

STRAND: STATISTICS

Unit 5 *Confidence Intervals*

TEXT

Contents

Section

- | | |
|-----|-------------------------------|
| 5.1 | Sampling Methods |
| 5.2 | The Distribution of $\bar{X}$ |
| 5.3 | Confidence Intervals |

5 Confidence Intervals

5.1 Sampling Methods

If everyone was the same, there would be no need for sampling or statistics; you could find out everything you needed to know from one person (or one event or one result). Statistics involves the study of variability so that estimates and predictions can be made in complex situations where there is no certain answer. The quality and usefulness of these predictions depend entirely on the quality of the data upon which they are based.

You are probably familiar with methods of sampling; some of them will be revised here.

You will need the 'fish' sheet included as the final page in this text.

Random samples

Note that the fish are numbered from 1 to 57.

Use 2-figure random numbers from a random number table, calculator or computer to select a sample of 5 fish from the population. The method is described here. For 3-figure random numbers from a calculator, decide in advance whether you will use the first two digits, the last two digits, or the first and last.

Some of your 2-figure numbers will be larger than 57. These can be ignored without affecting the fairness of the selection process.

Here is an example showing a line of random numbers from a table:

25	82	33	06	74	18	34	09
	↓			↓			
	ignore			ignore			

The fish selected are numbered

25 33 6 18 and 34.

Note that you must use random numbers **consecutively** from the table after making a random start. You may **not** move about at will selecting numbers from different parts of the table.

Measure the lengths of these 5 fish.

Find the mean length of your sample and repeat the method for 2 further samples.

Collect together the sample means and display your results on a stem and leaf diagram.

Do you think that the mean of all the sample means is a good estimate for the mean length of the fish?

Definitions

To clarify your ideas, precise definitions are given below.

Population

A **population** is the set of all elements of interest for a particular study. Quantities such as the population mean μ are known as **population parameters**.

Sample

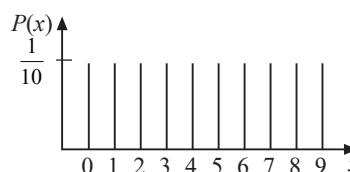
A **sample** is a subset of the population selected to represent the whole population. Quantities such as the sample mean $\bar{x}$ are known as **sample statistics** and are **estimates** of the corresponding population parameters.

Random sample

A **random sample** is a sample in which each member of the population has an equal chance of being selected. Random samples generate **unbiased** estimates of the population mean, whereas non-random samples may not be unbiased. Also, the variability within random samples can be mathematically predicted (as the next section will show).

In the next step we consider the significance of the sample size. The critical concept here is to recognise that as the sample size becomes larger, so the estimate of the sample mean should become closer to the true population mean.

The population of single-digit random numbers is theoretically infinite and consists of the numbers 0 to 9. These occur with equal probabilities and form a discrete uniform distribution, which you may have already met.



Clearly $p(x = 0) = \frac{1}{10}$

$$p(x = 1) = \frac{1}{10}$$

$$p(x = 9) = \frac{1}{10}$$



Worked Example 1

Use the formula

$$\mu = \sum x p(x) \quad , \quad \sigma^2 = \sum x^2 p(x) - \mu^2$$

to find the mean and variance of the population of single-digit random numbers.



Solution

We know that $p(0) = p(1) = \dots p(9) = \frac{1}{10}$

$$\begin{aligned}\mu &= \sum_{x=0}^9 x p(x) \\ &= 1 \cdot \frac{1}{10} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{1}{10} + \dots + 9 \cdot \frac{1}{10} \\ &= \frac{45}{10} \\ &= 4.5 \text{ (as expected)}\end{aligned}$$

$$\begin{aligned}\sigma^2 &= \sum_{x=0}^9 x^2 p(x) - (4.5)^2 \\ &= 1 \cdot \frac{1}{10} + 4 \cdot \frac{1}{10} + 9 \cdot \frac{1}{10} + 16 \cdot \frac{1}{10} + \dots + 81 \cdot \frac{1}{10} - (4.5)^2 \\ &= \frac{285}{10} - (4.5)^2 \\ &= 8.25\end{aligned}$$

giving the standard deviation of 2.87 to 2 decimal places.



Worked Example 2

(a) Here are a number of random samples of size 5 of single digit numbers.

