

Module 2E:

Inference for a Normal Population

Testing for equal variances; Welch's Approximate t test

Hypothesis testing for means using t tests Agenda

1. **Why** do we use Student t-tests instead of Z scores?

2. **What are the three types of t-tests**

- **One sample t tests**

- ☐ Assumptions

- ☐ When assumptions not met, use median and rank → **Signed test**

- **Paired t test**

- ☐ Assumptions

- **Two sample t test**

- ☐ Assumptions

- ☐ When variances aren't equal → **Welch's approximate t test**

- ☐ Other assumptions not met: median and rank → **Mann Whitney U test**

We won't have time to cover everything in detail – nor every example I give - so here is another reference that outlines the different t tests:

Assumptions of parametric tests:

- Random Samples
- Populations are normally distributed
- for two sample t-test: Populations have equal(ish) variances
 - if not → **Welch's approximate t-test**
 - How do we tell when populations don't have equal variances?

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

Methods:

- 1. The F-Test of equal variances**
- 2. Levene's test for homogeneity of variances**

Comparing Variances:

Hypotheses:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_A: \sigma_1^2 \neq \sigma_2^2$$

Methods:

1. The F-Test of equal variances

$$F = \frac{s_1^2}{s_2^2}$$

Basis of ANOVA 

- If variances are equal, this should be 1
- Put larger sample variance on top
 - forces it to be a one tailed test
- two different degrees of freedom:
 - **df = n_i - 1**
- very sensitive to assumption that both populations are normally distributed

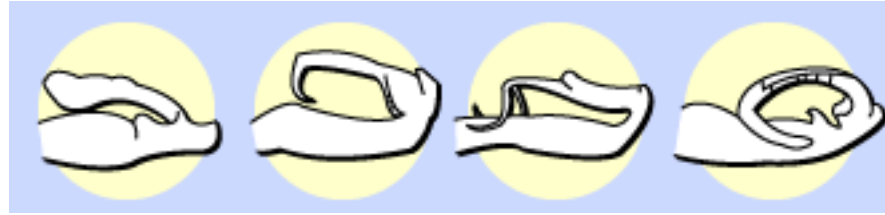
2. Levene's test for homogeneity of variances

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$



Example: Variation in insect genitalia (the above picture is from damsel fly) between polygamous species and monogamous species of insects: Lock and key or sexual selection? Goran Arnqvist expresses this data as a 'morphometric dimension' and we know that it is a normally distributed variable.

	Polygamous	Monogamous
Mean	-19.3	10.25
Sample Variance	243.9	2.27
Sample Size	7	9

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

	Polygamous	Monogamous
Mean	-19.3	10.25
Sample Variance	243.9	2.27
Sample Size	7	9

Example: Variation in insect genitalia between polygamous species and monogamous species.

$$s_1^2 = 243.9$$

$$s_2^2 = 2.27$$

$$F = \frac{s_1^2}{s_2^2} = \frac{243.9}{2.27} = 107.4$$

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

	Polygamous	Monogamous
Mean	-19.3	10.25
Sample Variance	243.9	2.27
Sample Size	7	9

Example: Variation in insect genitalia between polygamous species and monogamous species.

$$F = \frac{s_1^2}{s_2^2} = \frac{243.9}{2.27} = 107.4$$

- $df_1 = 7 - 1 = 6$
- $df_2 = 9 - 1 = 8$

$$F_{0.025, 6, 8} = 4.7$$

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

Example: Variation in insect genitalia between polygamous species and monogamous species.

Why is the critical value $\alpha/2$?

- by putting the larger variance in the numerator, we are forcing F to be greater than 1
- By the null hypothesis there is a 50:50 chance of either s^2 being greater so we want the higher tail to include just $\alpha/2$
 - *NOTE:* This is different than when the F ratio is used in ANOVA – because of the design of the ratio, MS_{group} will always be in numerator so it is just α

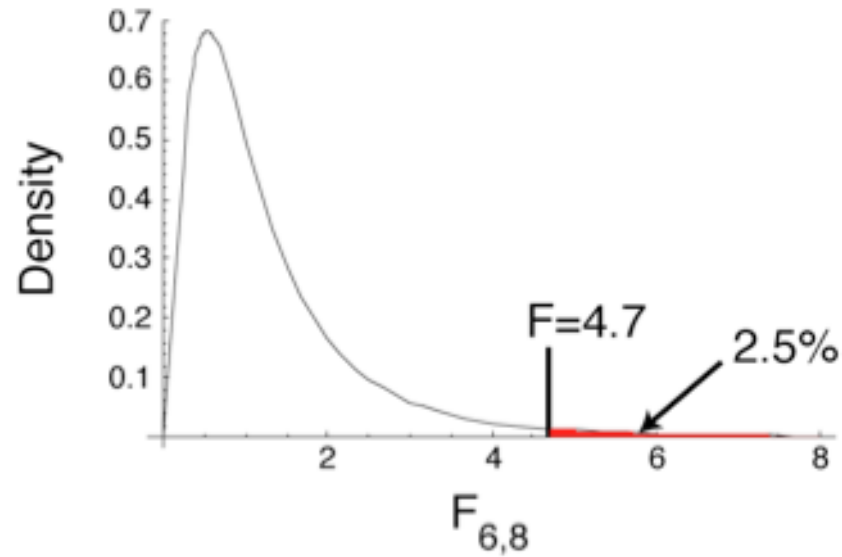
Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

Example: Variation in insect genitalia between polygamous species and monogamous species.



Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

Example: Variation in insect genitalia between polygamous species and monogamous species.

Conclusion:

- The $F = 107.4$ from the data is greater than the critical value for F with 6 and 8 dof so we can reject the null hypothesis that the variances of the two groups are equal.
- The variance in insect genitalia is much greater for polygamous species than monogamous species.

You could confirm this by determining the Confidence Interval of each of the two variances and seeing if they overlap...

$$\underline{6*243.9} < \sigma^2_1 < \underline{6*243.9} \quad ; \quad \underline{8*2.27} < \sigma^2_2 < \underline{8*2.27}$$

$$\chi^2_{0.025,6}$$

$$\chi^2_{0.975,6}$$

$$\chi^2_{0.025,8}$$

$$\chi^2_{0.975,8}$$

$$101.27 < \sigma^2_1 < 1180.16;$$

$$1.04 < \sigma^2_2 < 8.33$$

Comparing Variances:

Hypotheses:

$$H_0: \sigma^2_1 = \sigma^2_2$$

$$H_A: \sigma^2_1 \neq \sigma^2_2$$

Methods:

1. The F-Test of equal variances

2. Levene's test for homogeneity of variances

- more robust than F-test
- calculations are complex so you should know that it exists and why you would use it but you would probably use R or Python library to conduct the test.

3. Bartlett's test, Hartley's etc.

- they each have specific assumptions that must be met.

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have vastly unequal variances
- Often used as the 'default' method by programs such as R since it accounts for the Behrens-Fisher problem when variances are very unequal and still gives correct type I error rate when variances are equal (ish).



Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances
- Experimental Design:
 - 20 randomly selected burrowing owl nests
 - Randomly divided into two groups of 10 nests each
 - One group given extra dung; the other not
 - Counted number of dung beetles in owls' diets

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances
- Experimental Design:
 - 20 randomly selected burrowing owl nests
 - Randomly divided into two groups of 10 nests each
 - One group given extra dung; the other not
 - Counted number of dung beetles in owls' diets
- Summary of beetles consumed:

	Dung added	No Dung added
Mean	4.8	0.51
s	3.26	0.89

More Comparing Two Means

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances
- Summary of beetles consumed:

	Dung added	No Dung added
Mean	4.8	0.51
s	3.26	0.89

- Hypotheses:

H_0 : Owls catch the same number of dung beetles with or without extra dung ($\mu_1 = \mu_2$)

H_A : Owls do not catch the same number of dung beetles with or without extra dung ($\mu_1 \neq \mu_2$)

More Comparing Two Means

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances
- Complicated formula, but it works via tweaking (rounding down) the DEGREES OF FREEDOM

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$df = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{\frac{s_1^2}{n_1}}{n_1 - 1} + \frac{\frac{s_2^2}{n_2}}{n_2 - 1}}$$

More Comparing Two Means

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances
- Summary of beetles consumed:

	Dung added	No Dung added
Mean	4.8	0.51
s	3.26	0.89

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{4.8 - 0.51}{\sqrt{\frac{3.26^2}{10} + \frac{0.89^2}{10}}} = 4.01$$

$$df = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{s_1^2}{n_1 - 1} + \frac{s_2^2}{n_2 - 1}} = \frac{\frac{3.26^2}{10} + \frac{0.89^2}{10}}{\frac{3.26^2}{9} + \frac{0.89^2}{9}} = 10.33$$

More Comparing Two Means

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{4.8 - 0.51}{\sqrt{\frac{3.26^2}{10} + \frac{0.89^2}{10}}} = 4.01$$

$$df = \frac{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}{\frac{\frac{s_1^2}{n_1}}{n_1 - 1} + \frac{\frac{s_2^2}{n_2}}{n_2 - 1}} = \frac{\frac{3.26^2}{10} + \frac{0.89^2}{10}}{\frac{\frac{3.26^2}{10}}{9} + \frac{\frac{0.89^2}{10}}{9}} = 10.33$$

$$(\bar{Y}_1 - \bar{Y}_2) - t_{\alpha(2), df} SE_{\bar{Y}_1 - \bar{Y}_2} < \mu_1 - \mu_2 < (\bar{Y}_1 - \bar{Y}_2) + t_{\alpha(2), df} SE_{\bar{Y}_1 - \bar{Y}_2}$$

(you might notice it looks an awful lot like the other t statistic confidence intervals)

More Comparing Two Means

Welch's approximate t-test:

- it is used when comparing means of two populations that are normally distributed but have unequal variances

- Conclusion:

$$t_{0.05(2),10} = 2.23$$

Since $t = 4.01 > 2.23$, we can reject the null hypothesis with $P < 0.05$.

$$(\bar{Y}_1 - \bar{Y}_2) - t_{\alpha(2),df} SE_{\bar{Y}_1 - \bar{Y}_2} < \mu_1 - \mu_2 < (\bar{Y}_1 - \bar{Y}_2) + t_{\alpha(2),df} SE_{\bar{Y}_1 - \bar{Y}_2}$$

$$4.29 - 2.23*(1.069) < \mu_1 - \mu_2 < 4.29 + 2.23(1.069)$$

$$1.91 < \mu_1 - \mu_2 < 6.67$$

The above range does not contain “0” so it supports our rejection of the null hypothesis AND it is greater than “0” so extra dung near burrowing owls' nests increases the number of dung beetles eaten.