

The Normal Distribution

39.1	The Normal Distribution	2
39.2	The Normal Approximation to the Binomial Distribution	26
39.3	Sums and Differences of Random Variables	34

Learning outcomes

In a previous Workbook you learned what a continuous random variable was. Here you will examine the most important example of a continuous random variable: the normal distribution. The probabilities of the normal distribution have to be determined numerically. Tables of such probabilities, which refer to a simplified normal distribution called the standard normal distribution, which has mean 0 and variance 1, will be used to determine probabilities of the general normal distribution. Finally you will learn how to deal with combinations of random variables which is an important statistical tool applicable to many engineering situations.

The Normal Distribution

39.1

Introduction

Mass-produced items should conform to a specification. Usually, a mean is aimed for but due to random errors in the production process a tolerance is set on deviations from the mean. For example if we produce piston rings which have a target mean internal diameter of 45 mm then realistically we expect the diameter to deviate only slightly from this value. The deviations from the mean value are often modelled very well by the **normal distribution**. Suppose we decide that diameters in the range 44.95 mm to 45.05 mm are acceptable, then what proportion of the output is satisfactory? In this Section we shall see how to use the normal distribution to answer questions like this.



Prerequisites

Before starting this Section you should ...

- be familiar with the basic properties of probability
- be familiar with continuous random variables



Learning Outcomes

On completion you should be able to ...

- recognise the shape of the frequency curve for the normal distribution and the standard normal distribution
- calculate probabilities using the standard normal distribution
- recognise key areas under the frequency curve

1. The normal distribution

The normal distribution is the most widely used model for the distribution of a random variable. There is a very good reason for this. Practical experiments involve measurements and measurements involve errors. However you go about measuring a quantity, inaccuracies of all sorts can make themselves felt. For example, if you are measuring a length using a device as crude as a ruler, you may find errors arising due to:

- the calibration of the ruler itself;
- parallax errors due to the relative positions of the object being measured, the ruler and your eye;
- rounding errors;
- 'guesstimation' errors if a measurement is between two marked lengths on the ruler.
- mistakes.

If you use a meter with a digital readout, you will avoid some of the above errors but others, often present in the design of the electronics controlling the meter, will be present. Errors are unavoidable and are usually the sum of several factors. The behaviour of variables which are the sum of several other variables is described by a very important and powerful result called the Central Limit Theorem which we will study later in this Workbook. For now we will quote the result so that the importance of the normal distribution will be appreciated.

The central limit theorem

Let X be the sum of n independent random variables $X_i, i = 1, 2, \dots, n$ each having a distribution with mean μ_i and variance σ_i^2 ($\sigma_i^2 < \infty$), respectively, then the distribution of X has expectation and variance given by the expressions

$$E(X) = \sum_{i=1}^n \mu_i \quad \text{and} \quad V(X) = \sum_{i=1}^n \sigma_i^2$$

and becomes **normal** as $n \rightarrow \infty$.

Essentially we are saying that a quantity which represents the combined effect of a number of variables will be approximately normal no matter what the original distributions are provided that $\sigma^2 < \infty$. This statement is true for the vast majority of distributions you are likely to meet in practice. This is why the normal distribution is crucially important to engineers. A quotation attributed to Prof. G. Lippmann, (1845-1921, winner of the Nobel prize for Physics in 1908) 'Everybody believes on the law of errors, experimenters because they think it is a mathematical theorem and mathematicians because they think it is an experimental fact.'

You may think that anything you measure follows an approximate normal distribution. Unfortunately this is not the case. While the heights of human beings follow a normal distribution, weights do not. Heights are the result of the interaction of many factors (outside one's control) while weights principally depend on lifestyle (including how much and what you eat and drink!) In practice, it is found that weight is skewed to the right but that the square root of human weights is approximately normal.

The probability density function of a normal distribution with mean μ and variance σ^2 is given by the formula

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

This curve is always bell-shaped with the centre of the bell located at the value of μ . See Figure 1. The height of the bell is controlled by the value of σ . As with all normal distribution curves it is symmetrical about the centre and decays as $x \rightarrow \pm\infty$. As with any probability density function the area under the curve is equal to 1.

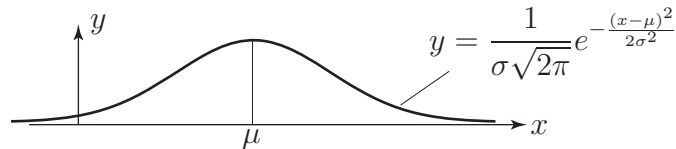


Figure 1

A normal distribution is completely defined by specifying its mean (the value of μ) and its variance (the value of σ^2 .) The normal distribution with mean μ and variance σ^2 is written $N(\mu, \sigma^2)$. Hence the distribution $N(20, 25)$ has a mean of 20 and a standard deviation of 5; remember that the second parameter is the variance which is the square of the standard deviation.



Key Point 1

A normal distribution has mean μ and variance σ^2 . A random variable X following this distribution is usually denoted by $N(\mu, \sigma^2)$ and we often write

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

Clearly, since μ and σ^2 can both vary, there are infinitely many normal distributions and it is impossible to give tabulated information concerning them all.

