Practice Session 8

Introduction to Type 1 and Type 2 Errors

In hypothesis testing, we either reject or fail to reject the null hypothesis. However, we can sometimes make the wrong decision. A Type 1 error results when we reject the null hypothesis, but the null hypothesis is actually true. A Type 2 error occurs when we fail to reject the null hypothesis, but the null hypothesis is actually false. One important fact to note is that the significance level that we use for p-values (typically 0.05) is the likelihood that we make a Type 1 error.

Type 1, Type 2, or Neither?

For each of the following statements, decide whether a Type 1 or a Type 2 Error was made. If no error was made, indicate that no error was committed.

a.) Type 1 Error

b.) No error

c.)

- No error
- Type 1 Error
- Type 2 Error
- No error

Single Sample T Test

According to a local statistics instructor, the average score on his first exam of the semester (`exam1`) is historically around 80%. Run a one sample T test to see if this year's students performed significantly better than the historic average. Use the formulas learned in class to perform the test, and do the calculations using R. The data is available in the `StatsGrades` data set from the `Lock5Data` library.

- $H_0 : \mu = 80$
- $H_A : \mu > 80$

```
library(Lock5Data)
exam1 = StatGrades$Exam1 # get vector of exam scores
n1 = length(StatGrades$Exam1) # get sample size (n = 50)

mean_exam1 = mean(StatGrades$Exam1) # compute sample mean
se_exam1 = sd(StatGrades$Exam1) / sqrt(n1) # compute standard error

T_stat_exam1 = (mean_exam1 - 80) / se_exam1
T_stat_exam1
```

```
[1] 0.7145975
```

```
# df = 50 - 1 = 49
pt(T_stat_exam1, df = n1 - 1, lower.tail = F)
```

```
[1] 0.2391243
```

Since our p-value is greater than 0.05, we will fail to reject the null hypothesis. We therefore do not have evidence that this year's exam scores are significantly higher than 80.

Two Sample T Test

The same local statistics instructor believes that students perform the same on the first exam (`exam1`) and the second exam (`exam2`). Run a two sample T test to see if this year's students performed differently on the two exams. Use the formulas learned in class to perform the test, and do the calculations using R. The data is available in the `StatsGrades` data set from the `Lock5Data` library.

```r
exam2 = StatGrades$Exam2 # get vector of exam scores
n2 = length(StatGrades$Exam2) # get sample size (n = 50)

mean_exam2 = mean(StatGrades$Exam2) # compute sample mean
```

Paired T Test

Rerun the two sample T test above, but this time do a paired T test. Note any differences in the results. What do you conclude about student performance on the two exams?

Let D be the true difference between the exam 1 and exam 2 scores. We will perform the following two-sided test:

- $H_0 : D = 0$
- $H_A : D \neq 0$

```r
diff_exams = exam2 - exam1
n_diff = length(diff_exams)

mean_diff = mean(diff_exams)
se_diff = sd(diff_exams) / sqrt(n_diff)

T_stat_diff = mean_diff / se_diff
```

```r
2 * pt(T_stat_diff, df = n_diff - 1, lower.tail = F)
```

```
[1] 0.01411695
```

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have evidence that this year's students performed differently on the two exams.