Contents 40

# Sampling Distributions and Estimation



## Learning outcomes

*You will learn about the distributions which are created when a population is sampled. For example, every sample will have a mean value; this gives rise to a distribution of mean values. We shall look at the behaviour of this distribution. We shall also look at the problem of estimating the true value of a population mean (for example) from a given sample.*

# Sampling Distributions

# 40.1

## Introduction

When you are dealing with large populations, for example populations created by the manufacturing processes, it is impossible, or very difficult indeed, to deal with the whole population and know the parameters of that population. Items such as car components, electronic components, aircraft components or ordinary everyday items such as light bulbs, cycle tyres and cutlery effectively form infinite populations. Hence we have to deal with samples taken from a population and estimate those population parameters that we need. This Workbook will show you how to calculate single number estimates of parameters - called point estimates - and interval estimates of parameters - called interval estimates or confidence intervals. In the latter case you will be able to calculate a range of values and state the confidence that the true value of the parameter you are estimating lies in the range you have found.

## Prerequisites

Before starting this Section you should ...

- understand and be able to calculate means and variances
- be familiar with the results and concepts met in the study of probability
- be familiar with the normal distribution

## Learning Outcomes

On completion you should be able to ...

- understand what is meant by the terms sample and sampling distribution
- explain the importance of sampling in the application of statistics
- explain the terms point estimate and the term interval estimate
- calculate point estimates of means and variances
- find interval estimates of population parameters for given levels of confidence

# 1. Sampling

## Why sample?

Considering samples from a distribution enables us to obtain information about a population where we cannot, for reasons of practicality, economy, or both, inspect the whole of the population. For example, it is impossible to check the complete output of some manufacturing processes. Items such as electric light bulbs, nuts, bolts, springs and light emitting diodes (LEDs) are produced in their millions and the sheer cost of checking every item as well as the time implications of such a checking process render it impossible. In addition, testing is sometimes destructive - one would not wish to destroy the whole production of a given component!

## Populations and samples

If we choose $n$ items from a population, we say that the size of the sample is $n$. If we take many samples, the means of these samples will themselves have a distribution which may be different from the population from which the samples were chosen. Much of the practical application of sampling theory is based on the relationship between the 'parent' population from which samples are drawn and the summary statistics (mean and variance) of the 'offspring' population of sample means. Not surprisingly, in the case of a normal 'parent' population, the distribution of the population and the distribution of the sample means are closely related. What is surprising is that even in the case of a non-normal parent population, the 'offspring' population of sample means is usually (but not always) normally distributed provided that the samples taken are large enough. In practice the term 'large' is usually taken to mean about 30 or more. The behaviour of the distribution of sample means is based on the following result from mathematical statistics.

## The central limit theorem

In what follows, we shall assume that the members of a sample are chosen at random from a population. This implies that the members of the sample are *independent*. We have already met the Central Limit Theorem. Here we will consider it in more detail and illustrate some of the properties resulting from it.

Much of the theory (and hence the practice) of sampling is based on the Central Limit Theorem. While we will not be looking at the proof of the theorem (it will be illustrated where practical) it is necessary that we understand what the theorem says and what it enables us to do. Essentially, the Central Limit Theorem says that if we take large samples of size $n$ with mean $\bar{X}$ from a population which has a mean $\mu$ and standard deviation $\sigma$ then the distribution of sample means $\bar{X}$ is normally distributed with mean $\mu$ and standard deviation $\dfrac{\sigma}{\sqrt{n}}$.

That is, the **sampling distribution of the mean** $\bar{X}$ follows the distribution

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

Strictly speaking we require $\sigma^2 < \infty$ , and it is important to note that no claim is made about the way in which the original distribution behaves, **and it need not be normal**. This is why the Central Limit Theorem is so fundamental to statistical practice. One implication is that a random variable which takes the form of a sum of many components which are random but not necessarily normal will itself be normal provided that the sum is not dominated by a small number of components. This explains why many biological variables, such as human heights, are normally distributed.

