# Chapter 7 - Inference for Numerical Data

Shana Green

**Working backwards, Part II.** (5.24, p. 203) A 90% confidence interval for a population mean is (65, 77). The population distribution is approximately normal and the population standard deviation is unknown. This confidence interval is based on a simple random sample of 25 observations. Calculate the sample mean, the margin of error, and the sample standard deviation.

```
u<-77
l<-65
n<-25 # Sample Size
df<-n-1 # Degree of Freedom
CL<-c(0.95)
me<-((u-l)/2) # Margin of Error
mu<-((u+l)/2) # Mean
t.value <- round(qt(CL, df), 2)
s<-round((sqrt(n)*me)/t.value, 2)

mu
```

```
## [1] 71
```

```
t.value
```

```
## [1] 1.71
```

```
s
```

```
## [1] 17.54
```

---

**SAT scores.** (7.14, p. 261) SAT scores of students at an Ivy League college are distributed with a standard deviation of 250 points. Two statistics students, Raina and Luke, want to estimate the average SAT score of students at this college as part of a class project. They want their margin of error to be no more than 25 points.

(a) Raina wants to use a 90% confidence interval. How large a sample should she collect?

```
std_dev <- 250
me <- 25
z<-qnorm(0.95)
standard_error <- me/z
n <- round((std_dev/standard_error)^2, digit=0)
n
```

```
## [1] 271
```

**The sample size should be at least 271.**

(b) Luke wants to use a 99% confidence interval. Without calculating the actual sample size, determine whether his sample should be larger or smaller than Raina's, and explain your reasoning.

**Luke would need a larger sample size than Raina because he would be using a higher confidence interval.**

(c) Calculate the minimum required sample size for Luke.

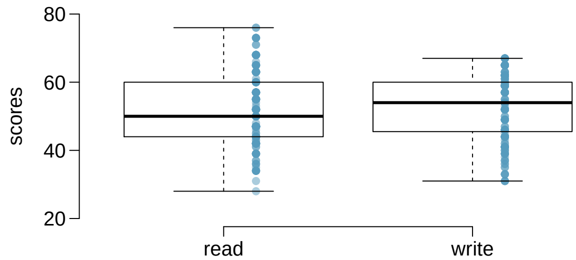
```
z<-qnorm(0.995)
standard_error <- me/z
n <- round((std_dev/standard_error)^2,digit=0)
n
```

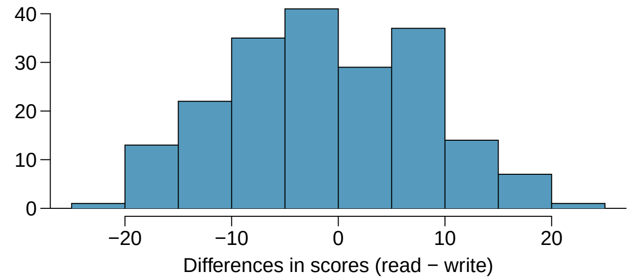
```
## [1] 663
```

**The sample size should be at least 663.**

---

**High School and Beyond, Part I.** (7.20, p. 266) The National Center of Education Statistics conducted a survey of high school seniors, collecting test data on reading, writing, and several other subjects. Here we examine a simple random sample of 200 students from this survey. Side-by-side box plots of reading and writing scores as well as a histogram of the differences in scores are shown below.




(a) Is there a clear difference in the average reading and writing scores?

```
mean(hsb2$read)
```

```
## [1] 52.23
```

```
mean(hsb2$write)
```

```
## [1] 52.775
```

**Based on both means, there is no clear difference between the two.**

(b) Are the reading and writing scores of each student independent of each other?

**I do not believe that the reading and writing scores are independent since each student has both scores.**

(c) Create hypotheses appropriate for the following research question: is there an evident difference in the average scores of students in the reading and writing exam?

$$H_0 : \mu_r - \mu_w = 0 \; H_A : \mu_r - \mu_w \neq 0$$

(d) Check the conditions required to complete this test.

```
n<-200
r<-mean(hsb2$read)
w<-mean(hsb2$write)
s<-8.887
df<-n-1 # Degree of Freedom
CL<-c(0.95)
me<-r-w # Margin of Error
mu<-((r+w)/2) # Mean
se<-round(s/sqrt(n),2)
t<-round(me/se,2)

p.value<-pt(t,df)


me
```

```
## [1] -0.545
```

```r
mu
```

```
## [1] 52.5025
```

```r
se
```

```
## [1] 0.63
```

```r
p.value
```

```
## [1] 0.1926743
```

**Since the p value is greater than .1, we do not have enough evidence to reject the null hypothesis.**

(e) The average observed difference in scores is $\widehat{x}_{read-write} = -0.545$, and the standard deviation of the differences is 8.887 points. Do these data provide convincing evidence of a difference between the average scores on the two exams?

```r
se
```

```
## [1] 0.63
```

```r
me
```

```
## [1] -0.545
```

```r
p <- round(2 * pt(me, df), 3)
p
```

```
## [1] 0.586
```

(f) What type of error might we have made? Explain what the error means in the context of the application.

**Type II error: Incorrectly reject the alternative hypothesis.**

**Based on (c) we may come across a type II error by rejecting $H_A$.**

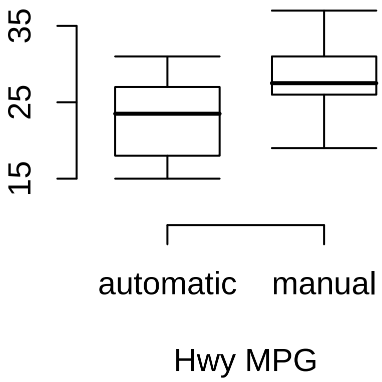
(g) Based on the results of this hypothesis test, would you expect a confidence interval for the average difference between the reading and writing scores to include 0? Explain your reasoning.

