

Practice Session 9

Part 1

Introduction to Type I and Type II Errors

In hypothesis testing, we need to decide whether we are going to reject the null hypothesis, or fail to reject the null hypothesis. We sometimes, however, make the wrong decision. A **Type I** error occurs when we reject the null hypothesis, but the null hypothesis is actually true. A **Type II** error occurs when we fail to reject the null hypothesis, but the null hypothesis is actually false. The probability that we make a Type I error is also known as the significance level. Typically, we set our significance level to be 0.05 in hypothesis testing.

Question 1: Type I, Type II, or Neither?

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

a) Shiba ran a hypothesis test with the following hypotheses:

- $H_0 : \mu = 50$
- $H_A : \mu > 50$

She rejected the null hypothesis; the true mean is actually 50.

b) Joe wanted to test if the mean bag weight of Lays Potato chips is less than 5 oz.

- $H_0 : \mu = 5$
- $H_A : \mu < 5$

He failed to reject the null; the true mean is actually 5 oz.

c) A test for an infectious disease gives a positive result if a patient is believed to have the disease, and a negative result otherwise. However, the test can sometimes be wrong. Label the following statements as a Type I error, Type II error, or neither. Assume the null hypothesis is that the patient doesn't have the disease.

- A positive test result; the patient actually has the disease
- A positive test result; the patient doesn't actually have the disease
- A negative test result; the patient actually has the disease
- A negative test result; the patient doesn't actually have the disease

In the context of this question, which error (Type I or Type II) would be more problematic? If you could make one of these errors less likely to occur, which one would it be, and why?

Introduction to the One Sample t-test

When our population standard deviation σ is unknown, we can do inference on population means μ using a t-distribution. Let n be the sample size, \bar{x} be the sample mean, and s be the sample standard deviation. The formula for the estimated standard error (SE) of the sample mean, and the t statistic are below.

$$SE = \frac{s}{\sqrt{n}}$$

$$t \text{ Statistic} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t(df = n - 1)$$

Where $df = n - 1$ is the degrees of freedom of the test. This formula is for the one sample t-test.

Question 2: One Sample t-test

According to a local statistics instructor, the average score on his first exam of the semester (**Exam1**) is historically around 80%. The data is available in the **StatsGrades** data set from the **Lock5Data** library.

a) First visualize the data using a boxplot and a histogram. Do you see any outliers?

```
# Your code here
```

Run a one sample t-test to see if this year's students performed significantly better than the historic average.

- b) State the null and alternative hypotheses using symbols
- c) Calculate the t-statistic

```
# Your code here
```

- d) Calculate the p-value of the test, and state your conclusion.

```
# Your code here
```

- e) Use the `dt()` function and `plot()` to visualize the t-distribution. The degrees of freedom will be the same as for the test. Add a vertical red line to the plot at the observed sample mean.

```
# Your code here
```

Introduction to the Two Sample t-test

Similar to the one sample case, we can compute a standard error and t-statistic to perform a two sample test. Suppose we have two samples, sample 1 and sample 2. To conduct the test, we need the following information:

- n_1 = sample size for sample 1
- n_2 = sample size for sample 2
- \bar{x}_1 = sample mean for sample 1
- \bar{x}_2 = sample mean for sample 2
- s_1 = sample standard deviation for sample 1
- s_2 = sample standard deviation for sample 2

$$SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$t \text{ Statistic} = \frac{\bar{x}_2 - \bar{x}_1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t(df = \min\{n_1 - 1, n_2 - 1\})$$

Where $df = \min\{n_1 - 1, n_2 - 1\}$ is the degrees of freedom for the test.

Question 3: 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). Again, the data is available in the `StatsGrades` data set from the `Lock5Data` library.

- a) First visualize both sets of exam scores using side-by-side boxplots. Comment on what you notice.

```
# Your code here
```

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.

- b) State the null and alternative hypotheses using symbols
c) Calculate the t-statistic

```
# Your code here
```

- d) Calculate the p-value of the test, and state your conclusion.

```
# Your code here
```

- e) Suppose you were more interested in how students do on the first exam compared to the `Final`. Run a hypothesis test to see if student scores improved on the final compared to the first exam.

```
# Your code here
```

Question 4: Paired t-test

Rerun the two sample t-test that compares Exam 1 to Exam 2, 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? Why does it make sense to do a paired t-test in this case?

- a) State the null and alternative hypotheses using symbols.
b) Calculate the difference vector by subtracting the Exam 1 scores from the Exam 2 scores.

```
# Your code here
```

- c) Calculate the mean of the difference vector, save this as `D_bar`.

```
# Your code here
```

d) Calculate the t-statistic

e) Calculate the p-value of the test, and state your conclusion.

```
# Your code here
```

Part 2 Permutations tests for one mean and difference between two means:

Test for the difference between two means with permutations:

The same local statistics instructor believes that students perform the same on the first exam (**Exam1**) and the second exam (**Exam2**). Again, the data is available in the **StatsGrades** data set from the **Lock5Data** library.

1.) Create a null distribution and create its histogram.

```
# Your code here
```

2.) Calculate the p-value

```
# Your code here
```

Test for one mean:

(Let go back to the question of testing one mean).

According to a local statistics instructor, the average score on his first exam of the semester (**Exam1**) is historically around 80%. The data is available in the **StatsGrades** data set from the **Lock5Data** library. Create a bootstrap distribution with a mean of the hypothesised value

1.) Calculate the mean of the data.

```
# Your code here
```

2.) Subtract this mean from the data.

```
# Your code here
```

3.) Add the hypothesized value to the data.

```
# Your code here
```

4.) Now create a bootstrap distribution with the last modified data you have created in the previous step.

```
# Your code here
```

5.) Calculate the p-value

```
# Your code here
```