t-distribution
Cohen's d
Practice Problem

3 — t-Tests

3.1 t-distribution

The t-Test is best to use when we do not know the population standard deviation. Instead we use the sample standard deviation.

Definition 3.1 — t-stat. The t-Test statistic can be computed very similarly to the z-stat, to compute the t-stat we compute:

$$t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

We also have to compute the degrees of freedom (df) for the sample: $df = n - 1$

Like the Z-stat we can use a table to get the proportion below or between a specific value. T-tests are also great for testing two sample means (i.e. paired t-tests), we modify the formula to become:

$$\frac{(x_2 - x_1) - (\mu_2 - \mu_1)}{\frac{\sqrt{(s_1^2 + s_2^2}}{n}}$$

■ **Example 3.1** ■

3.1.1 Cohen's d

Definition 3.2 — Cohen's d. Cohen's d measures the effect size of the strength of a phenomenon. Cohen's d gives us the distance between means in standardized units. Cohen's d is computed by:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

where $s = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$

3.2 Practice Problem

Problem 3.1 Pizza company A wants to know if they deliver Pizza faster than Company B. The following table outlines there delivery times:

Company A	Company B
20.4	20.2
24.2	16.9
15.4	18.5
21.4	17.3
20.2	20.5
18.5	
21.5	

Table 3.1: Pizza Companies Delivery Times

Problem 3.2 Use Cohen's d to measure the effect size between the two times.

4 — t-Tests continued

4.1 Standard Error

Definition 4.1 — Standard Error. The Standard error is the standard deviation of the sample means over all possible samples (of a given size) drawn from the population. It can be computed by:

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

The standard error for two samples can be computed with:

$$SE = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Definition 4.2 — Pooled Variance. Pooled variance is a method for estimating variance given several different samples taken in different circumstances where the mean may vary between samples but the true variance is assumed to remain the same. The pooled variance is computed by using:

$$S_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

R We can use pooled variance to compute standard error that is:

$$SE = \sqrt{\frac{S_p^2}{n_1} + \frac{S_P^2}{n_2}}$$