For example, if we produce piston rings which have a target mean internal diameter of 45 mm then we may realistically expect the actual diameter to deviate from this value. Such deviations are well-modelled by the normal distribution. Suppose we decide that diameters in the range 44.95 mm to 45.05 mm are acceptable, we may then ask the question ‘What proportion of our manufactured output is satisfactory?’

Without tabulated data concerning the appropriate normal distribution we cannot easily answer this question (because the integral used to calculate areas under the normal curve is intractable.)

Since tabulated data allow us to apply the distribution to a wide variety of statistical situations, and we cannot tabulate all normal distributions, we tabulate only one - the standard normal distribution - and convert all problems involving the normal distribution into problems involving the standard normal distribution.

2. The standard normal distribution

At this stage we shall, for simplicity, consider what is known as a standard normal distribution which is obtained by choosing particularly simple values for μ and σ .



Key Point 2

The **standard normal distribution** has a mean of zero and a variance of one.

In Figure 2 we show the graph of the standard normal distribution which has probability density function $y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

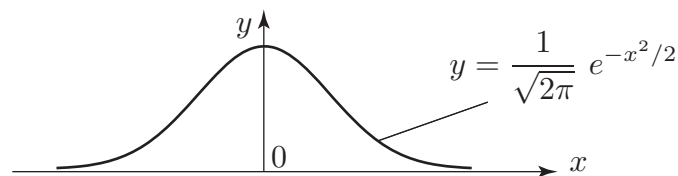


Figure 2: The standard normal distribution curve

The result which makes the standard normal distribution so important is as follows:



Key Point 3

If the behaviour of a continuous random variable X is described by the distribution $N(\mu, \sigma^2)$ then the behaviour of the random variable $Z = \frac{X - \mu}{\sigma}$ is described by the standard normal distribution $N(0, 1)$.

We call Z the **standardised normal variable** and we write

$$Z \sim N(0, 1)$$



Example 1

If the random variable X is described by the distribution $N(45, 0.000625)$ then what is the transformation required to obtain the standardised normal variable?

Solution

Here, $\mu = 45$ and $\sigma^2 = 0.000625$ so that $\sigma = 0.025$. Hence $Z = (X - 45)/0.025$ is the required transformation.



Example 2

When the random variable $X \sim N(45, 0.000625)$ takes values between 44.95 and 45.05, between which values does the random variable Z lie?

Solution

$$\text{When } X = 45.05, Z = \frac{45.05 - 45}{0.025} = 2$$

$$\text{When } X = 44.95, Z = \frac{44.95 - 45}{0.025} = -2$$

Hence Z lies between -2 and 2 .



The random variable X follows a normal distribution with mean 1000 and variance 100. When X takes values between 1005 and 1010, between which values does the standardised normal variable Z lie?

Your solution

Answer

The transformation is $Z = \frac{X - 1000}{10}$.

$$\text{When } X = 1005, Z = \frac{5}{10} = 0.5$$

$$\text{When } X = 1010, Z = \frac{10}{10} = 1.$$

Hence Z lies between 0.5 and 1.

3. Probabilities and the standard normal distribution

Since the standard normal distribution is used so frequently a table of values has been produced to help us calculate probabilities - located at the end of the Workbook. It is based upon the following diagram:

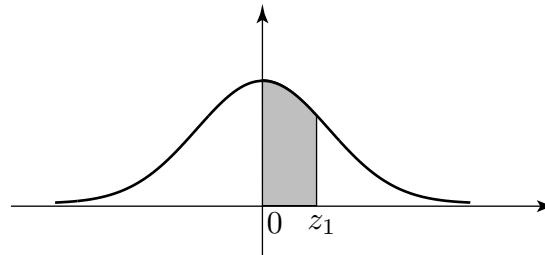


Figure 3

Since the total area under the curve is equal to 1 it follows from the symmetry in the curve that the area under the curve in the region $Z > 0$ is equal to 0.5. In Figure 3 the shaded area is the probability that Z takes values between 0 and z_1 . When we 'look-up' a value in the table we obtain the value of the shaded area.



Example 3

What is the probability that Z takes values between 0 and 1.9? (Refer to the table of normal probabilities at the end of the Workbook.)

Solution

The row beginning '1.9' and the column headed '0' is the appropriate choice and its entry is 4713. This is to be read as 0.4713 (we omitted the '0.' in each entry for clarity.) The interpretation is that the probability that Z takes values between 0 and 1.9 is 0.4713.



Example 4

What is the probability that Z takes values between 0 and 1.96?

Solution

This time we want the row beginning 1.9 and the column headed '6'. The entry is 4750 so that the required probability is 0.4750.



Example 5

What is the probability that Z takes values between 0 and 1.965?

Solution

There is no entry corresponding to 1.965 so we take the average of the values for 1.96 and 1.97. (This linear interpolation is not strictly correct but is acceptable.)

The two values are 4750 and 4756 with an average of 4753. Hence the required probability is 0.4753.



What are the probabilities that Z takes values between

- (a) 0 and 2 (b) 0 and 2.3 (c) 0 and 2.33 (d) 0 and 2.333?

