9999 |
| 2.67 | 0.4962 | 0.0038 | 0.9962 | | | | |
| 2.68 | 0.4963 | 0.0037 | 0.9963 | 4.00 | 0.5000 | 0.0000 | 1.0000 |
| 2.69 | 0.4964 | 0.0036 | 0.9964 | | | | |
| 2.70 | 0.4965 | 0.0035 | 0.9965 | | | | |
| 2.71 | 0.4966 | 0.0034 | 0.9966 | | | | |
| 2.72 | 0.4967 | 0.0033 | 0.9967 | | | | |
| 2.73 | 0.4968 | 0.0032 | 0.9968 | | | | |
| 2.74 | 0.4969 | 0.0031 | 0.9969 | | | | |
| 2.75 | 0.4970 | 0.0030 | 0.9970 | | | | |
| 2.76 | 0.4971 | 0.0029 | 0.9971 | | | | |
| 2.77 | 0.4972 | 0.0028 | 0.9972 | | | | |
| 2.78 | 0.4973 | 0.0027 | 0.9973 | | | | |
| 2.79 | 0.4974 | 0.0026 | 0.9974 | | | | |

**Mathematics Learning Centre**
**T** +61 2 9351 4061
**F** +61 2 9351 5797
**E** mlc.enquiries@sydney.edu.au
**sydney.edu.au/mlc**

MATHEMATICS
LEARNING CENTRE

THE UNIVERSITY OF
SYDNEY

Produced by UPS, the University of Sydney, January 2010.


ABN 15 211 513 464
CRICOS 00026A