In the case where the original distribution is normal, the relationship between the original distribution $X \sim N(\mu, \sigma)$ and the distribution of sample means $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ is shown below.

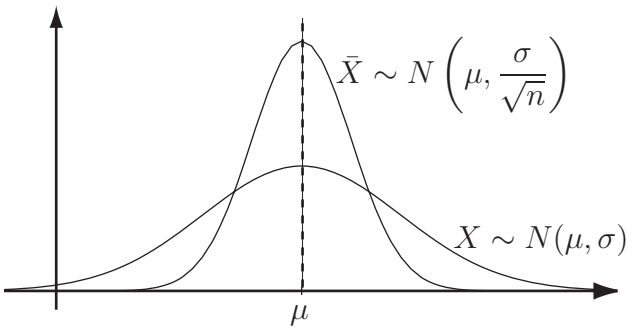


**Figure 1**

The distributions of $X$ and $\bar{X}$ have the same mean $\mu$ but $\bar{X}$ has the smaller standard deviation $\frac{\sigma}{\sqrt{n}}$

The theorem says that we must take *large* samples. If we take *small* samples, **the theorem only holds if the original population is normally distributed**.

## Standard error of the mean

You will meet this term often if you read statistical texts. It is the name given to the standard deviation of the population of sample means. The name stems from the fact that there is some uncertainty in the process of predicting the original population mean from the mean of a sample or samples.

**Key Point 1**

For a sample of $n$ independent observations from a population with variance $\sigma^2$, the **standard error of the mean** is $\sigma_n = \frac{\sigma}{\sqrt{n}}$.

Remember that this quantity is simply the standard deviation of the distribution of sample means.

## Finite populations

When we sample without replacement from a population which is not infinitely large, the observations are not independent. This means that we need to make an adjustment in the standard error of the mean. In this case the standard error of the sample mean is given by the related but more complicated formula

$$\sigma_{n,N} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

where $\sigma_{n,N}$ is the standard error of the sample mean, $N$ is the population size and $n$ is the sample size.

Note that, in cases where the size of the population $N$ is large in comparison to the sample size $n$, the quantity

$$\frac{N-n}{N-1} \approx 1$$

so that the standard error of the mean is approximately $\sigma/\sqrt{n}$.

### Illustration - a distribution of sample means

It is possible to illustrate some of the above results by setting up a small population of numbers and looking at the properties of small samples drawn from it. Notice that the setting up of a small population, say of size 5, and taking samples of size 2 enables us to deal with the totality of samples, there are $\binom{5}{2} = \frac{5!}{2!3!} = 10$ distinct samples possible, whereas if we take a population of 100 and draw samples of size 10, there are $\binom{100}{10} = \frac{100!}{10!90!} = 51,930,928,370,000$ possible distinct samples and from a practical point of view, we could not possibly list them all let alone work with them!

Suppose we take a population consisting of the five numbers 1, 2, 3, 4 and 5 and draw samples of size 2 to work with. The complete set of possible samples is:

$$(1,2),(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5)$$

For the parent population, since we know that the mean $\mu = 3$, then we can calculate the standard deviation by

$$\sigma = \sqrt{\frac{(1-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(5-3)^2}{5}} = \sqrt{\frac{10}{5}} = 1.4142$$

For the population of sample means,

$$1.5, 2, 2.5, 3, 2.5, 3, 3.5, 3.5, 4, 4.5$$

their mean and standard deviation are given by the calculations:

$$\frac{1.5+2+2.5+3+2.5+3+3.5+3.5+4+4.5}{10} = 3$$

and

$$\sqrt{\frac{(1.5-3)^2+(2-3)^2+\cdots+(4-3)^2+(4.5-3)^2}{10}} = \sqrt{\frac{7.5}{10}} = 0.8660$$

We can immediately conclude that the mean of the population of sample means is the same as the population mean $\mu$.

