Chapter 7

Example: The following example uses the dataset **faithful** which contains the duration of eruption (in minutes) and waiting times between eruptions for Old Faithful geyser in Yellowstone National Park, Wyoming, USA.

```
# look up info on the faithful dataset
?faithful

# load the faithful dataset
data(faithful)
```

We can use the **t.test** function to run a hypothesis (see chapter 8) and to create a confidence interval for an unknown parameter.

Suppose we want to create a 95% confidence interval for the true mean duration of eruptions for Old Faithful. By default, R will produce a 95% confidence interval unless otherwise specified. Since the population variance of eruption time for Old Faithful is unknown, the confidence interval will be created using the t-distribution.

```
t.test(faithful$eruptions)

##
##  One Sample t-test
##
## data:  faithful$eruptions
## t = 50.397, df = 271, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  3.351534 3.624032
## sample estimates:
## mean of x
##  3.487783
```

In the output, we see that the 95% confidence interval for the true mean eruption duration is $3.351534 < \mu < 3.624032$. If someone were to claim that the average eruption duration of Old Faithful is 4 minutes, we would reject this claim.

We could create a 99% confidence interval for the true mean eruption duration as well. This will be a wider confidence interval since we are requiring higher confidence.

```
t.test(faithful$eruptions,conf.level = 0.99)

##
##  One Sample t-test
##
## data:  faithful$eruptions
## t = 50.397, df = 271, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 99 percent confidence interval:
##  3.308257 3.667309
## sample estimates:
## mean of x
##  3.487783
```

The 99% confidence interval for the true mean eruption duration is $3.308257 < \mu < 3.667309$. We can also create one-sided confidence intervals by using the **alternative** argument with either **greater** or less. Below, a 90% upper one-sided confidence interval is created.

```
t.test(faithful$eruptions,conf.level = 0.90,alternative = "greater")

##
##  One Sample t-test
##
## data:  faithful$eruptions
## t = 50.397, df = 271, p-value < 2.2e-16
## alternative hypothesis: true mean is greater than 0
## 90 percent confidence interval:
##  3.398876      Inf
## sample estimates:
## mean of x
##  3.487783
```

The 90% upper one-sided confidence interval is $3.398876 < \mu$. We are 90% confident that the true mean eruption duration of Old Faithful is at least 3.339 minutes.