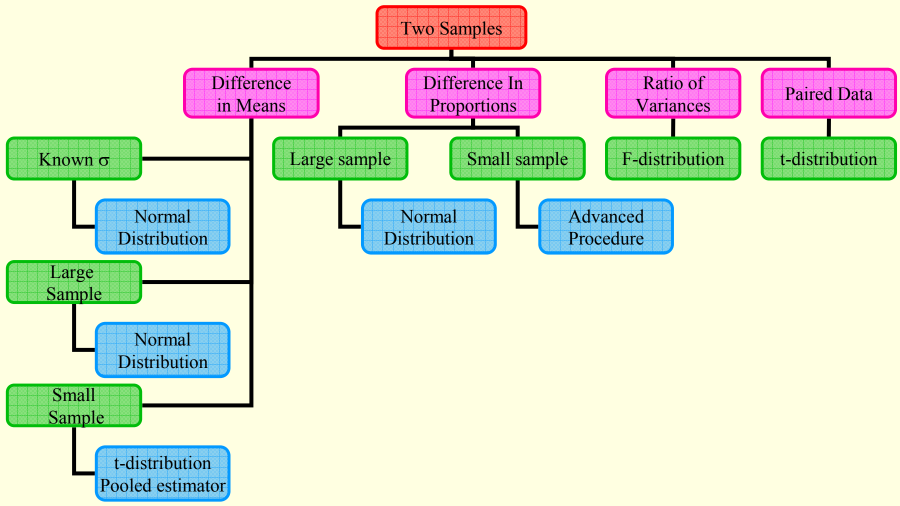
University of Jordan
Chemical Engineering Department
Process Modeling by Statistical Methods– 905331

Lecture 08: Testing of Hypothesis – Two Samples

Dr. Ali Khalaf Al-Matar

Introduction

- In the previous part we discussed inferences regarding a single population parameter.
- In this lecture we are interested in situations involving means, proportions, and variances of two different population distributions. Namely, we deal with
  - $X_1, \ldots, X_m$ is a random sample from a population with mean μ_1 and variance σ_1^2.
  - $Y_1, \ldots, Y_n$ is a random sample from a population with mean μ_2 and variance σ_2^2.
  - The X and Y samples are independent of one another.



Difference Between Two Population Means

- The natural estimator of μ_1-μ_2 is the difference in the two sample means.

$$E(\bar{X} - \bar{Y}) = E(\bar{X}) - E(\bar{Y}) = \mu_1 - \mu_2$$

$$V(\bar{X} - \bar{Y}) = V(\bar{X}) + V(\bar{Y}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$\sigma_{\bar{X} - \bar{Y}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

- Consider $\mu_1 - \mu_2$ as a parameter θ, then its estimator is $\hat{\theta} = \overline{X} - \overline{Y}$ with a standard deviation $\sigma_{\hat{\theta}}$. When the variances of the two populations are known, the test statistic takes the form $(\hat{\theta} - \text{null value})/\sigma_{\hat{\theta}}$ which is used widely in the single sample test of hypothesis as a test statistic.
- When the population variances are not known, substitute them with the sample variances.

Normal Population with Known Variances

Null hypothesis $H_0 : \mu_1 - \mu_2 = \Delta_0$

Test Statistic value: $z = \dfrac{\overline{x}_1 - \overline{x}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$z \geq z_\alpha$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$z \leq -z_\alpha$
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

$\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}$

Confidence interval for $\mu_1 - \mu_2$ with a confidence level 100(1-α)%

Normal Population with Unknown Variances

- Replace population variances with their sample equivalents

Null hypothesis $H_0: \mu_1 - \mu_2 = \Delta_0$

Test Statistic value: $z = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

Alternative hypothesis	Rejection region for level α test
$H_1: \mu_1 - \mu_2 > \Delta_0$	$z \geq z_\alpha$
$H_1: \mu_1 - \mu_2 < \Delta_0$	$z \leq -z_\alpha$
$H_1: \mu_1 - \mu_2 \neq \Delta_0$	either $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$

$$\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Provided that *m* and *n* are both large, a confidence interval for $\mu_1 - \mu_2$ with a confidence level $100(1-\alpha)\%$

- A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested; formulation 1 is the standard chemistry, and formulation 2 has a new drying ingredient that should reduce the drying time. From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient. Ten specimens are painted with formulation 1, and another 10 specimens are painted with formulation 2; the 20 specimens are painted in random order. The two sample average drying times are 121 minutes and 112 minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using $\alpha = 0.05$?

1. The quantity of interest is the difference in mean drying times, $\mu_1 - \mu_2$, and $\Delta_0 = 0$
2. H_0: $\mu_1 - \mu_2 = 0$, or H_0: $\mu_1 = \mu_2$.
3. H_1: $\mu_1 > \mu_2$. We want to reject H_0 if the new ingredient reduces mean drying time
4. $\alpha = 0.05$
5. The test statistic is

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

   where $\sigma_1^2 = \sigma_2^2 = (8)^2 = 64$ and $n_1 = n_2 = 10$.
6. Reject H_0: $\mu_1 = \mu_2$ if $z_0 > 1.645 = z_{0.05}$.
7. Computations: Since $\bar{x}_1 = 121$ minutes and $\bar{x}_2 = 112$ minutes, the test statistic is

$$z_0 = \frac{121 - 112}{\sqrt{\frac{(8)^2}{10} + \frac{(8)^2}{10}}} = 2.52$$

8. Conclusion: Since $z_0 = 2.52 > 1.645$, we reject H_0: $\mu_1 = \mu_2$ at the $\alpha = 0.05$ level and conclude that adding the new ingredient to the paint significantly reduces the drying time. Alternatively, we can find the P-value for this test as

$$P\text{-value} = 1 - \Phi(2.52) = 0.0059$$

   Therefore, H_0: $\mu_1 = \mu_2$ would be rejected at any significance level $\alpha \geq 0.0059$.