Your solution

Answer

- (a) The entry is 4772; the probability is 0.4772.
(b) The entry is 4893; the probability is 0.4893.
(c) The entry is 4901; the probability is 0.4901.
(d) The entry for 2.33 is 4901, that for 2.34 is 4904.

Linear interpolation gives a value of $4901 + 0.3(4904 - 4901)$ i.e. about 4902; the probability is 0.4902.

Note from Table 1 that as Z increases from 0 the entries increase, rapidly at first and then more slowly, toward 5000 i.e. a probability of 0.5. This is consistent with the shape of the curve.

After $Z = 3$ the increase is quite slow so that we tabulate entries for values of Z rising by increments of 0.1 instead of 0.01 as in the rest of Table 1.

4. Calculating other probabilities

In this Section we see how to calculate probabilities represented by areas other than those of the type shown in Figure 3.

Case 1

Figure 4 illustrates what we do if both Z values are positive. By using the properties of the standard normal distribution we can organise matters so that any required area is always of 'standard form'.

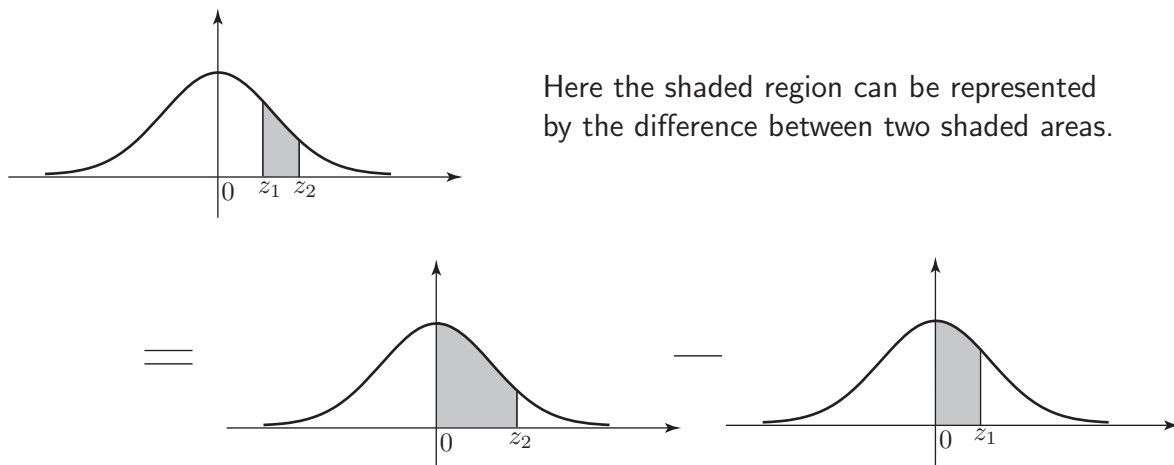


Figure 4



Example 6

Find the probability that Z takes values between 1 and 2.

Solution

Using Table 1:

$P(Z = z_2)$ i.e. $P(Z = 2)$ is 0.4772

$P(Z = z_1)$ i.e. $P(Z = 1)$ is 0.3413.

Hence $P(1 < Z < 2) = 0.4772 - 0.3413 = 0.1359$

Remember that with a continuous distribution, $P(Z = 1)$ is meaningless (will have zero probability) so that $P(1 \leq Z \leq 2)$ is interpreted as $P(1 < Z < 2)$.

Case 2

The following diagram illustrates the procedure to be followed when finding probabilities of the form $P(Z > z_1)$.

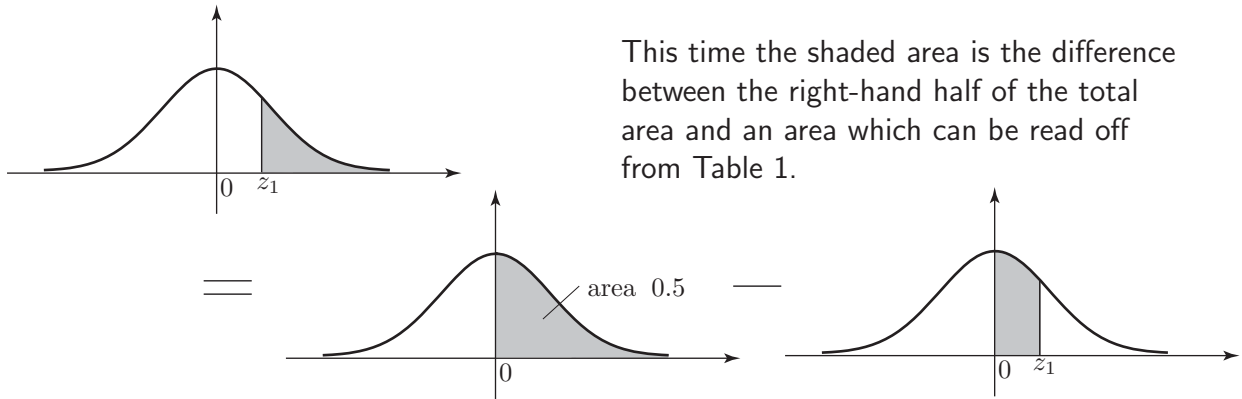


Figure 5



Example 7

What is the probability that $Z > 2$?