1. { 2, 4, 0, 9, 6 }
2. { 3, 5, 6, 1, 1 }
3. { 5, 1, 4, 6, 4 }
4. { 9, 4, 2, 8, 4 }
5. { 5, 7, 4, 1, 9 }
6. { 9, 1, 1, 7, 9 }
7. { 9, 4, 2, 8, 4 }
8. { 9, 2, 4, 4, 9 }

For each sample, find its mean, $\bar{x}$, and calculate the mean and standard deviation of $\bar{x}$ values.

- (b) Repeat the process with 8 sets of random numbers of size 10.
- (c) Compare your values of the mean and standard deviation.



Solution

(a) The mean values of the 8 sets are as follows:

- | | | | |
|----|-----|----|-----|
| 1. | 4.2 | 5. | 5.2 |
| 2. | 3.2 | 6. | 5.4 |
| 3. | 4.0 | 7. | 5.4 |
| 4. | 5.4 | 8. | 5.6 |

So the set of values of $\bar{x}$ is:

{ 4.2, 3.2, 4.0, 5.4, 5.2, 5.4, 5.4, 5.6 }

and hence

mean of $\bar{x}$ values = 4.8

standard deviation of $\bar{x}$ values ≈ 0.82 (variance = 0.68)

(b) Here are 8 sets of random numbers of size 10.

1. { 1, 4, 4, 2, 2, 5, 8, 2, 8, 2 }
2. { 0, 3, 5, 5, 7, 4, 0, 0, 3, 6 }
3. { 7, 7, 1, 7, 3, 4, 8, 2, 5, 4 }
4. { 1, 2, 2, 5, 2, 5, 9, 5, 5, 7 }
5. { 7, 3, 6, 7, 3, 0, 4, 5, 3, 5 }
6. { 8, 5, 4, 7, 5, 1, 6, 0, 2, 1 }
7. { 7, 1, 0, 4, 3, 4, 4, 9, 4, 2 }
8. { 4, 3, 8, 4, 5, 2, 1, 4, 9, 7 }

The mean values are given below:

1. 3.8
2. 3.3
3. 4.8
4. 4.3
5. 4.3
6. 3.9
7. 3.8
8. 4.7

The set of values of $\bar{x}$ gives

$$\text{mean} = 4.1125$$

$$\text{standard deviation} \approx 0.473 \quad \left(\text{variance} = \frac{1431}{6400} \right)$$

- (c) Both means of the sample means give reasonable estimates of the population mean (which is 4.5), but the standard deviation of the means is considerably decreased for the larger sets ($n = 10$ rather than 5).

5.2 The Distribution of $\bar{X}$

Consider the idea of taking samples from a population. If it is a large population, it is possible (but perhaps not practical) to take a large number of samples, all of the same size from that population. For each sample, the mean $\bar{x}$ can be calculated. The value of $\bar{x}$ will vary from sample to sample and, as a result, is itself a random variable having its own distribution.

The value of the mean from any one sample is known as $\bar{x}$. If the distribution of all the possible values of the sample means is considered, this theoretical distribution is known as the **distribution of $\bar{X}$** .

So $\bar{X}$ itself is a random variable which takes different values for different random samples selected from a population.

In general, sample means are usually less variable though, than individual values. This is because, within a sample of size $n = 10$, say, large and small values in the sample tend to cancel each other out when $\bar{x}$ is calculated. In the example in the last section, even in a sample of ten random digits, $\bar{x}$ is unlikely to take a value greater than 7 or less than 2. Larger samples will generate values of $\bar{X}$ which are even more restricted in range (less variable).

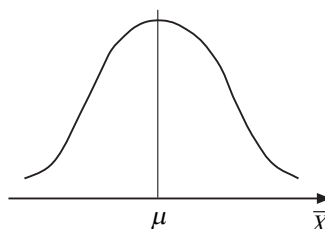
There is an inverse relationship between the size of the samples and the variance of $\bar{X}$. Also, the distribution of $\bar{X}$ tends to be a peaked distribution (with mean and mode at μ) which approaches a normal distribution for large samples.

The Central Limit Theorem

The Central Limit Theorem describes the distribution of $\bar{X}$ if **all** possible random samples (of a given size) are selected from a population. The following results hold.

