

# Introduction to Confidence Intervals

**40.2**



## Introduction

In this Section you will learn how to construct a variety of confidence intervals for the mean and variance of a distribution, which have wide applicability in engineering.



## Prerequisites

Before starting this Section you should ...

- ① be familiar with the results and concepts met in probability
- ② be able to calculate probabilities using the standard normal distribution
- ③ know how to use the table of standard normal probabilities



## Learning Outcomes

After completing this Section you should be able to ...

- ✓ calculate a variety of intervals for the mean of a distribution
- ✓ calculate a variety of intervals for the variance of a distribution

# 1. Confidence 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 9,

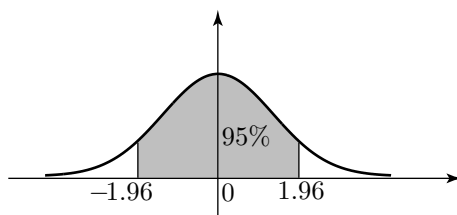


Figure 9

The shaded area is 95% of the total area. If we look at the entry in the Table of values (page 15) 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% confidence 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.



We wish to find the 99% confidence interval for  $Z$  about its mean, i.e. the value of  $z_1$  in Figure 10

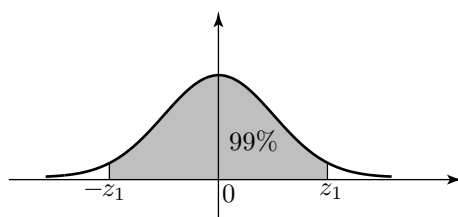


Figure 10 The shaded area is 99% of the total area.

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**

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% confidence interval.

**Your solution**

$(-2.58, 2.58)$  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  $Z$

- (i) which is exceeded on 5% of occasions
- (ii) which is exceeded on 99% of occasions.

**Your solution**

(i) 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  $Z_1 = 1.65$ .

(ii) 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  $z_1 = -2.33$ .

## 2. Confidence Intervals - Generalised Normal Distribution

We saw in sub-section 1 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  $x_2 = \mu + 1.96\sigma$ .

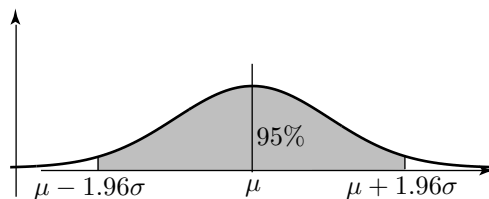


Figure 6

**Example** 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% confidence limits on the internal diameter of a single pipe?

**Solution**

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

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

i.e. 46.08 mm and 53.92 mm.

The **confidence interval** is (46.08, 53.92).

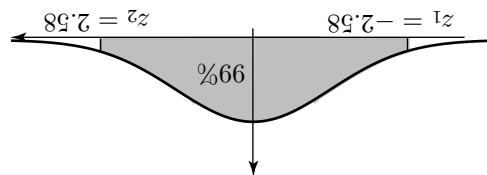


What is the 99% confidence interval for the lifetime of a bulb when the lifetimes of such bulbs are normally distributed with a mean of 2000 hours and standard deviation of 40 hours?

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**

Figure 7



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

**Your solution**

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

Next, find the values for  $x_1$  and  $x_2$  for the given problem.

**Your solution**

$$x_2 = 2000 + 2.58 \times 40 = 2103.2 \text{ hours}$$
$$x_1 = 2000 - 2.58 \times 40 = 1896.8 \text{ hours}$$

Finally, write down the 99% confidence interval for the lifetimes.

**Your solution**

(1896.8 hours, 2103.2 hours).