Solution

$P(0 < Z < 2) = 0.4772$ (from Table 1). Hence the probability is $0.5 - 0.4772 = 0.0228$.

Case 3

Here we consider the procedure to be followed when calculating probabilities of the form $P(Z < z_1)$. Here the shaded area is the sum of the left-hand half of the total area and a 'standard' area.

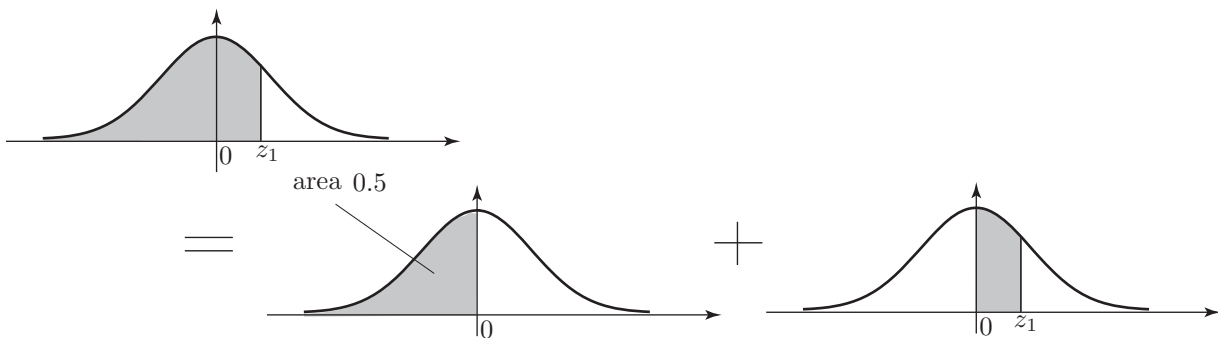


Figure 6

**Example 8**

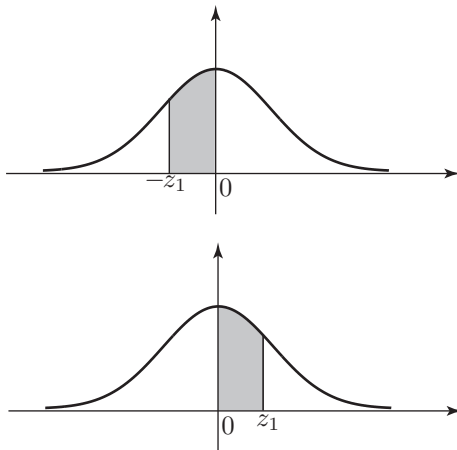
What is the probability that $Z < 2$?

Solution

$$P(Z < 2) = 0.5 + 0.4772 = 0.9772.$$

Case 4

Here we consider what needs to be done when calculating probabilities of the form $P(-z_1 < Z < 0)$ where z_1 is positive. This time we make use of the symmetry in the standard normal distribution curve.



By symmetry this shaded area is equal in value to the one above.

Figure 7**Example 9**

What is the probability that $-2 < Z < 0$?

Solution

The area is equal to that corresponding to $P(0 < Z < 2) = 0.4772$.

Case 5

Finally we consider probabilities of the form $P(-z_2 < Z < z_1)$. Here we use the sum property and the symmetry property.

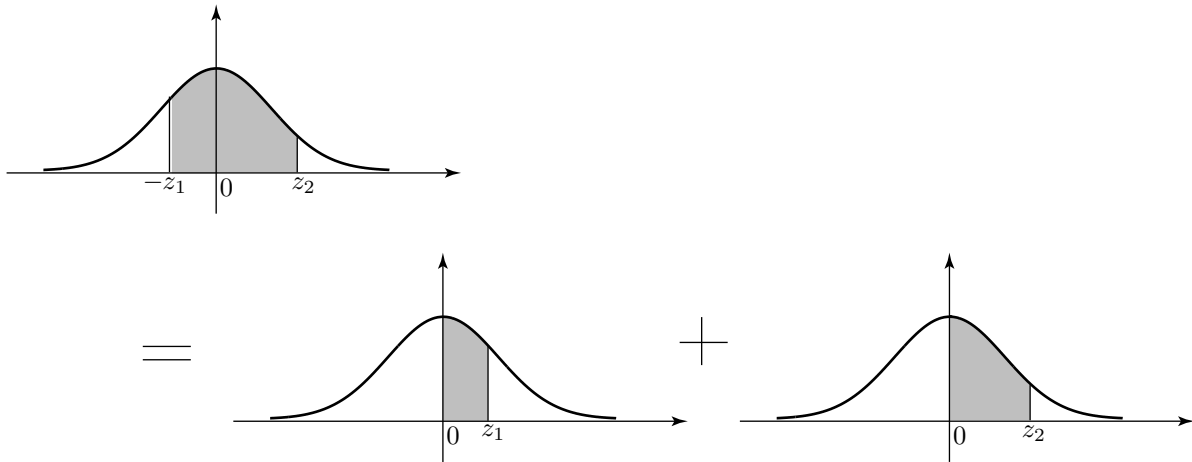


Figure 8



Example 10

What is the probability that $-1 < Z < 2$?

