MATH 1342

Section 8.3

Comparing Two Means

Two – sample *t* – tests compare the responses to two treatments or characteristics of two populations.

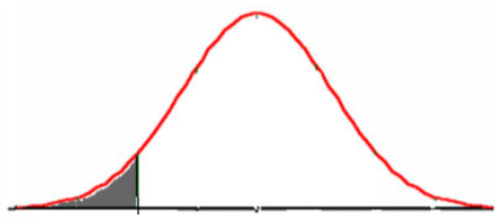
There is a separate sample from each treatment or population. These tests are quite different than the matched pairs *t* – test discussed in section 8.1.

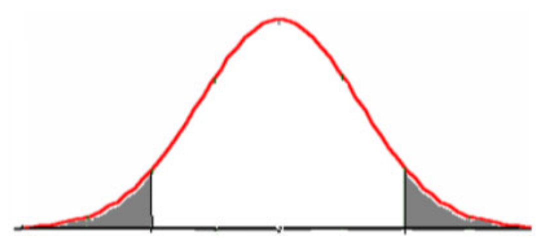
How can we tell the difference between dependent and independent populations/samples?

In the wording of the questions, are the two samples connected in some way? (Is there a before/after relationship?) Then it would be a matched pairs t-test.
If the two samples are unrelated and unconnected, then this is a two-sample t-test.

The null and alternate hypotheses would be:

$$H_0: \mu_1 = \mu_2 \quad \text{or} \quad H_0: \mu_1 = \mu_2 \quad \text{or} \quad H_0: \mu_1 = \mu_2$$
$$H_a: \mu_1 > \mu_2 \qquad\quad H_a: \mu_1 < \mu_2 \qquad\quad H_a: \mu_1 \neq \mu_2$$





And the assumptions for a two-sample *t* – test are:

1. We have two independent SRSs, from two distinct populations and we measure the same variable for both samples.
2. Both populations are normally distributed with unknown means and standard deviations. (Or if each given sample size is greater than or equal to 30.)

Two-Sample t-test statistic:

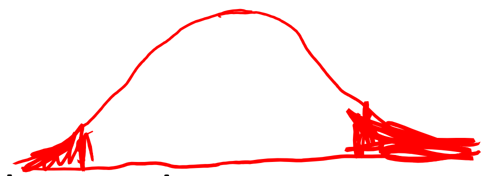
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \overbrace{(\mu_1 - \mu_2)}^{=0}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The degrees of freedom is equal to the smaller of $n_1 - 1$ and $n_2 - 1$.

Example:

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$



```
qt(0.02/2,64)
[1] -2.386037
```

The president of an all-female school stated in an interview that she was sure that the students at her school studied more, on average, than the students at a neighboring all-male school. The president of the all-male school responded that he thought the mean study time for each student body was undoubtedly about the same and suggested that a study be undertaken to clear up the controversy.

Accordingly, independent samples were taken at the two schools with the following results:

School	Sample Size	Mean Study Time (hrs)	Standard deviation (hrs)
All Female (Group 1)	65	18.56	4.35
All Male (Group 2)	75	17.95	4.87

$\alpha = .02$

Determine, at the 2% level of significance, if there is a significant difference between the mean studying times of the students in the two schools based on these samples.

Rejection Region: $t < -2.386$ or $t > 2.386$

Workspace:

School	Sample Size	Mean Study Time (hrs)	Standard deviation (hrs)
All Female	65	18.56	4.35
All Male	75	17.95	4.87

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

0

```
> (18.56-17.95)/sqrt(4.35^2/65+4.87^2/75)
[1] 0.782733
```

This does not fall within our rejection region

```
2*pt(-0.782733,64)
1] 0.4366706
```

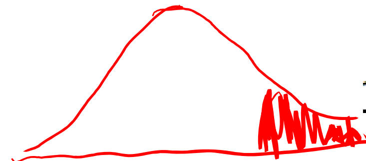
p-value: 43.7% (alpha : 2%)
p > alpha
FRHo

This means that that the average study at the two schools is roughly the same

Example:

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$



```
qt(1-0.05,39)
1] 1.684875
```

Rej Reg:
t > 1.685

A study was conducted to determine whether remediation in basic mathematics enabled students to be more successful in an elementary statistics course. Samples of final exam scores were taken from students who had remediation and from students who did not. Here are the results of the study:

	Remedial (Grp 1)	Non-remedial (Grp 2)
Sample size	100	40
Mean Exam Grade	83.0	76.5
Std Dev for Exam	2.76	4.11

$p < \alpha$

RH_0

Test, at the 5% (α) level, whether the remediation helped the students to be more successful.

$t =$

```
(83.0-76.5)/sqrt(2.76^2/100+4.11^2/40)
1] 9.206406
```

```
1-pt(9.206406,39)
1] 1.260747e-11
```

$p \approx 0$

Popper 29:

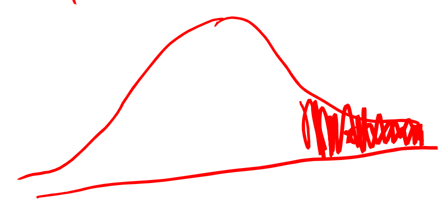
You conduct a study comparing the average cost of a 3-bedroom home in New York (Population 1) versus Texas (Population 2). You select a simple random sample of 3-bedroom homes purchased within the last year from each population. The results are described below.

Test, at a 5% level of significance, if the cost of a home in New York is more than the cost of a home in Texas.

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

	New York	Texas
Sample Size	75	135
Mean Cost (in thousands)	450	210
Standard Deviation	75	100



Popper 29, Continued:

	New York	Texas
Sample Size	75	135
Mean Cost (in thousands)	450	210
Standard Deviation	75	100

1. What is the Null Hypothesis?
   a. $\mu_1 = \mu_2$ b. $\mu_1 < \mu_2$ c. $\mu_1 > \mu_2$ d. $\mu_1 \neq \mu_2$
2. What is the Alternate Hypothesis?
   a. $\mu_1 = \mu_2$ b. $\mu_1 < \mu_2$ c. $\mu_1 > \mu_2$ d. $\mu_1 \neq \mu_2$
3. What is the critical value?
   a. -1.99 b. 1.35 c. -1.67 d. 1.67
4. What is the test statistic?
   a. 1.965 b. 19.66 c. 0 d. 1.32
5. What is the p-value?
   a. 0 b. 1.56 c. 5.67 d. 15.32
6. What is your conclusion?
   a. New York is higher cost (overwhelming) b. New York is higher cost (strong)
   c. New York is higher cost (weak) d. New York is not higher cost

```
qt(1-0.05,74)
[1] 1.665707
```

```
(450-210)/sqrt(75^2/75+100^2/135)
[1] 19.65668
```

```
1-pt(19.65668,74)
[1] 0
```

$p < \alpha$ and $p < 1\%$