Chapter 5: 5.6, 5.14, 5.20, 5.32, 5.48

Max Wagner

October 31, 2015

5.6

```
x <- 77 - 65
x2 <- x / 2
answer <- 65 + x2
```

The mean is the midpoint: 71

ME is: 6

DF are: 24

tdf: 1.71

forumla: ME = t(s/sqrt(n)) -> 6 = 1.71(s/sqrt(25)) -> s = 17.5439

5.14

a.

SD = 250

ME = 25

(1 - CI / 2) = 0.05

z = 1.65

```
((1.65 * 250) / 25) ^ 2
```

```
## [1] 272.25
```

b.

The sample size would need to be larger, in order to be more accurate. This can be seen with the z score difference.

c.

Z score because the sample size is > 30 and the SD is known.

5.20

a.

Reading score tends to be lower than writing score, but the difference is not massive, as some students seem to be better in reading than writing.

b.

It is unlikely that they are completely independent from each other as reading and writing are tied.

c.

Null: there is no difference in reading and writing scores, ALT: there is a difference in reading and writing scores

d.

The samples should be independent, random, greater than 30, and have no skew.

e.

The pvalue shown below is not less than the needed .05. This means that there is no difference between reading and writing scores.

```
diff <- -.545
sd <- 8.887
n <- 200

se <- sd / sqrt(n)
t <- diff / se
df <- 200 - 1
p <- pt(t, df = df);p
```

```
## [1] 0.1934182
```

f.

The source I would imagine is from there being a link between reading and writing scores, AKA they are not independent. In this case we would have incorrectly accepted the null hypothesis.

g.

It is very likely to include 0, is there is indeed no difference between reading and writing scores.

5.32

Null: difference = 0, Alt: difference != 0

```r
mean_a <- 16.12
mean_m <- 19.85
sd_a <- 3.58
sd_m <- 4.51
n <- 26
df <- n - 1
diff_means <- abs(mean_a - mean_m)
se <- sqrt((sd_a ^ 2 / n) + (sd_m ^ 2 / n))
t <- diff_means / se
p <- pt(t, df); 1 - p
```

```
## [1] 0.001441807
```

The p value is less than 0.05, so we reject the null hypothesis, and say that the means fuel milage differs.

5.48

a.

Null: there is no difference between all means, Alt: 1 or more means differ

b.

I assume that the samples are independent and random. The samples should also be normal. In this case the HS and Bachelor's boxplots show a significant number of outliers.

c.

ANOVA	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			