Solution

$$P(-1 < Z < 0) = P(0 < Z < 1) = 0.3413$$

$$P(0 < Z < 2) = 0.4772$$

Hence the required probability $P(-1 < Z < 2)$ is 0.8185.

Other cases can be handled by a combination of the ideas already used.



Find the following probabilities.

- (a) $P(0 < Z < 1.5)$ (b) $P(Z > 1.8)$
(c) $P(1.5 < Z < 1.8)$ (d) $P(Z < 1.8)$
(e) $P(-1.5 < Z < 0)$ (f) $P(Z < -1.5)$
(g) $P(-1.8 < Z < -1.5)$ (h) $P(-1.5 < Z < 1.8)$

(A simple sketch of the standard normal curve will help.)

Your solution

Answer

- (a) 0.4332 (direct from Table 1)
(b) $0.5 - 0.4641 = 0.0359$
(c) $P(0 < Z < 1.8) - P(0 < Z < 1.5) = 0.4641 - 0.4332 = 0.0309$
(d) $0.5 + 0.4641 = 0.9641$
(e) $P(-1.5 < Z < 0) = P(0 < Z < 1.5) = 0.4332$
(f) $P(Z < -1.5) = P(Z > 1.5) = 0.5 - 0.4332 = 0.0668$
(g) $P(-1.8 < Z < -1.5) = P(1.5 < Z < 1.8) = 0.0309$
(h) $P(0 < Z < 1.5) + P(0 < Z < 1.8) = 0.8973$

5. The cumulative distribution function

We know that the normal probability density function $f(x)$ is given by the formula

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

and so the cumulative distribution function $F(x)$ is given by the formula

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(u-\mu)^2/2\sigma^2} du$$

In the case of the cumulative distribution for the standard normal curve, we use the special notation $\Phi(z)$ and, substituting 0 and 1 for μ and σ^2 , we obtain

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du$$

The shape of the curve is essentially 'S' -shaped as shown in Figure 9. Note that the curve runs from $-\infty$ to $+\infty$. As you can see, the curve approaches the value 1 asymptotically.

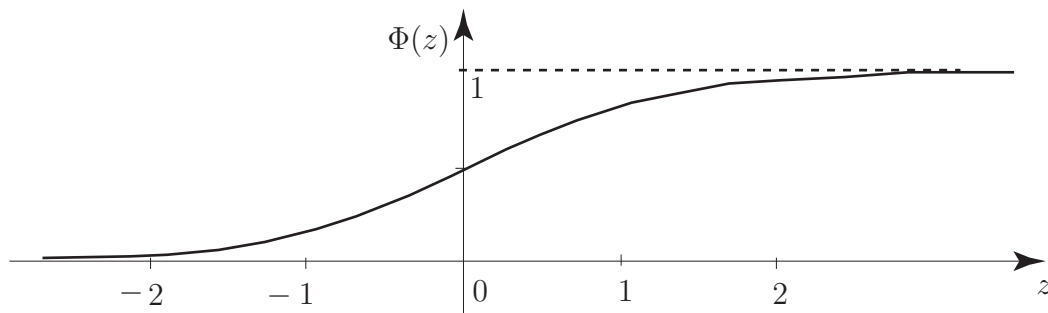


Figure 9

Comparing the integrals

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(u-\mu)^2/2\sigma^2} du \quad \text{and} \quad \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-v^2/2} dv$$

shows that

$$v = \frac{u - \mu}{\sigma} \quad \text{and so} \quad dv = \frac{du}{\sigma}$$

and $F(x)$ may be written as

$$\begin{aligned} F(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-v^2/2} \sigma dv \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-v^2/2} dv = \Phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

We already know, from the basic definition of a cumulative distribution function, that

$$P(a < X < b) = F(b) - F(a)$$

so that we may write the probability statement above in terms of $\Phi(z)$ as

$$P(a < X < b) = F(b) - F(a) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

The value of $\Phi(z)$ is measured from $z = -\infty$ to any ordinate $z = z_1$ and represents the probability $P(Z < z_1)$.

The values of $\Phi(z)$ start as shown below:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	5040	5080	5120	5160	5199	5239	5279	5319	5359
0.1	.5398	5438	5478	5517	5577	5596	5636	5675	5714	5753
0.2	.5793	5832	5871	5909	5948	5987	6026	6064	6103	6141

You should compare the values given here with the values given for the normal probability integral (Table 1 at the end of the Workbook). Simply adding 0.5 to the values in the latter table gives the values of $\Phi(z)$. You should also note that the diagrams shown at the top of each set of tabulated values tells you whether you are looking at the values of $\Phi(z)$ or the values of the normal probability integral.

Exercises

1. If a random variable X has a standard normal distribution find the probability that it assumes a value:
 - (a) less than 2.00
 - (b) greater than 2.58
 - (c) between 0 and 1.00
 - (d) between -1.65 and -0.84
2. If X has a standard normal distribution find k in each of the following cases:
 - (a) $P(X < k) = 0.4$
 - (b) $P(X < k) = 0.95$
 - (c) $P(0 < X < k) = 0.1$

Answers

- 1 (a) 0.9772 (b) 0.0049 (c) 0.3413 (d) 0.1510
2 (a) -0.2533 (b) 1.6450 (c) 0.2533

6. Applications of the normal distribution

We have, in the previous subsection, noted that the probability density function of a normal distribution X is

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

This curve is always 'bell-shaped' with the centre of the bell located at the value of μ . The height of the bell is controlled by the value of σ . See Figure 10.