1. The expected mean of all possible sample means is μ , the population mean; i.e.

$$E(\bar{X}) = \mu$$



2. The variance of the sample means is the population variance divided by the sample size

$$V(\bar{X}) = \frac{\sigma^2}{n}$$

As n increases, the variance of $\bar{X}$ decreases and as $n \rightarrow \infty$

$$V(\bar{X}) \rightarrow 0$$

3. If all possible values of $\bar{X}$ are calculated for a given sample size $n \geq 30$ a normal distribution is formed irrespective of the distribution of the original population:
i.e. for $n \geq 30$

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

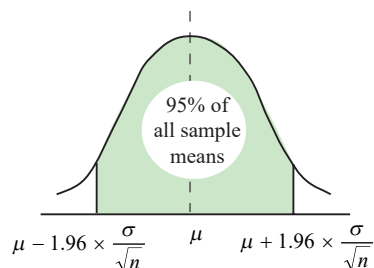
Note that these results are true only for **random samples**. For non-random samples you cannot make predictions in terms of mean, variance or distribution of $\bar{X}$.

The Central Limit Theorem can be used in practical situations to identify **unusual** samples, which are not typical of the population from which they have been selected.

For a given size of sample, the distribution of all possible sample means forms a normal distribution.

The mean of this distribution is μ (the overall population mean) and the variance is $\frac{\sigma^2}{n}$, giving the

standard deviation as $\frac{\sigma}{\sqrt{n}}$. This is often referred to as the **standard error**.



Distribution of all possible sample means for random samples of size n

Referring to normal distribution tables, 95% of any normal distribution lies between $z = -1.96$ and $z = +1.96$. So for a particular sample size n , 95% of all sample means should lie within 1.96 times the standard error each side of μ .

A sample mean outside this range may, in general, be :

- a genuine 'freak' result; after all, 5% of random samples do give means outside these limits;
- a result from a random sample selected from a different population;
- a sample selected from the population specified but not a random sample (e.g. a high proportion of children with above average IQs due to school selection procedures).



Worked Example 1

A survey of adults aged 16-64 living in Great Britain, by the Office of Population Censuses and Surveys (OPCS), found that adult females had a mean height of 160.9 cm with standard deviation of 6 cm.

A sample of fifty female students is found to have a mean height of 162 cm. Are their heights typical of the general population?



Solution

The population mean is given by

$$\mu = 160.9 \text{ cm}$$

Since the sample size is $n = 50$, the standard error is given by

$$\frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{50}} = 0.849$$

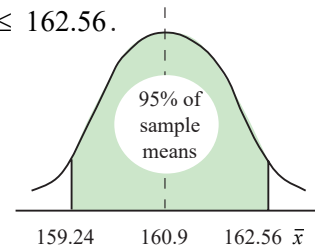
Thus the range of values for 95% of all sample means ($n = 50$) is

$$160.9 - (1.96 \times 0.849) \leq \bar{x} \leq 160.9 + (1.96 \times 0.849)$$

$$160.9 - 1.66 \leq \bar{x} \leq 160.9 + 1.66$$

So 95% of all $\bar{x}$ values should lie in the range $159.24 \leq \bar{x} \leq 162.56$.

You can see that the sample mean of 162 cm obtained from the fifty students is within the range of typical values for $\bar{x}$. So there is no evidence to suggest that this sample is not typical of the population in terms of height.



These ideas can be used to identify unusual sample means for large ($n \geq 30$) random samples selected from any population, or for small samples selected from a normal population provided the value of the population variance is known.



Exercises

1. IQ (Intelligence Quotient) scores are measured on a test which is constructed to give individual scores forming a normal distribution with a mean of 100 points and standard deviation of 15 points. A random sample of 10 students achieves a mean IQ score of 110 points. Is this sample typical of the general population?
2. A large group of female students is found to have a mean pulse rate (resting) of 75 beats per minute and standard deviation of 12 beats.

Later, a class of 30 students is found to have a mean pulse rate of 82 beats per minute. What are your conclusions?

3. Over the summer months, samples of adult specimens of freshwater shrimps are taken from a slow moving stream. Their lengths are measured and found to have a mean of 39 mm and standard deviation of 5.3 mm. During the winter, a small sample of 10 shrimps is found to have a mean length of 41 mm.
 - (a) Have the shrimps continued growing in the colder weather?
 - (b) What assumptions have you had to make in order to answer the question?
4. Afzal believes that the packets of crisps in the school tuck-shop are underweight. He takes a sample of ten packets of salt and vinegar crisps and finds their mean weight is 24.6 g. As the weight stated on the packets is 25 g, he writes to the manufacturer to complain. He receives the following reply:

Dear Sir,

Thank you for your letter of 5th July. We do share your concern over the weight of crisps in our packets of salt and vinegar crisps.