Using the results given above the value of $\sigma_{n,N}$ should be given by the formula

$$\sigma_{n,N} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

with $\sigma = 1.4142,\ N = 5$ and $n = 2$. Using these numbers gives:

$$\sigma_{2,5} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{1.4142}{\sqrt{2}}\sqrt{\frac{5-2}{5-1}} = \sqrt{\frac{3}{4}} = 0.8660 \text{ as predicted.}$$

Note that in this case the 'correction factor' $\sqrt{\frac{N-n}{N-1}} \approx 0.8660$ and is significant. If we take samples of size 10 from a population of 100, the factor becomes

$$\sqrt{\frac{N-n}{N-1}} \approx 0.9535$$

and for samples of size 10 taken from a population of 1000, the factor becomes

$$\sqrt{\frac{N-n}{N-1}} \approx 0.9955.$$

Thus as $\sqrt{\frac{N-n}{N-1}} \to 1$, its effect on the value of $\frac{\sigma}{\sqrt{n}}$ reduces to insignificance.


Two-centimetre number 10 woodscrews are manufactured in their millions but packed in boxes of 200 to be sold to the public or trade. If the length of the screws is known to be normally distributed with a mean of 2 cm and variance 0.05 cm$^2$, find the mean and standard deviation of the sample mean of 200 boxed screws. What is the probability that the sample mean length of the screws in a box of 200 is greater than 2.02 cm?

**Your solution**

**Answer**
Since the population is very large indeed, we are effectively sampling from an infinite population. The mean and standard deviation are given by

$$\mu = 2 \text{ cm} \quad \text{and} \quad \sigma_{200} = \frac{\sqrt{0.05}}{\sqrt{200}} = 0.016 \text{ cm}$$

Since the parent population is normally distributed the means of samples of 200 will be normally distributed as well.
Hence $\mathrm{P}(\text{sample mean length} > 2.02) = \mathrm{P}(z > \dfrac{2.02 - 2}{0.016}) = \mathrm{P}(z > 1.25) = 0.5 - 0.3944 = 0.1056$

# 2. Statistical estimation

When we are dealing with large populations (the production of items such as LEDs, light bulbs, piston rings etc.) it is extremely unlikely that we will be able to calculate population parameters such as the mean and variance directly from the full population.

We have to use processes which enable us to estimate these quantities. There are two basic methods used called point estimation and interval estimation. The essential difference is that point estimation gives single numbers which, in the sense defined below, are best estimates of population parameters, while interval estimates give a range of values together with a figure called the confidence that the true value of a parameter lies within the calculated range. Such ranges are usually called **confidence intervals**.

Statistically, the word 'estimate' implies a defined procedure for finding population parameters. In statistics, the word 'estimate' does not mean a guess, something which is rough-and-ready. What the word does mean is that an agreed precise process has been (or will be) used to find required values and that these values are 'best values' in some sense. Often this means that the procedure used, which is called the 'estimator', is:

(a) **consistent** in the sense that the difference between the true value and the estimate approaches zero as the sample size used to do the calculation increases;

(b) **unbiased** in the sense that the expected value of the estimator is equal to the true value;

(c) **efficient** in the sense that the variance of the estimator is small.

Expectation is covered in Workbooks 37 and 38. You should note that it is not always possible to find a 'best' estimator. You might have to decide (for example) between one which is

consistent, biased and efficient

and one which is

consistent, unbiased and inefficient

when what you really want is one which is

**consistent, unbiased** and **efficient.**

## Point estimation

We will look at the point estimation of the mean and variance of a population and use the following notation.

### Notation

| | Population | Sample | Estimator |
|---|---|---|---|
| **Size** | $N$ | $n$ | |
| **Mean** | $\mu$ or $\mathsf{E}(x)$ | $\bar{x}$ | $\hat{\mu}$ for $\mu$ |
| **Variance** | $\sigma^2$ or $\mathsf{V}(x)$ | $s^2$ | $\hat{\sigma}^2$ for $\sigma^2$ |

### Estimating the mean

This is straightforward.

