ECN 102, Summer 2020

Week 2 Section
Two Sample Test Examples

Two Sample Tests

- **Two-Sample Test:** We have data for two groups and we want to compare them.
- We can test comparisons of the means: a **difference in means** test.
- We can test comparisons of variance: a **difference in variances** test.

Two Sample Tests

- Two versions of the difference in means test.
- One assumes the groups have the same variance
- One assumes the groups do not have the same variance
- So let's test for difference of variances first so we can then choose which difference of means test to do.

Difference in Variances

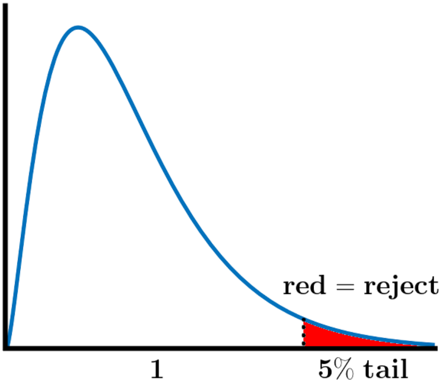
- Group A has population mean μ_A and population variance σ_A^2 with sample mean $\bar{x}_A$ and sample variance s_A^2
- Group B has population mean μ_B and population variance σ_B^2 with sample mean $\bar{x}_B$ and sample variance s_B^2
- The group with higher sample variance is group A: $s_A^2 > s_B^2$
- If group A has higher population variance, then $\sigma_A^2/\sigma_B^2 > 1$

$$H_0 : \frac{\sigma_A^2}{\sigma_B^2} \leq 1 \qquad \text{against} \qquad H_1 : \frac{\sigma_A^2}{\sigma_B^2} > 1$$

- The test statistic is

$$F \equiv \frac{s_A^2}{s_B^2} \sim F(n_A - 1, n_B - 1)$$

We reject the null hypothesis if we find too much evidence against it, i.e. if we calculate an F-statistic that is far away from 1 in the positive direction.



Difference in Variances: Cats!

Let's look at data for the weights of cats.
`http://www.wimivo.com/courses/2020Su_ECN102/cats.csv`

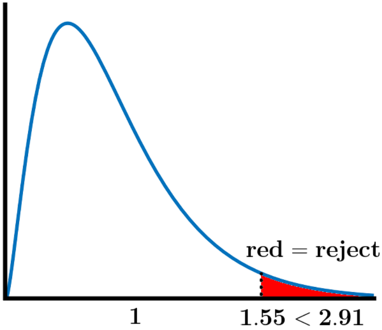
- Let's go to R.
- Male cats: $\bar{x}_A = 6.393$, $s_A^2 = 1.062$, $n_A = 97$
- Female cats: $\bar{x}_B = 5.202$, $s_B^2 = 0.365$, $n_B = 47$
- Made male cats Group A so that $\sigma_A^2/\sigma_B^2 > 1$ if true.
- The test statistic is

$$F = \frac{1.062}{0.365} \approx 2.911$$

- 5% significance critical value: `qf(1-0.05,96,46)` $= 1.551$

Difference in Variances: Cats!

The test statistic is $F = 2.91$, the critical value is 1.55



Reject null at 5% significance, conclude male cats have higher variance: more skinny male cats and more chonky male cats.

```
pf(2.91, 96, 46, lower.tail = FALSE)
var.test(male, female, alternative = "greater")
```

Difference in Means, Unequal Variances

- Group A has population mean μ_A and population variance σ_A^2 with sample mean $\bar{x}_A$ and sample variance s_A^2
- Group B has population mean μ_B and population variance σ_B^2 with sample mean $\bar{x}_B$ and sample variance s_B^2
- If they do have the same mean, then $\mu_A - \mu_B = 0$

$$H_0 : \mu_A - \mu_B = 0$$
$$H_1 : \mu_A - \mu_B \neq 0$$

- Use the test-statistic

$$t \equiv \frac{(\bar{x}_A - \bar{x}_B) - 0}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}} \sim T(n_A + n_B - 2)$$

- Then testing proceeds in the typical way.

Difference in Means: Cats!

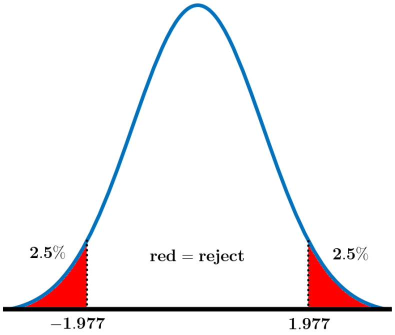
- Male cats: $\bar{x}_A = 6.393$, $s_A^2 = 1.062$, $n_A = 97$
- Female cats: $\bar{x}_B = 5.202$, $s_B^2 = 0.365$, $n_B = 47$
- The test statistic is

$$t = \frac{(6.393 - 5.202) - 0}{\sqrt{\frac{1.062}{97} + \frac{0.365}{47}}} \approx 8.710$$

- Because $n_A + n_B$ is large, we can use the 5% critical value for normal distribution: 1.96.
- But better to use `qt(1-0.05/2, 97+47-2)` $= 1.977$

Difference in Means: Cats!

The test statistic is $t = 8.71$, the critical value is 1.977



t is in rejection region: conclude that the mean weight of female cats is not equal to the mean weight of male cats

```
2*pt(8.71, 97+47-2, lower.tail = FALSE)
t.test(male, female, var.equal = FALSE)
```