Over a period of time, we have found that the standard deviation of the weights of individual packets is a little below 1 g. For this reason we believe that your sample mean weight of 24.6 g comes well within the normal limits of acceptability.

Yours faithfully,

Do you agree with Afzal or do you agree with the manufacturers?

Would it help Afzal to take a larger sample?

5.3 Confidence Intervals

In many situations the value of μ , the population mean, may not be known for the variable being measured. Is it possible to estimate the value of μ in such cases?

The best estimate of μ is the value of $\bar{x}$ obtained from a random sample. As the estimate consists of a single value, $\bar{x}$, it is referred to as a **point estimate**. (Other less reliable point estimates can be obtained from the sample median or mid-range.) The sample mean $\bar{x}$ is an **unbiased estimator** for μ , but even so, the value of $\bar{x}$ obtained from any particular random sample is unlikely to give the exact value of μ . In fact, as an unbiased estimator, half the values of $\bar{x}$ will under estimate μ , while half will give over estimates.

In order to 'hedge our bets' a range of values may be given which should include the value of μ . This is called an interval estimate or **confidence interval**.

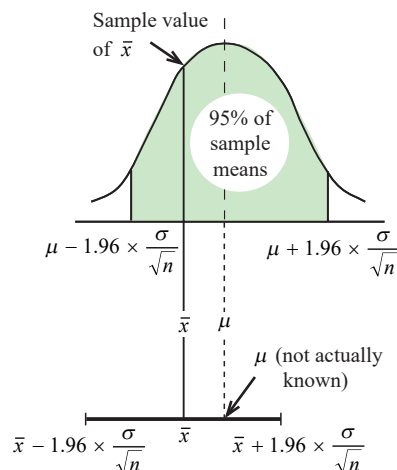
The ideas introduced in earlier sections can be used to construct such an interval estimate. There are two distinct cases to consider.

Population Variance Known

The distribution of all possible sample means, $\bar{X}$ forms a normal distribution, with a mean μ , at the true population mean. (The variance of this distribution is $\frac{\sigma^2}{n}$ and decreases for larger sized samples.

In reality, you are unlikely to know μ and all you have is one sample result $\bar{x}$. (Now $\bar{x}$ could lie anywhere in the distribution, as shown in the diagram opposite.)

Distribution of all possible values of $\bar{x}$ obtained from random samples of size n .



In order to estimate μ , a range of values can be taken around $\bar{x}$ which hopefully will include the true value of μ .

The **95% confidence interval** for μ is found by taking a range of 1.96 times the standard error $\frac{\sigma}{\sqrt{n}}$ either side of $\bar{x}$; that is

$$\left(\bar{x} - 1.96 \times \frac{\sigma}{\sqrt{n}} , \bar{x} + 1.96 \times \frac{\sigma}{\sqrt{n}} \right)$$

Providing $\bar{x}$ lies within the central 95% of the distribution of all possible sample means, the confidence interval will include μ . This will happen for 95 random samples out of 100. If the sample is a 'freak' sample and the sample value of $\bar{x}$ lies at one of the extreme ends of the distribution, the confidence interval will not include μ . This will happen for only 5 random samples (roughly) out of every 100.

If this margin of error is to be reduced a wider interval (which will include μ for, say, 99 samples out of 100) can be constructed. This is called a **99% confidence interval**.



Worked Example 1

What values of the standard variable z trap 99% of the distributions?



Solution

Tables of the normal distribution give $z = \pm 2.58$ to give 0.5% of the distribution in each tail.

Since the standard error is $\frac{\sigma}{\sqrt{n}}$, the range of values for which 99% of all sample means should lie is

$$\left(\bar{x} - 2.58 \times \frac{\sigma}{\sqrt{n}} , \bar{x} + 2.58 \times \frac{\sigma}{\sqrt{n}} \right)$$

This gives the required confidence interval.

Note that, for 99% confidence interval, 1 sample in every 100 the confidence interval will **not** include μ .

