

DSSH 6301 - HW 05 Solutions

Problem 1

$$n = 100$$

$$\bar{x} = 104$$

$$\sigma = 22$$

```
n <- 100
x_bar <- 104
mu <- 100
sigma <- 22
```

Part a

Null Hypothesis: The average person is as smart as Sarah Palin, ie the population mean is 100. $H_0 : \mu = 100$.

Part b

Alternative Hypothesis: The average person is smarter than 100 (one-tailed) or the average person is different from 100 (two-tailed). $H_{a_1} : \mu > 100$ or $H_{a_2} : \mu \neq 100$.

Part c

$$t = \frac{\bar{x} - \mu}{se}$$

```
stan_err <- function(sigma, n)
  return(sigma / sqrt(n))

se <- stan_err(sigma, n)
t <- (x_bar - mu) / se
t
```

```
## [1] 1.818182
```

Part d

Here we perform a two-tailed test, although a one-tailed is also reasonable. The two-tailed test is more conservative and will also reject the null for extreme values in the opposite direction of what was expected.

Part e

We go with the standard $\alpha = 0.05$, although other levels are also reasonable.

The threshold values for a two-tailed test are:

```
thres <- qt(c(0.025, 1-0.025), n-1)
thres
```

```
## [1] -1.984217  1.984217
```

Part f

The threshold value for a one-tailed test is:

```
thres <- qt(1-0.05, n-1)
thres
```

```
## [1] 1.660391
```

Since our test statistic of 1.818 is greater than this value, we reject the null under the one-tailed test.

Part g

However, since our test statistic is not greater than 1.984, we fail to reject the null under the two-tailed test.

Part h

```
x_bar + qt(c(0.025, 0.975), n-1)*se
```

```
## [1] 99.63472 108.36528
```

The 95% CI $\in [99.63472, 108.36528]$.

Once again you can see that, since the 95% CI encompasses the null value (100), this is equivalent to the two-tailed test failing to reject the null.

Part i

The p-value for the two-tailed test is:

```
pt(t, n-1, lower.tail=F)*2
```

```
## [1] 0.07206032
```

Again, since this is greater than 0.05, we fail to reject the null. We cannot say for certain that the average person is any different from Sarah Palin in IQ.

Problem 2

```
n_men <- 50
x_barmen <- 1124
sigma_men <- 200

n_women <- 50
x_barwomen <- 1245
sigma_women <- 200
```

Part a

```
se1 <- stan_err(sigma_men, n_men)
se2 <- stan_err(sigma_women, n_women)

se_diff <- sqrt(se1^2 + se2^2)

diff <- x_barmen - x_barwomen
t <- diff / se_diff
t
```

```
## [1] -3.025
```

```
df <- n_men + n_women - 1

# The 95% CI:
diff + qt(c(0.025, 0.975), df)*se_diff
```

```
## [1] -200.36868 -41.63132
```

If the two sets of scores were significantly different, the CI should not include 0. Since it does not include 0, we can reject the null hypothesis that they are the same.

We can also see that our test statistic (t) is in the rejection region for a two-tailed test.

```
qt(c(0.025, 0.975), df)
```

```
## [1] -1.984217 1.984217
```

Part b

The answer here depends on what your research hypothesis was. If you believed that men were better than women in this sport and you performed a one-tailed test, your rejection region would be statistics greater than `qt(0.95,df)` and you would only reject the null if you got a high enough test statistic. But in this case, you would be making a serious error, since as it happens, women are in fact significantly better than men, not the reverse, and the test statistic is large but negative. A two-tailed test would catch this and correctly reject the null that the two groups are equivalent. This again illustrates the value of the two-tailed test.

Problem 3

H_a : Drinking the night before a test helps performance on the test.

H_0 : Drinking the night before a test does not help test performance.

```
n_treatment <- 50
mean_treatment <- 78
sd_treatment <- 10

n_control <- 50
mean_control <- 75
sd_control <- 5
```

Part a

```
se1 <- stan_err(sd_treatment, n_treatment)
se2 <- stan_err(sd_control, n_control)

se_diff <- sqrt(se1^2 + se2^2)
se_diff
```

```
## [1] 1.581139
```

```
t <- (mean_treatment - mean_control) / se_diff
t
```

```
## [1] 1.897367
```

```
df <- se_diff^2 / (se1^4/(n_treatment-1) + se2^4/(n_control-1))
df
```

```
## [1] 28.82353
```

```
p_val <- 2*pt(t, df, lower.tail=F)
p_val
```

```
## [1] 0.06783864
```

The treatment group did not do better than the control group. We fail to reject the null at the $\alpha = 0.05$ level.

Problem 4

Part a

```
set.seed(1)
n <- 10000
population_std_norm <- rnorm(n)
t.test(population_std_norm)
```

```
##
## One Sample t-test
##
## data: population_std_norm
## t = -0.6457, df = 9999, p-value = 0.5185
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.02638126 0.01330718
## sample estimates:
## mean of x
## -0.006537039
```

This test compares the population to a mean of 0 (the default μ for the `t.test` function). We find a p-value of 0.519; a CI of [-0.02638126, 0.01330718]; and a test statistic of -0.6457 (which is neither less than -1.96 nor greater than 1.96). All of these mean that we fail to reject the null, and find no evidence that the mean is different from 0.

Part b

```
t.test(population_std_norm, rnorm(n))
```

```
##
## Welch Two Sample t-test
##
## data: population_std_norm and rnorm(n)
## t = -0.1657, df = 19988.72, p-value = 0.8684
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.03011178 0.02541783
## sample estimates:
## mean of x mean of y
## -0.006537039 -0.004190066
```

This test compares the means of two normal populations of mean 0 and sd 1 (the default for `rnorm`). Whether via the p-value (which is greater than 0.05); the CI (which includes 0); or the t statistic (which is not in the rejection region), we again fail to reject the null that these two samples are not different.

Part c

We construct a second variable that is 0.1 greater than the first, plus some random noise (the `3*rnorm(n)` part):

```
population_std_norm_2 <- population_std_norm + 3*rnorm(n) + .1
t.test(population_std_norm - population_std_norm_2)
```

```
##
## One Sample t-test
##
## data: population_std_norm - population_std_norm_2
## t = -4.0575, df = 9999, p-value = 4.997e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -0.18192293 -0.06340499
## sample estimates:
## mean of x
## -0.122664
```

With a p-value < 0.05 , we reject the null that these two samples are the same. The “mean of x” (where x is the difference between the two groups) is -0.122, pretty close to the truth of -0.1.

Part d

```
xbar <- mean(population_std_norm)
sd <- sd(population_std_norm)

se <- stan_err(sd, n)

t <- xbar / se

p_val <- pt(t, n-1)*2
p_val
```

```
## [1] 0.5184722
```

```
# 95% CI
xbar + qt(c(0.025, 0.975), n-1)*se
```

```
## [1] -0.02638126 0.01330718
```