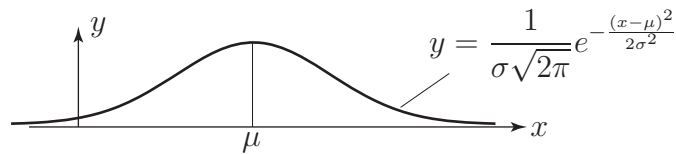


Figure 10

We now show, by example, how probabilities relating to a general normal distribution X are determined. We will see that being able to calculate the probabilities of a standard normal distribution Z is crucial in this respect.



Example 11

Given that the variate X follows the normal distribution $X \sim N(151, 15^2)$, calculate:

- (a) $P(120 \leq X \leq 155)$; (b) $P(X \geq 185)$

Solution

The transformation used in this problem is $Z = \frac{X - \mu}{\sigma} = \frac{X - 151}{15}$

(a)

$$\begin{aligned} P(120 \leq X \leq 155) &= P\left(\frac{120 - 151}{15} \leq Z \leq \frac{155 - 151}{15}\right) \\ &= P(-2.07 \leq Z \leq 0.27) \\ &= 0.4808 + 0.1064 = 0.5872 \end{aligned}$$

(b)

$$\begin{aligned} P(X \geq 185) &= P\left(Z \geq \frac{185 - 151}{15}\right) \\ &= P(Z \geq 2.27) \\ &= 0.5 - 0.4884 = 0.0116 \end{aligned}$$

We note that, as for any continuous random variable, we can only calculate the probability that

- X lies between two given values;
- X is greater than a given value;
- X is less than a given value.

rather than for individual values.



A worn, poorly set-up machine is observed to produce components whose length X follows a normal distribution with mean 20 cm and variance 2.56 cm Calculate:

- the probability that a component is at least 24 cm long;
- the probability that the length of a component lies between 19 and 21 cm.

Your solution

Answer

The transformation used is $Z = \frac{X - 20}{1.6}$ giving

$$P(X \geq 24) = P\left(Z \geq \frac{24 - 20}{1.6}\right) = P(Z \geq 2.5) = 0.5 - 0.4938 = 0.0062$$

and

$$P(19 < X < 21) = P\left(\frac{19 - 20}{1.6} < Z < \frac{21 - 20}{1.6}\right) = P(-0.625 < Z < 0.625) = 0.4681$$



Example 12

Piston rings are mass-produced. The target internal diameter is 45 mm but records show that the diameters are normally distributed with mean 45 mm and standard deviation 0.05 mm. An acceptable diameter is one within the range 44.95 mm to 45.05 mm. What proportion of the output is unacceptable?

Solution

There are many words in the statement of the problem; we must read them carefully to extract the necessary information. If X is the diameter of a piston ring then $X \sim N(45, (0.05)^2)$.

The transformation is $Z = \frac{X - \mu}{\sigma} = \frac{X - 45}{0.05}$.

The upper limit of acceptability is $x_2 = 45.05$ so that $z_2 = (45.05 - 45)/0.05 = 1$.

The lower limit of acceptability is $x_1 = 44.95$ so that $z_1 = (44.95 - 45)/0.05 = -1$.

The range of 'acceptable' Z values is therefore -1 to 1 . Figure 11 below.

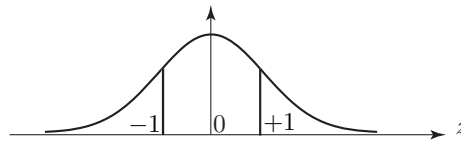


Figure 11

Using the symmetry of the curves $P(-1 < Z < 1) = 2 \times P(0 < Z < 1) = 2 \times 0.3413 = 0.6826$.

Thus the proportion of unacceptable items is $1 - 0.6826 = 0.3174$, or 31.74%.



Example 13

If the standard deviation is halved by improved production practices what is now the proportion of unacceptable items?

Solution

Now $\sigma = 0.025$ so that:

$$z_2 = \frac{45.05 - 45}{0.025} = 2 \quad \text{and} \quad z_1 = -2$$

Then $P(-2 < Z < 2) = 2 \times P(0 < Z < 2) = 2 \times 0.4772 = 0.9544$. Hence the proportion of unacceptable items is reduced to $1 - 0.9544 = 0.0456$ or 4.56%.

We observe that less of the area under the curve now lies outside the interval $(44.95, 45.05)$.

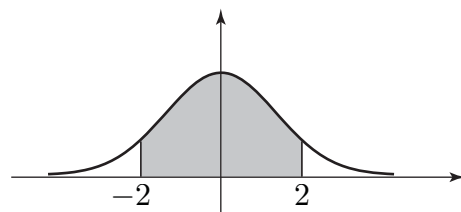


Figure 12



The resistance of a strain gauge is normally distributed with a mean of 100 ohms and a standard deviation of 0.2 ohms. To meet the specification, the resistance must be within the range 100 ± 0.5 ohms.