For other confidence intervals, e.g. 90%, 98%, you can look up the appropriate z value in the normal distribution tables.



Worked Example 2

A random sample of 100 men is taken and their mean height is found to be 180 cm. The population variance $\sigma^2 = 49 \text{ cm}^2$. Find the 95% confidence interval for μ , the mean height of the population.



Solution

$$\text{Lower limit} = \bar{x} - 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$\text{upper limit} = \bar{x} + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

when the standard error is given by

$$\frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{49}{100}} = 0.7$$

Hence the 95% confidence interval for μ is given by

$$= (180 \pm 1.96 \times 0.7) \text{ cm}$$

$$= (180 \pm 1.37) \text{ cm}$$

So μ should lie between 178.63 and 181.37.

It should be noted at this stage that a **parameter** is a measure of a population, e.g. the population mean μ , or population variances, etc.; whilst a **statistic** is a similar measure taken from a sample, e.g. the sample mean $\bar{x}$.

So a statistic is an **estimator** of a parameter. When investigating a practical problem, it is unlikely that information concerning all the items in a given population will be available. Knowledge will normally be limited to one sample, from which tentative conclusions may be drawn concerning the whole population from which the sample is taken. The larger the sample, the greater the confidence in the estimation.



Worked Example 3

The lifetimes of 10 light bulbs were observed (in hours) as

1052 1271 836 962 1019 1051 512 1027 1219 1040

Assuming that the standard deviation for light bulbs of this type is 80 hours,

- find the 95% confidence interval for the mean lifetime of this type of bulb;
- find the % of the confidence interval that has a total range of 80 hours;
- determine the sample size, n , needed to restrict the range of the 95% confidence interval to 50 hours.



Solution

- (a) Using the data, $\bar{x} = 998.9$, and taking $\sigma = 80$ (given in the question) the 95% confidence interval is given by

$$\left(\bar{x} - \frac{1.96\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \right)$$

$$\text{i.e. } \left(998.9 - \frac{1.96 \times 80}{\sqrt{10}}, 998.9 + \frac{1.96 \times 80}{\sqrt{10}} \right)$$

This gives

$$(998.9 - 49.6, 998.9 + 49.6)$$

$$\text{i.e. } (949.3, 1048.5)$$

- (b) For a confidence interval of range 80,

$$80 = \frac{2 \times z \times 80}{\sqrt{10}} \Rightarrow z \approx 1.58$$

For a normal distribution, with $z = 1.58$, the probability of being in the shaded region as shown in the diagram is, from tables,

$$0.94295$$

This gives the probability, p , of being in the tail as

$$p = 1 - 0.94295$$

$$= 0.05705$$

With two tails (see diagrams) the fraction covered is

$$2 \times 0.05705 = 0.1141$$

Hence the confidence interval is based on the fraction

$$1 - 0.1141 = 0.8859$$

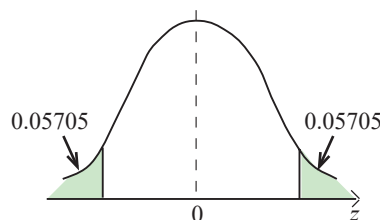
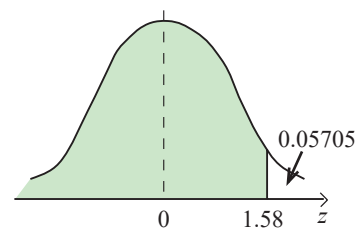
i.e. 89% confidence interval.

- (c) Now we have

$$50 = \frac{2 \times 1.96 \times 80}{\sqrt{n}}$$

$$\text{i.e. } \sqrt{n} = 6.272 \Rightarrow n = 39.34$$

So $n = 40$ will ensure that the range is restricted to 50 hours.



Population Variance Unknown

If the population variance is unknown and a large sample is taken, then the variance must be established from the sample itself. If s^2 is the **sample** variance, the best estimate for the **population** variance is given by

$$\begin{aligned}\hat{\sigma}^2 &= \frac{ns^2}{(n-1)} && \text{(this is not derived here)} \\ &= \frac{n}{(n-1)} \left(\frac{1}{n} \sum x^2 - \bar{x}^2 \right) \\ &= \frac{1}{(n-1)} (\sum x^2 - n\bar{x}^2)\end{aligned}$$

[This quantity is shown on calculators as σ_{n-1} or s_{n-1} .]

