Unit #7 (b): Hypothesis Testing

9.7.3-9.7.5, 9.9.4

(Two Sample)

Learning Objectives

At the end of this unit, students should be able to:

1. State and interpret hypotheses about two samples.

2. Write and compute the test statistic for the a test comparing two population means with large samples (when the standard deviation is known and unknown). Use the test statistic to conduct the test and make a decision (using either a rejection region or a p-value).

3. Distinguish between lower-tail, upper-tail, and two-tailed tests. Reason about when each should be used.

4. Write and compute the test statistic for the two sample t-test. Recognize when this test applies. Use the test statistic to conduct the test and make a decision (using either a rejection region or a p-value).

5. Write and compute the test statistic for the test for proportion differences. Recognize when this test applies. Use the test statistic to conduct the test and make a decision (using either a rejection region or a p-value).

6. Describe the F distribution, including what it is parameterized by, how it arises, and what it is used to test.

Pop #1 has mean μ_1, var σ_1^2
Pop #2 has mean μ_2, var σ_2^2

$H_0: \mu_1 = \mu_2$; $H_1: \mu_1 \neq \mu_2$ $\boxed{OR}$ $H_1: \mu_1 < \mu_2$
$\boxed{OR}$ $H_1: \mu_1 > \mu_2$

Test Procedures for Normal Populations with Known Variances

Null Hypothesis: $H_0: \mu_1 - \mu_2 = c$, $c \in \mathbb{R}$

- Let $X_1, \ldots, X_{n_1}$ be a sample from pop #1.
- Let $Y_1, \ldots, Y_{n_2}$ be a sample from pop #2.

(samples are ind)

Test Statistic: $$\frac{\bar{X} - \bar{Y} - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = Z \sim N(0,1)$$

Alternative Hypothesis | Rejection Region for Level α Test

$\mu_1 - \mu_2 > c$ | $Z \geq Z_\alpha$

$\mu_1 - \mu_2 < c$ | $Z \leq -Z_\alpha$

$\mu_1 - \mu_2 \neq c$ | $Z \leq -Z_{\alpha/2}$ or $Z \geq Z_{\alpha/2}$

Test Procedures for Normal Populations with Known Variances

n_1

μ_1 Analysis of a random sample consisting of 20 specimens of cold-rolled steel to determine yield strengths resulted in a sample average strength of $\bar{x} = 29.8$ ksi.

$H_0: \mu_1 - \mu_2 = 0$

n_2

$H_1: \mu_1 - \mu_2 \neq 0$

A second random sample of 25 two-sided galvanized steel specimens gave a sample average strength of $\bar{y} = 34.7$ ksi.

σ_1 σ_2

Assuming that the two yield-strength distributions are normal with standard deviations of 4.0 and 5.0, respectively, does the data indicate that the corresponding true average yield strengths μ_1 and μ_2 are different? Let's carry out a test at significance level = 0.01. $= \alpha$

$$Z = \frac{29.8 - 34.7}{\sqrt{\frac{4^2}{20} + \frac{5^2}{25}}} = -3.65$$

$H_0: \mu_1 - \mu_2 = 0$
$H_1: \mu_1 - \mu_2 \neq 0$

Large Sample Tests

The assumptions of normal population distributions and known standard deviations are fortunately unnecessary when both sample sizes are sufficiently large. WHY?

Furthermore, using S_x^2 and S_y^2 in place of σ_x^2 and σ_y^2 gives a variable whose distribution is approximately standard normal:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_x^2}{n_1} + \frac{S_y^2}{n_2}}} \sim N(0,1)$$

These tests are usually appropriate if both $n_1 > 30$ and $n_2 > 30$.

Talk today @ 3:00pm in ECCR 245: "Bias in, Bias Out: Predictive Models in the Criminal Justice system".

Large Sample Tests

Example: Data on daily calorie intake both for a sample of teens who said they did not typically eat fast food and another sample of teens who said they did usually eat fast food.

Eat Fast Food	Sample Size	Sample Mean	Sample SD
No (X) (μ_1)	663 $= n_1$	2258 $= \bar{X}$	1519 $= s_x$
Yes (Y) (μ_2)	413 $= n_2$	2637 $= \bar{Y}$	1138 $= s_y$

Does this data provide strong evidence for concluding that true average calorie intake for teens who typically eat fast food exceeds more than 200 calories per day the true average intake for those who don't typically eat fast food?