$$\hat{\mu} = \bar{x}$$

is a sensible estimate since the difference between the population mean and the sample mean disappears with increasing sample size. We can show that this estimator is unbiased. Symbolically we have:

$$\hat{\mu} = \frac{x_1 + x_2 + \cdots x_n}{n}$$

so that

$$\begin{aligned} \mathsf{E}(\hat{\mu}) = &= \frac{\mathsf{E}(x_1) + \mathsf{E}(x_2) + \cdots + \mathsf{E}(x_n)}{n} \\ &= \frac{\mathsf{E}(X) + \mathsf{E}(X) + \cdots + \mathsf{E}(X)}{n} \\ &= \mathsf{E}(X) \\ &= \mu \end{aligned}$$

Note that the expected value of $x_1$ is $\mathsf{E}(X)$, i.e. $\mathsf{E}(x_1) = \mathsf{E}(X)$. Similarly for $x_1, x_2, \cdots, x_n$.

### Estimating the variance

This is a little more difficult. The true variance of the population is $\sigma^2 = \dfrac{\sum(x-\mu)^2}{N}$ which suggests the estimator, calculated from a sample, should be $\hat{\sigma}^2 = \dfrac{\sum(x-\mu)^2}{n}$.

However, we do not know the true value of $\mu$, but we do have the estimator $\hat{\mu} = \bar{x}$.

Replacing $\mu$ by the estimator $\hat{\mu} = \bar{x}$ gives

$$\hat{\sigma}^2 = \frac{\sum(x-\bar{x})^2}{n}$$

This can be written in the form

$$\hat{\sigma}^2 = \frac{\sum(x-\bar{x})^2}{n} = \frac{\sum x^2}{n} - (\bar{x})^2$$

Hence

$$\mathsf{E}(\hat{\sigma}^2) = \frac{\mathsf{E}(\sum x^2)}{n} - \mathsf{E}\{(\bar{X})^2\} = \mathsf{E}(X^2) - \mathsf{E}\{(\bar{X})^2\}$$

We already have the important result

$$\mathsf{E}(x) = \mathsf{E}(\bar{x}) \quad \text{and} \quad \mathsf{V}(\bar{x}) = \frac{\mathsf{V}(x)}{n}$$

Using the result $\mathsf{E}(x) = \mathsf{E}(\bar{x})$ gives us

$$\begin{aligned}
\mathsf{E}(\hat{\sigma}^2) &= \mathsf{E}(x^2) - \mathsf{E}\{(\bar{x})^2\} \\
&= \mathsf{E}(x^2) - \{\mathsf{E}(x)\}^2 - \mathsf{E}\{(\bar{x})^2\} + \{\mathsf{E}(\bar{x})\}^2 \\
&= \mathsf{E}(x^2) - \{\mathsf{E}(x)\}^2 - (\mathsf{E}\{(\bar{x})^2\} - \{\mathsf{E}(\bar{x})\}^2) \\
&= \mathsf{V}(x) - \mathsf{V}(\bar{x}) \\
&= \sigma^2 - \frac{\sigma^2}{n} \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}$$

This result is **biased**, for an unbiased estimator the result should be $\sigma^2$ not $\frac{n-1}{n}\sigma^2$.

Fortunately, the remedy is simple, we just multiply by the so-called Bessel's correction, namely $\frac{n}{n-1}$ and obtain the result

$$\hat{\sigma}^2 = \frac{n}{n-1}\frac{\sum(x-\bar{x})^2}{n} = \frac{\sum(x-\bar{x})^2}{n-1}$$

There are two points to note here. Firstly (and rather obviously) you should not take samples of size 1 since the variance cannot be estimated from such samples. Secondly, you should check the operation of any hand calculators (and spreadsheets!) that you use to find out exactly what you are calculating when you press the button for standard deviation. You might find that you are calculating either

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \quad \text{or} \quad \hat{\sigma}^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

It is just as well to know which, as the first formula assumes that you are calculating the variance of a population while the second assumes that you are estimating the variance of a population from a random sample of size $n$ taken from that population.

From now on we will assume that we divide by $n-1$ in the sample variance and we will simply write $s^2$ for $s^2_{n-1}$.