This result is used in the next example.



Worked Example 4

A user of a certain gauge of steel wire suspects that its breaking strength, in newtons (N), is different from that specified by the manufacturer. Consequently the user tests the breaking strength, x N, of each of a random sample of nine lengths of wire and obtains the following *ordered* results.

72.2 72.9 73.4 73.8 74.1 74.5 74.8 75.3 75.9

$$[\sum x = 666.9 \quad \sum x^2 = 49\,428.25]$$

Calculate the mean and the variance of the sample values.

Hence calculate a 95% confidence interval for the mean breaking strength.

Comment upon the manufacturer's claims that the breaking strength of the wire has a mean of 75.



Solution

For the sample,

$$\begin{aligned}\text{mean} &= \frac{\sum x_i}{n} \\ &= \frac{666.9}{9} \\ &= 74.1\end{aligned}$$

$$\begin{aligned}\text{variance} &= \frac{1}{n} \sum x^2 - \bar{x}^2 \\ &= \frac{49428.25}{9} - (74.1)^2 \\ &\approx 1.218\end{aligned}$$

The estimate of the population variance is given by

$$\begin{aligned}\hat{\sigma}^2 &= \frac{n}{(n-1)} s^2 \\ &= \frac{9}{8} \times 1.218 \\ &= 1.370 \\ \Rightarrow \sigma &\approx 1.170\end{aligned}$$

The 95% confidence interval is now given by

$$\begin{aligned}&\left(\bar{x} - 1.96 \times \frac{\hat{\sigma}}{\sqrt{n}}, \bar{x} + 1.96 \times \frac{\hat{\sigma}}{\sqrt{n}} \right) \\ \Rightarrow &\left(74.1 - 1.96 \times \frac{1.170}{\sqrt{9}}, 74.1 + 1.96 \times \frac{1.170}{\sqrt{9}} \right) \\ \Rightarrow &(73.34, 74.86)\end{aligned}$$

So the manufacturer's claim of a mean of 75 N is unlikely to be true since it is not included in the 95% confidence interval.



Exercises

- A sample of size 250 has mean 57.1 and standard deviation 11.8.
 - Find the standard error of the mean.
 - Give 95% confidence limits for the mean of the population.
- A company making cans for lemonade wishes to print 'Average contents x ml' on their cans, and to be 99% confident that the true mean volume is greater than x ml. The volume of lemonade in a can is known to have a standard deviation of 3.2 ml, and a random sample of 50 cans contained a mean volume of 503.6 ml. What volume of x should be stated?

- Butter is sold in packs marked as salted or unsalted and the masses of the packs of both types of butter are known to be normally distributed. The mean mass of the *salted* packs of butter is 225.38 g and the standard deviation for both packs is 8.45 g.

A sample of 12 of the *unsalted* packs of butter had masses, measured to the nearest gram, as follows.

219 226 217 224 223 216 221 228 215 229 225 229

- Find a 95% confidence interval for the mean mass of *unsalted* packs of butter.
- Using the given data for *salted* packs, that is the mean and standard deviation, calculate limits between which 90% of the masses of *salted* packs of butter will lie.

- (c) Estimate the size of sample which should be taken in order to be 95% sure that the sample mean of the masses of *salted* packs does not differ from the true mean by more than 3 g.
- (d) State, giving a reason, whether or not you would use the same sample size to be 95% sure of the same accuracy when sampling *unsalted* packs of butter.

4. The lengths of a sample of 100 rods produced by a machine are given below.

Length (cm)	5.60-5.62	5.62-5.64	5.64-5.66	5.66-5.68	5.68-5.70	5.70-5.72	5.72-5.74	5.74-5.76	5.76-5.78	5.78-5.80
Number of rods	1	3	5	5	8	20	24	16	12	6

- (a) Find the mean and standard deviation of the lengths in this sample.
- (b) (i) Estimate the standard error of the mean, and give 95% confidence limits for the true mean length, μ , of rods produced by the machine.
- (ii) Explain carefully the meaning of these confidence limits.

By taking a larger sample, the manufacturers wish to find 95% confidence limits for μ which differ by less than 0.004 cm.

- (c) Find the smallest sample size needed to do this.

For use with Section 5.1, Sampling Methods

