Chapter 10
Estimating Proportions with Confidence

For Chapter 10, this manual will show how to calculate multipliers for confidence intervals which use the normal distribution and also provide a demonstration for calculating confidence intervals for proportions.

Calculating the $z^*$ multiplier

The confidence interval for the population proportion uses the multiplier $z^*$. $z^*$ is the value such that the area between $-z^*$ and $+z^*$ under the standard normal curve is equal to the desired confidence interval. To find the multiplier for a 95% confidence interval, for example, we would want to find $z^*$ so that $P(-z^* \leq Z \leq z^*) = 0.95$ where $Z$ is a standard normal random variable. We will use R to calculate the multipliers shown in Table 10.1. These multipliers correspond to 90, 95, 98, and 99% confidence intervals. The inverse cumulative distribution function for the standard normal will be used to determine $z^*$ once the appropriate percentile is calculated. The percentile is $1 - (1 - CL) / 2$, where $CL$ denotes the confidence level such as 0.90, 0.95, etc. Thus for the 95% confidence interval, the percentile would be $p = 1 - (1 - 0.95) / 2 = 1 - 0.025 = 0.975$. (Note: In the just shown arithmetic, $(1 - CL) / 2 = 0.025$ is known as the level of significance and denoted by $\alpha$. This will be discussed in Chapter 11.) Below is the R code and the R output showing the $z^*$ multipliers for 90, 95, 98, and 99% confidence intervals respectively being 1.645, 1.960, 2.326, and 2.576.

```r
> CL <- c(0.90, 0.95, 0.98, 0.99) # confidence levels
> p <- 1-(1-CL)/2 # percentiles
> p
[1] 0.950 0.975 0.990 0.995
> qnorm(p, mean=0, sd=1) # z* multipliers
[1] 1.644854 1.959964 2.326348 2.575829
```

Example 10.3

Calculate a confidence interval for a population proportion

The R function `prop.test()` performs a hypothesis test on proportion data (see Chapter 11) and also returns a confidence interval for the population proportion. To calculate a confidence interval, you must provide `prop.test()` with the number of successes, the number of trials, and the desired confidence level. In Example 10.3 we wish to calculate the 90% confidence interval for a random survey of Americans. The number of trials is 1003 and the number of successes is 56% of 1003 which is $0.56 \times 1003 = 561$. Using `prop.test(561, 1003, conf.level=0.9)` we find the 90% confidence interval being 0.5329 through 0.5854. (Note: This will be close, but not exactly the same to the confidence interval shown in the text because of the use of a continuity correction in an attempt to achieve greater precision. Adding the option “correct=F”, `prop.test(561, 1003, conf.level=0.9, correct=F)`, would eliminate the continuity correction. See help(prop.test).)
Example 10.10

Calculating a confidence interval for a difference in two proportions

The R function \texttt{prop.test()} can be used with two sample proportions to calculate a confidence interval for the difference between the two population proportions. You only need to provide it a vector of successes and a vector of the respective number of trials.

In Example 10.10 we want to estimate the difference between the proportions with gender for seatbelt wearing habits. We would expect to have 0 within the confidence interval if there is no difference. 915 of the 1467 females wear their seatbelts when driving. 771 of the 1575 sampled males wear their seatbelts when driving. Below is the one line \texttt{prop.test()} command and the resulting R output which gives us a 95\% confidence interval of 0.0986 to 0.1698. (Using the correct=F option to ignore the continuity correction would have given the interval of 0.0921 to 0.1692.)

\begin{verbatim}
> prop.test( 561, 1003, conf.level=0.9)

1-sample proportions test with continuity correction
data: 561 out of 1003, null probability 0.5X-squared = 13.8824, df = 1, p-value = 0.0001946alternative hypothesis: true p is not equal to 0.590 percent confidence interval:
 0.5329114 0.5854074
sample estimates:
p0.559322
\end{verbatim}

\begin{verbatim}
> prop.test( c(915, 771), c(1467, 1575), conf.level=.95 )

2-sample test for equality of proportions with continuity correction
data: c(915, 771) out of c(1467, 1575)X-squared = 54.8244, df = 1, p-value = 1.318e-13alternative hypothesis: two.sided95 percent confidence interval:
0.0855335 0.16984279
sample estimates:
  prop 1    prop 2
0.6237219 0.4895238
\end{verbatim}