$H_0: \mu_2 - \mu_1 \leq 200$

$H_1: \mu_2 - \mu_1 > 200$

Let's investigate by carrying out a test of hypotheses at a significance level of 0.05. "α"

$$Z = \frac{(\bar{Y} - \bar{X}) - (\overbrace{\mu_2 - \mu_1}^{200})}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}} = 2.2 > 1.64$$

$\Rightarrow$ Reject H_0.

$Z_\alpha = 1.64$

$H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 < 0$

Two Sample t-Test

When the population distribution are both normal, the standardized variable

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}} \sim t_v$$

has approximately a *t* distribution with df *v* estimated from the data by

$$\nu = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}}$$

Skip

Two Sample t-Test

Null Hypothesis:

Test Statistic:

Alternative Hypothesis | Rejection Region for Level α Test

Skip

A Test for Proportions

Theoretically, we know that:

has approximately a standard normal distribution when H_0 is true.

However, this Z cannot serve as a test statistic because the value of p is unknown—H_0 asserts only that there is a common value of p, but does not say what that value is.

A Test for Proportions

Under the null hypothesis, we assume that $p_1 = p_2 = p$, instead of separate samples of size m and n from two different populations (two different binomial distributions). So, we really have a single sample of size $m + n$ from one population with proportion p.

The total number of individuals in this combined sample having the characteristic of interest is $X + Y$.

The estimator of p is then

A Test for Proportions

Null Hypothesis:

Test Statistic:

Alternative Hypothesis | Rejection Region for Level α Test

The *F* Test for Equality of Variances

The F Distribution

The *F* probability distribution has two parameters, denoted by v_1 and v_2. The parameter v_1 is called the *numerator degrees of freedom*, and v_2 is the *denominator degrees of freedom*.

A random variable that has an *F* distribution cannot assume a negative value. The density function is complicated and will not be used explicitly, so it's not shown.

There is an important connection between an *F* variable and chi-squared variables.

$\frac{X_1}{v_1} = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

The F Distribution

If X_1 and X_2 are independent chi-squared rv's with v_1 and v_2 df, respectively, then the rv