(a) What percentage of gauges are unacceptable?

First, state the upper and lower limits of acceptable resistance and find the Z -values which correspond:

Your solution

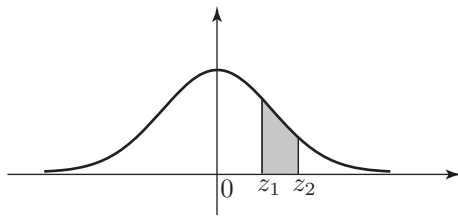
Answer

$$x_1 = 99.5, \quad x_2 = 100.5 \quad Z = \frac{X - 100}{0.2} \quad \{(0.2)^2 = 0.04\} \quad \text{so that } z_1 = -2.5 \text{ and } z_2 = 2.5$$

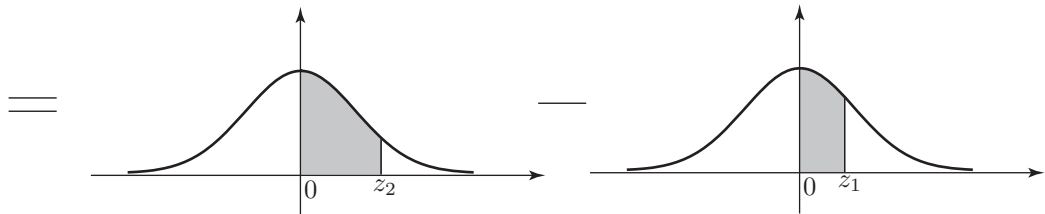
Now, using a suitable sketch, calculate the probability that $z_1 < Z < z_2$:

Your solution

Answer



Here the shaded region can be represented by the difference between two shaded areas.



The shaded area (see diagram) is 0.4938 (from the table of values on page 15). Using symmetry,

$$\begin{aligned} P(-2.5 < Z < 2.5) &= 2 \times 0.4938 \\ &= 0.9876. \end{aligned}$$

Hence the proportion of acceptable gauges is 98.76%.

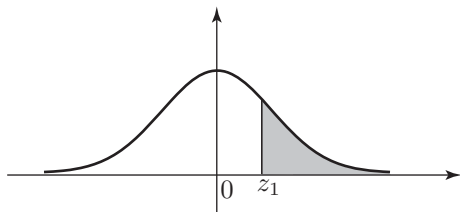
Therefore the proportion of unacceptable gauges is 1.24%.

(b) To what value must the standard deviation be reduced if the proportion of unacceptable gauges is to be no more than 0.2%?

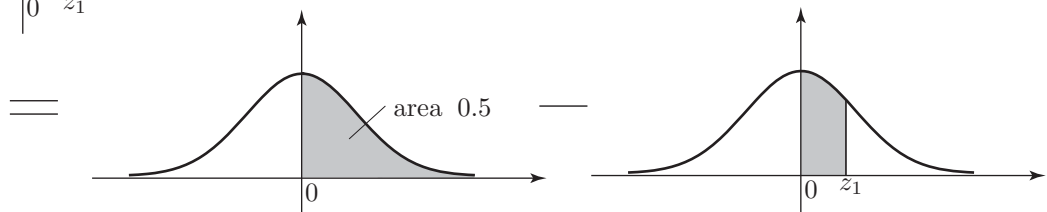
First sketch the standard normal curve marking on it the lower and upper values z_1 and z_2 and appropriate areas:

Your solution

Answer



This time the shaded area is the difference between the right-hand half of the total area and an area which can be read off from Table 1.



Now use the Table to find z_2 , and hence write down the value of z_1 :

Your solution

Answer

$$z_2 = 3.1 \quad \text{so that} \quad z_1 = -3.1$$

Finally, rewrite $Z = \frac{X - \mu}{\sigma}$ to make σ the subject. Put in values for z_2 , x_2 and μ hence evaluate σ :

Your solution

Answer

$$\sigma = \frac{X - \mu}{Z} = \frac{100.5 - 100}{3.1} = 0.16 \text{ (2 d.p.)}$$

7. Probability intervals - standard normal distribution

We use probability models to make predictions in situations where there is not sufficient data available to make a definite statement. Any statement based on these models carries with it a **risk** of being proved incorrect by events. Notice that the normal probability curve extends to infinity in both directions. **Theoretically** any value of the normal random variable is possible, although, of course, values far from the mean position (zero) are very unlikely.