## Interval estimation

We will look at the process of finding an interval estimation of the mean and variance of a population and use the notation used above.

### Interval estimation for the mean

This interval is commonly called the Confidence Interval for the Mean.

Firstly, we know that while the sample mean $\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$ is a good estimator of the population mean $\mu$. We also know that the calculated mean $\bar{x}$ of a sample of size $n$ is unlikely to be exactly equal to $\mu$. We will now construct an interval around $\bar{x}$ in such a way that we can quantify the confidence that the interval actually contains the population mean $\mu$.

Secondly, we know that for sufficiently large samples taken from a large population, $\bar{x}$ follows a normal distribution with mean $\mu$ and standard deviation $\frac{\sigma}{\sqrt{n}}$.

Thirdly, looking at the following extract from the normal probability tables,

| $Z = \dfrac{X-\mu}{\sigma}$ | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.9 | .4713 | 4719 | 4726 | 4732 | 4738 | 4744 | 4750 | 4756 | 4762 | 4767 |

we can see that $2 \times 47.5\% = 95\%$ of the values in the standard normal distribution lie between $\pm 1.96$ standard deviation either side of the mean.

So before we see the data we may say that

$$\mathsf{P}\left(\mu - 1.96\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

After we see the data we say with $95\%$ confidence that

$$\mu - 1.96\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + 1.96\frac{\sigma}{\sqrt{n}}$$

which leads to

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}$$

This interval is called a 95% confidence interval for the mean $\mu$.

Note that while the $95\%$ level is very commonly used, there is nothing sacrosanct about this level. If we go through the same argument but demand that we need to be $99\%$ certain that $\mu$ lies within the confidence interval developed, we obtain the interval

$$\bar{x} - 2.58\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}$$

since an inspection of the standard normal tables reveals that $99\%$ of the values in a standard normal distribution lie within 2.58 standard deviations of the mean.

The above argument assumes that we know the population variance. In practice this is often not the case and we have to estimate the population variance from a sample. From the work we have seen above, we know that the best estimate of the population variance from a sample of size n is given by the formula

$$\hat{\sigma}^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

It follows that if we do not know the population variance, we must use the estimate $\hat{\sigma}$ in place of $\sigma$. Our 95% and 99% confidence intervals (for large samples) become

$$\bar{x} - 1.96\frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{\hat{\sigma}}{\sqrt{n}} \qquad \text{and} \qquad \bar{x} - 2.58\frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \bar{x} + 2.58\frac{\hat{\sigma}}{\sqrt{n}}$$

where

$$\hat{\sigma}^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

When we do not know the population variance, we need to estimate it. Hence we need to gauge the confidence we can have in the estimate.

In small samples, when we need to estimate the variance, the values 1.96 and 2.58 need to be replaced by values from the Student's $t$-distribution. See HELM 41.

## Example 1

After 1000 hours of use the weight loss, in gm, due to wear in certain rollers in machines, is normally distributed with mean $\mu$ and variance $\sigma^2$. Fifty independent observations are taken. (This may be regarded as a "large" sample.) If observation $i$ is $y_i$, then $\sum_{i=1}^{50} y_i = 497.2$ and $\sum_{i=1}^{50} y_i^2 = 5473.58$.

Estimate $\mu$ and $\sigma^2$ and give a 95% confidence interval for $\mu$.

**Solution**

We estimate $\mu$ using the sample mean: $\bar{y} = \dfrac{\sum y_i}{n} = \dfrac{497.2}{50} = 9.944$ gm

We estimate $\sigma^2$ using the sample variance:

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum (y_i - \bar{y})^2 = \frac{1}{n-1}\left\{\sum y_i^2 - \frac{1}{n}\left[\sum y_i\right]^2\right\} \\ &= \frac{1}{49}\left\{5473.58 - \frac{1}{50}497.2^2\right\} = 10.8046 \text{ gm}^2 \end{aligned}$$

The estimated standard error of the mean is $\sqrt{\dfrac{s^2}{n}} = \sqrt{\dfrac{10.8046}{50}} = 0.4649$ gm