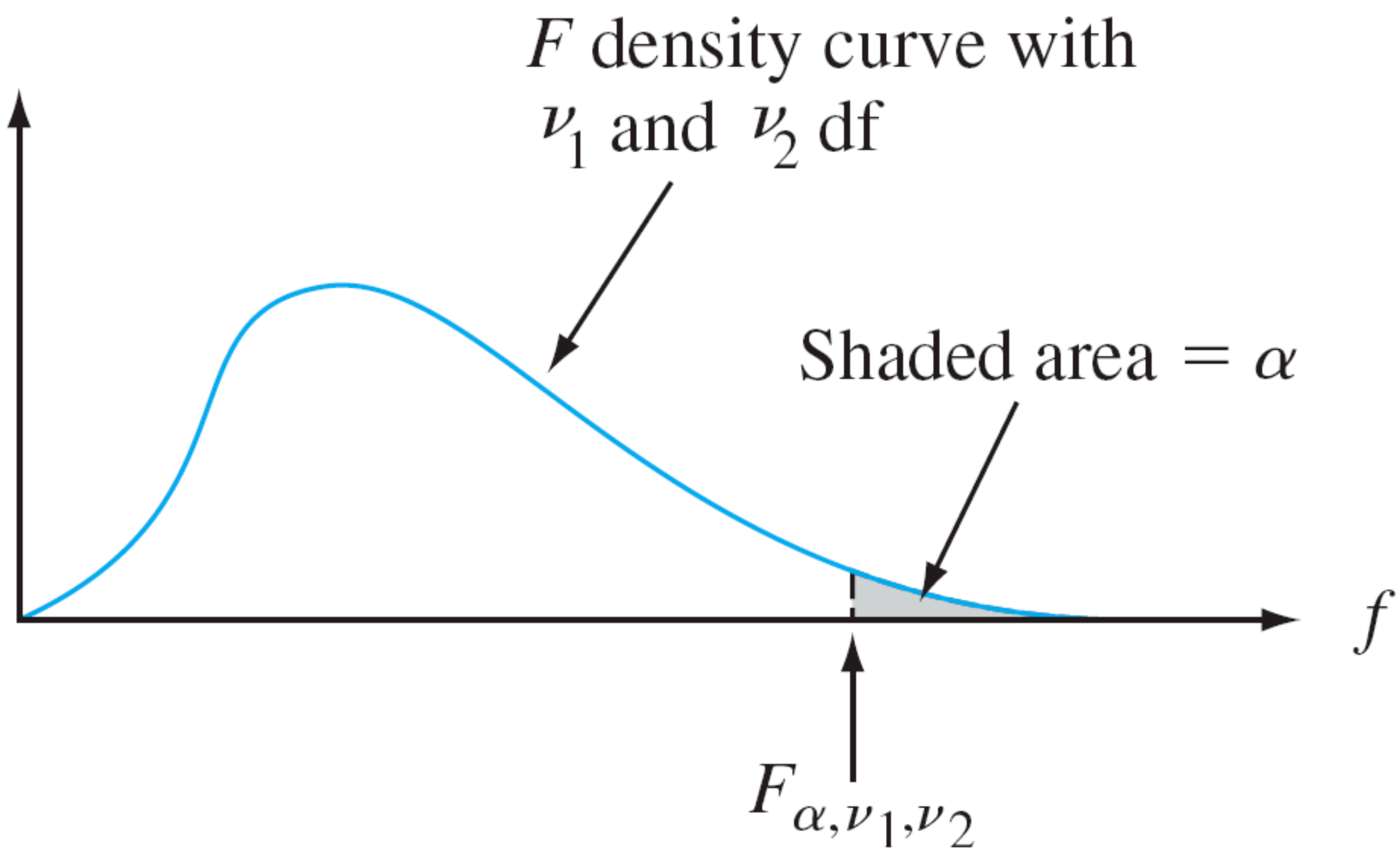
$$F = \frac{X_1/v_1}{X_2/v_2}$$

can be shown to have an *F* distribution.

Recall that a chi-squared distribution was obtain by summing squared standard Normal variables (such as squared deviations for example). So a scaled ratio of two variances is a ratio of two scaled chi-squared variables.

The F Distribution

Figure below illustrates a typical *F* density function.



The F Distribution

We use F_{α,v_1,v_2} for the value on the horizontal axis that captures of the area under the *F* density curve with v_1 and v_2 df in the upper tail.

The density curve is not symmetric, so it would seem that both upper- and lower-tail critical values must be tabulated. This is not necessary, though, because of the fact that

$$F_{1-\alpha,v_1,v_2} = 1/F_{\alpha,v_2,v_1}.$$

For example, $F_{.05,6,10} = 3.22$ and $F_{.95,10,6} = 0.31 = 1/3.22$.

F-Test for Equal Variances

A test procedure for hypotheses concerning the ratio σ_1^2/σ_2^2 is based on the following result.

Theorem

Let $X_1, \ldots, X_m$ be a random sample from a normal distribution with variance σ_1^2, let $Y_1, \ldots, Y_n$ be another random sample (independent of the X_i's) from a normal distribution with variance σ_2^2, and let S_1^2 and S_2^2 denote the two sample variances. Then the rv

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

has an F distribution with $v_1 = m - 1$ and $v_2 = n - 1$.

F-Test for Equal Variances

This theorem results from combining the fact that the variables $(n - 1)S_2^2/\sigma_2^2$ and $(m - 1)S_1^2/\sigma_1^2$ each have a chi-squared distribution with $m – 1$ and $n – 1$ df, respectively.

Because F involves a ratio rather than a difference, the test statistic is the ratio of sample variances.

The claim that $\sigma_1^2 = \sigma_2^2$ is then rejected if the ratio differs by too much from 1.

F-Test for Equal Variances

Null Hypothesis: $H_0\colon \sigma_1^2 = \sigma_2^2$

Test Statistic: $f = s_1^2/s_2^2$

Alternative Hypothesis | Rejection Region for Level α Test

$H_a\colon \sigma_1^2 > \sigma_2^2$	$f \geq F_{\alpha,m-1,n-1}$
$H_a\colon \sigma_1^2 < \sigma_2^2$	$f \leq F_{1-\alpha,m-1,n-1}$
$H_a\colon \sigma_1^2 \neq \sigma_2^2$	either $f \geq F_{\alpha/2,m-1,n-1}$ or $f \leq F_{1-\alpha/2,m-1,n-1}$

F-Test for Equal Variances

<u>Example</u>: On the basis of data reported in the article "Serum Ferritin in an Elderly Population" (*J. of Gerontology*, 1979: 521–524), the authors concluded that the ferritin distribution in the elderly had a smaller variance than in the younger adults. (Serum ferritin is used in diagnosing iron deficiency.)

For a sample of 28 elderly men, the sample standard deviation of serum ferritin (mg/L) was $s_1 = 52.6$; for 26 young men, the sample standard deviation was $s_2 = 84.2$.

Does this data support the conclusion as applied to men? Use alpha = .01.