Consider the diagram in Figure 13:

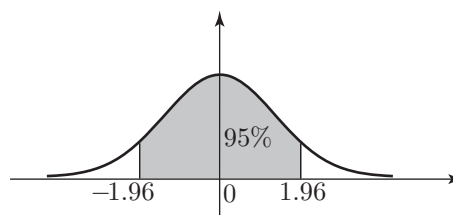


Figure 13

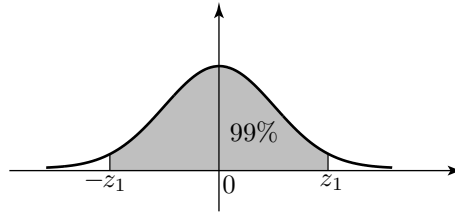
The shaded area is 95% of the total area. If we look at the entry in Table 1 (at the end of the Workbook) corresponding to $Z = 1.96$ we see the value 4750. This means that the probability of Z taking a value between 0 and 1.96 is 0.475. By symmetry, the probability that Z takes a value between -1.96 and 0 is also 0.475. Combining these results we see that

$$P(-1.96 < Z < 1.96) = 0.95 \text{ or } 95\%$$

We say that the 95% probability interval for Z (about its mean of 0) is $(-1.96, 1.96)$. It follows that there is a 5% chance that Z lies outside this interval.



Find the 99% probability interval for Z about its mean, i.e. the value of z_1 in the diagram:



The shaded area is 99% of the total area

First, note that 99% corresponds to a probability of 0.99. Find z_1 such that

$$P(0 < Z < z_1) = \frac{1}{2} \times 0.99 = 0.495 :$$

Your solution

Answer

We look for a table value of 4950. The nearest we get is 4949 and 4951 corresponding to $Z = 2.57$ and $Z = 2.58$ respectively. We choose $Z = 2.58$.

Now quote the 99% probability interval:

Your solution

Answer

$(-2.58, 2.58)$ or $-2.58 < Z < 2.58$.

Notice that the risk of Z lying outside this wider interval is reduced to 1%.



Find the value of Z

- (a) which is exceeded on 5% of occasions
- (b) which is exceeded on 99% of occasions.

Your solution**Answer**

(a) The value is z_1 , where $P(Z > z_1) = 0.05$. Hence $P(0 < Z < z_1) = 0.5 - 0.05 = 0.45$. This corresponds to a table entry of 4500. The nearest values are 4495 ($Z = 1.64$) and 4505 ($Z = 1.65$).

Hence the required value is $Z_1 = 1.65$.

(b) Values less than z_1 occur on 1% of occasions. By symmetry values greater than $(-z_1)$ occur on 1% of occasions so that $P(0 < z < -z_1) = 0.49$. The nearest table corresponding to 4900 is 4901 ($Z = 2.33$).

Hence the required value is $z_1 = -2.33$.

8. Probability intervals - general normal distribution

We saw in subsection 3 that 95% of the area under the standard normal curve lay between $z_1 = -1.96$ and $z_2 = 1.96$. Using the formula $Z = \frac{X - \mu}{\sigma}$ in the re-arrangement $X = \mu + Z\sigma$. We can see that 95% of the area under the general normal curve lies between $x_1 = \mu - 1.96\sigma$ and $x_2 = \mu + 1.96\sigma$.

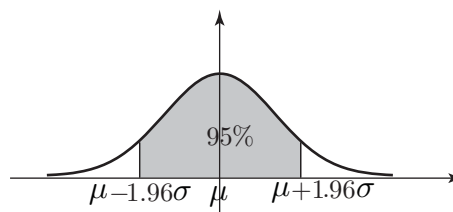


Figure 14



Example 14

Suppose that the internal diameters of mass-produced pipes are normally distributed with mean 50 mm and standard deviation 2 mm. What are the 95% probability limits on the internal diameter of a single pipe?

Solution

Here $\mu = 50$ $\sigma = 2$ so that the 95% probability limits are

$$50 \pm 1.96 \times 2 = 50 \pm 3.92\text{mm}$$

i.e. 46.08 mm and 53.92 mm.

The probability interval is (46.08, 53.92).

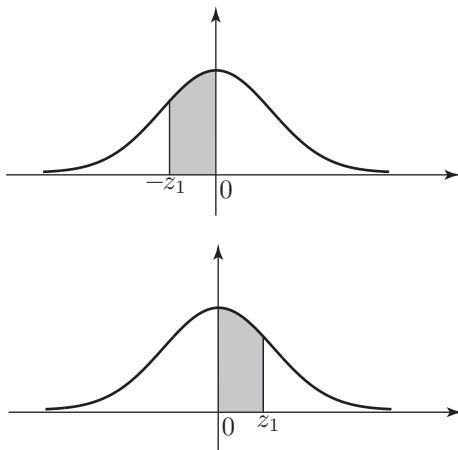


What is the 99% probability interval for the lifetime of a bulb when the lifetimes of such bulbs are normally distributed with a mean of 2000 hours and standard deviation of 40 hours?

First sketch the standard normal curve marking the values z_1, z_2 between which 99% of the area under the curve is located:

Your solution

Answer



By symmetry this shaded area is equal in value to the one above.

Now deduce the corresponding values x_1, x_2 for the general normal distribution:

Your solution

Answer

$$x_1 = \mu - 2.58\sigma, \quad x_2 = \mu + 2.58\sigma$$

Next, find the values for x_1 and x_2 for the given problem:

Your solution

Answer

$$x_1 = 2000 - 2.58 \times 40 = 1896.8 \text{ hours}$$

$$x_2 = 2000 + 2.58 \times 40 = 2103.2 \text{ hours}$$

Finally, write down the 99% probability interval for the lifetimes:

Your solution

Answer

(1896.8 hours, 2103.2 hours).