**I would expect a CI because the evidence was not strong enough to suggest that average reading and writing scores differ.**

---

**Fuel efficiency of manual and automatic cars, Part II.** (7.28, p. 276) The table provides summary statistics on highway fuel economy of cars manufactured in 2012. Use these statistics to calculate a 98% confidence interval for the difference between average highway mileage of manual and automatic cars, and interpret this interval in the context of the data.

| | Hwy MPG | |
|---|---|---|
| | Automatic | Manual |
| Mean | 22.92 | 27.88 |
| SD | 5.29 | 5.01 |
| n | 26 | 26 |




$$H_0 : \mu_a - \mu_m = 0 \; H_A : \mu_a - \mu_m \neq 0$$

```
n <- 26
a<-0.5


manual <- 27.88
sd_manual <- 5.01


auto <- 22.92
sd_auto<- 5.29


SE<- ((sd_auto^2 / n) + (sd_manual^2 / n)) ^ a

difference<-manual-auto

# T value of 98% confidence interval
T<-qt(0.01,df=25)
T<--T
```

```r
lower<-round(difference-T*SE,3)
higher<-round(difference+T*SE,3)

c(lower,higher)
```

```
## [1] 1.409 8.511
```

**The 98% confidence interval for the difference between average highway mileage of manual and automatic cars is between 1.409 and 8.511.**

**Email outreach efforts.** (7.34, p. 284) A medical research group is recruiting people to complete short surveys about their medical history. For example, one survey asks for information on a person's family history in regards to cancer. Another survey asks about what topics were discussed during the person's last visit to a hospital. So far, as people sign up, they complete an average of just 4 surveys, and the standard deviation of the number of surveys is about 2.2. The research group wants to try a new interface that they think will encourage new enrollees to complete more surveys, where they will randomize each enrollee to either get the new interface or the current interface. How many new enrollees do they need for each interface to detect an effect size of 0.5 surveys per enrollee, if the desired power level is 80%?

**If we assume $\alpha = 0.05$, then 304 enrollees are needed for each interface to detect an effect size of 0.5 surveys per enrollee, if the desired power level is 80%.**

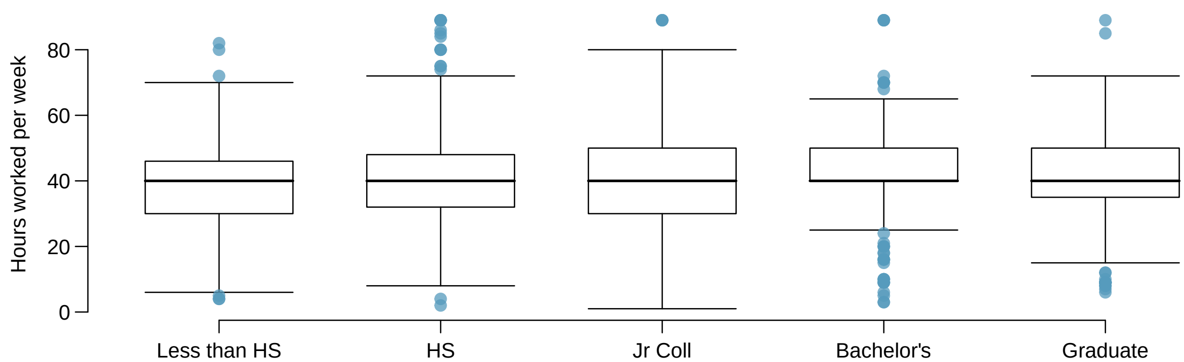
```
sd_1 <- 2.2
sd_2 <- 2.2
effect_size <- 0.5
standard_error <- 0.5/(0.84+1.96)
n <- round((sd_1^2+sd_2^2)/standard_error^2,digits=0)
n
```

```
## [1] 304
```

---

**Work hours and education.** The General Social Survey collects data on demographics, education, and work, among many other characteristics of US residents.47 Using ANOVA, we can consider educational attainment levels for all 1,172 respondents at once. Below are the distributions of hours worked by educational attainment and relevant summary statistics that will be helpful in carrying out this analysis.

| | *Educational attainment* | | | | | |
|---|---|---|---|---|---|---|
| | Less than HS | HS | Jr Coll | Bachelor's | Graduate | Total |
| Mean | 38.67 | 39.6 | 41.39 | 42.55 | 40.85 | 40.45 |
| SD | 15.81 | 14.97 | 18.1 | 13.62 | 15.51 | 15.17 |
| n | 121 | 546 | 97 | 253 | 155 | 1,172 |




(a) Write hypotheses for evaluating whether the average number of hours worked varies across the five groups.

$$H_0: \text{The average number of hours worked doesn't vary across the five groups}$$

$$H_A: \text{The average number of hours worked varies across the five groups.}$$

(b) Check conditions and describe any assumptions you must make to proceed with the test.

**The data from all 1,172 respondents are independent across groups.**

(c) Below is part of the output associated with this test. Fill in the empty cells.

| | Df | Sum Sq | Mean Sq | F-value | Pr(>F) |
|---|---|---|---|---|---|
| degree | 4 | 2006.16 | 501.54 | 2.188984 | 0.0682 |
| Residuals | 1167 | 267,382 | 229.12 | | |
| Total | 1171 | 269388.16 | | | |

(d) What is the conclusion of the test?

**Since the p $>$ 0.05, we can reject the null hypothesis and conclude that there is not a significant difference between the groups.**