Section 8-5 Comparing Variation in Two Samples

Comparing Variation in Two Samples

Assumptions

1. The two populations are *independent* of each other.
2. The two populations are each normally distributed.

When performing a hypothesis test, compare σ_1^2 to σ_2^2. Choices for H_1 are ≠ or >. (only can do right tail tests)

Mar 29-1:59 PM

Test Statistic for Hypothesis Tests with Two Variances

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

s_1^2 = *larger* of the two sample variances

n_1 = size of the sample with the *larger* variance

σ_1^2 = variance of the population from which the sample with the *larger* variance was drawn

The symbols s_2^2, n_2, and σ_2^2 are used for the other sample and population.

Mar 29-2:01 PM

p.768 -773

Table A-5 F-Distribution

(α value specified)

Numerator degrees of freedom (df_1) larger

Denominator degrees of freedom (df_2) smaller

	1	2	3	4	.	.	.
1							
2							
3							
4							
5							
.							
.							
.							

Mar 29-2:10 PM

F - distribution

Critical Values: Using Table A-5, we obtain critical *F* values that are determined by the following three values:

1. The significance level α.
2. Numerator degrees of freedom (df_1) = $n_1 - 1$ (top row)
3. Denominator degrees of freedom (df_2) = $n_2 - 1$ (column on left)

F -distribution is NOT symmetric. All values are positive (similar to X^2). When looking up the values the table will only give you the value for the right side. The left value is the reciprocal of the right critical value.

ex) α = .05 2 tailed test n_1=25 and n_2 = 20.
use α= .025 df top row = 24 and df column = 19.

.4078 2.4523

Mar 29-2:03 PM

- All one-tailed tests will be right-tailed.
- All two-tailed tests will need only the critical value to the right.
- When degrees of freedom is not listed exactly, use the critical values on either side as an interval. Use interpolation only if the test statistic falls within the interval.

Mar 29-2:06 PM

If the two populations do have equal variances, then $F = \frac{s_1^2}{s_2^2}$ will be close to 1 because s_1^2 and s_2^2 are close in value.

If the two populations have radically different variances, then $F = \frac{s_1^2}{s_2^2}$ will be a large number.

Remember, the larger sample variance will be s_1^2.

Consequently, a value of *F* near 1 will be evidence in favor of the conclusion that $\sigma_1^2 = \sigma_2^2$.

But a large value of *F* will be evidence against the conclusion of equality of the population variances.

Mar 29-2:07 PM

Coke Versus Pepsi

Ex1) **Data Set 1 in Appendix B includes the weights (in pounds) of samples of regular Coke and regular Pepsi. Sample statistics are shown. Use the 0.05 significance level to test the claim that the weights of regular Coke and the weights of regular Pepsi have the same standard deviation.**

	Regular Coke	Regular Pepsi
n	36	36
$\bar{x}$	0.81682	0.82410
s	0.007507	0.005701

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

1. $H_o : \sigma_1^2 = \sigma_2^2$

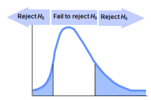
$H_1 : \sigma_1^2 \neq \sigma_2^2$

2. Value of $F = \frac{s_1^2}{s_2^2}$

$= \frac{0.007507^2}{0.005701^2}$

$= 1.7339$

3.



0 .5333 F = above 1.8752

4. Fail to Reject: There is sufficient evidence to support the standard deviations are the same.

Mar 29-2:11 PM

Ex2) In a sample of 40 math majors, they had a mean score of 75.0 and a standard deviation of 15.0. In a sample of 60 English majors, they had a mean score of 70.0 and standard deviation of 14.0. Determine if the English majors have less variability compared to the math majors. Use $\alpha = .05$

This a one-tailed test, so have to use greater than symbol for the alternative hypothesis.

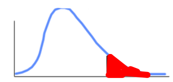
1. Math (group #1) > English (group #2)

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

H_1: $\sigma_1^2 > \sigma_2^2$

2. Test Statistic: $= \frac{S_1^2}{S_2^2}$ $\frac{\text{math}}{\text{English}}$ $\frac{15.0}{14.0}$ $= 1.0714$

3. Critical Value:



df top row =39 (math)
df column = 59 (English)

1.5943

4. Fail to reject. There is not sufficient evidence to support that the English majors have less variability than the math majors.

Mar 29-2:18 PM