The 95% confidence interval for $\mu$ is $\bar{y} \pm 1.96\sqrt{\dfrac{s^2}{n}}$. That is $9.479 < \mu < 10.409$

### Exercises

1. The voltages of sixty nominally 10 volt cells are measured. Assuming these to be independent observations from a normal distribution with mean $\mu$ and variance $\sigma^2$, estimate $\mu$ and $\sigma^2$. Regarding this as a "large"sample, find a 99% confidence interval for $\mu$. The data are:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 10.3 | 10.5 | 9.6 | 9.7 | 10.6 | 9.9 | 10.1 | 10.1 | 9.9 | 10.5 |
| 10.1 | 10.1 | 9.9 | 9.8 | 10.6 | 10.0 | 9.9 | 10.0 | 10.3 | 10.1 |
| 10.1 | 10.3 | 10.5 | 9.7 | 10.1 | 9.7 | 9.8 | 10.3 | 10.2 | 10.2 |
| 10.1 | 10.5 | 10.0 | 10.0 | 10.6 | 10.9 | 10.1 | 10.1 | 9.8 | 10.7 |
| 10.3 | 10.4 | 10.4 | 10.3 | 10.4 | 9.9 | 9.9 | 10.5 | 10.0 | 10.7 |
| 10.1 | 10.6 | 10.0 | 10.7 | 9.8 | 10.4 | 10.3 | 10.0 | 10.5 | 10.1 |

2. The natural logarithms of the times in minutes taken to complete a certain task are normally distributed with mean $\mu$ and variance $\sigma^2$. Seventy-five independent observations are taken. (This may be regarded as a "large" sample.) If the natural logarithm of the time for observation $i$ is $y_i$, then $\sum y_i = 147.75$ and $\sum y_i^2 = 292.8175$.

   Estimate $\mu$ and $\sigma^2$ and give a 95% confidence interval for $\mu$.

   Use your confidence interval to find a 95% confidence interval for the median time to complete the task.

**Answers**

1. $\sum y_i = 611.0$, $\sum y_i^2 = 6227.34$ and $n = 60$. We estimate $\mu$ using the sample mean:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{611.0}{60} = 10.1833 \text{ V}$$

We estimate $\sigma^2$ using the sample variance:

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum(y_i - \bar{y})^2 = \frac{1}{n-1}\left\{\sum y_i^2 - \frac{1}{n}\left[\sum y_i\right]^2\right\} \\ &= \frac{1}{59}\left\{6227.34 - \frac{1}{59}611.0^2\right\} = 0.090226 \end{aligned}$$

The estimated standard error of the mean is

$$\sqrt{\frac{s^2}{n}} = \sqrt{\frac{0.090226}{60}} = 0.03878 \text{ V}$$

The 99% confidence interval for $\mu$ is $\bar{y} \pm 2.58\sqrt{s^2/n}$. That is

$$10.08 < \mu < 10.28$$

2. We estimate $\mu$ using the sample mean:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{147.75}{75} = 1.97$$

We estimate $\sigma^2$ using the sample variance:

$$\begin{aligned} s^2 &= \frac{1}{n-1}\sum(y_i - \bar{y})^2 = \frac{1}{n-1}\left\{\sum y_i^2 - \frac{1}{n}\left[\sum y_i\right]^2\right\} \\ &= \frac{1}{74}\left\{292.8175 - \frac{1}{75}147.75^2\right\} = 0.02365 \end{aligned}$$

The estimated standard error of the mean is

$$\sqrt{\frac{s^2}{n}} = \sqrt{\frac{0.02365}{75}} = 0.01776$$

The 95% confidence interval for $\mu$ is $\bar{y} \pm 1.96\sqrt{s^2/n}$. That is

$$1.935 < \mu < 2.005$$

The 95% confidence interval for the median time, in minutes, to complete the task is

$$e^{1.935} < M < e^{2.005}$$

That is

$$6.93 < M < 7.42$$