Two-sample *t*-Test

- Many problems involve at least one sample with a small size and unknown variances
- Subject to the following two assumptions
  - Both populations are normal, so that $X_1,\ldots,X_n$ is a random sample from a normal distribution and so is $Y_1,\ldots,Y_m$ (with the X's and Y's independent of each other), and
  - The values of the two population variances σ_1^2 and σ_2^2 are equal, so that their common value can be denoted by σ^2 (which is unknown).

The Pooled Estimator of σ^2

- The natural estimator of $\mu_1 - \mu_2$ is the difference in the two sample means. However the variance is simplified since σ is the same

$$E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$$

$$V(\bar{X}_1 - \bar{X}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

- The pooled estimator of the common variance σ^2, denoted σ_p^2, is defined by

$$s_p^2 = \frac{(n_1 - 1)}{n_1 + n_2 - 2} s_1^2 + \frac{(n_2 - 1)}{n_1 + n_2 - 2} s_2^2$$
$$= w s_1^2 + (1 - w) s_2^2$$

- The test statistic has a *t* distribution with *m* + *n* - 2 degrees of freedom (**df**)

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

The Pooled *t*-Test

Null hypothesis $H_0 : \mu_1 - \mu_2 = \Delta_0$

Test Statistic value: $t = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t \geq t_{\alpha, n_1+n_2-2}$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t \leq -t_{\alpha, n_1+n_2-2}$
$H_1 : \mu_1 - \mu_2 \neq \Delta_0$	either $t \geq t_{\alpha, n_1+n_2-2}$ or $t \leq -t_{\alpha, n_1+n_2-2}$

$\bar{x} - \bar{y} \pm t_{\alpha/2, m+n-2} s_p \sqrt{\dfrac{1}{m} + \dfrac{1}{n}}$

The confidence interval for $\mu_1 - \mu_2$ with a confidence level $100(1-\alpha)\%$

- Two catalysts are being analyzed to determine how they affect the mean yield of a chemical process. Specifically, catalyst 1 is currently in use, but catalyst 2 is acceptable. Since catalyst 2 is cheaper, it should be adopted, providing it does not change the process yield. A test is run in the pilot plant and results in the data shown in Table 10-1. Is there any difference between the mean yields? Use $\alpha = 0.05$, and assume equal variances.

Table 10-1 Catalyst Yield Data, Example 10-5

Observation Number	Catalyst 1	Catalyst 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$

1. The parameters of interest are μ_1 and μ_2, the mean process yield using catalysts 1 and 2, respectively, and we want to know if $\mu_1 - \mu_2 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, or H_0: $\mu_1 = \mu_2$
3. H_1: $\mu_1 \neq \mu_2$
4. $\alpha = 0.05$
5. The test statistic is

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

6. Reject H_0 if $t_0 > t_{0.025,14} = 2.145$ or if $t_0 < -t_{0.025,14} = -2.145$.
7. Computations: From Table 10-1 we have $\bar{x}_1 = 92.255$, $s_1 = 2.39$, $n_1 = 8$, $\bar{x}_2 = 92.733$, $s_2 = 2.98$, and $n_2 = 8$. Therefore

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(7)(2.39)^2 + 7(2.98)^2}{8 + 8 - 2} = 7.30$$

$$s_p = \sqrt{7.30} = 2.70$$

and

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{2.70\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{92.255 - 92.733}{2.70\sqrt{\frac{1}{8} + \frac{1}{8}}} = -0.35$$

8. Conclusions: Since $-2.145 < t_0 = -0.35 < 2.145$, the null hypothesis cannot be rejected. That is, at the 0.05 level of significance, we do not have strong evidence to conclude that catalyst 2 results in a mean yield that differs from the mean yield when catalyst 1 is used.

Test Procedure when $\sigma_1^2 \neq \sigma_2^2$

- The t test is robust in the presence of mild departures from the basic assumptions.
- The t test is more robust for departures from assumptions when $m = n$ than when $m \neq n$.
- Use the Smith-Satterthwaite test when $\sigma_1^2 \neq \sigma_2^2$,
- The number of degrees of freedom is estimated from the data

Smith-Satterthwaite Test

Null hypothesis $H_0 : \mu_1 - \mu_2 = \Delta_0$

Test Statistic value: $t = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu_1 - \mu_2 > \Delta_0$	$t \ge t_{\alpha, n_1+n_2-2}$
$H_1 : \mu_1 - \mu_2 < \Delta_0$	$t \le -t_{\alpha, n_1+n_2-2}$
$H_1 : \mu_1 - \mu_2 \ne \Delta_0$	either $t \ge t_{\alpha, n_1+n_2-2}$ or $t \le -t_{\alpha, n_1+n_2-2}$

Degrees of Freedom: $\nu = \dfrac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}}$

- Arsenic concentration in public drinking water supplies is a potential health risk. An article in the *Arizona Republic* (Sunday, May 27, 2001) reported drinking water arsenic concentrations in parts per billion (ppb) for 10 metropolitan Phoenix communities and 10 communities in rural Arizona. The data follow:

Metro Phoenix ($\bar{x}_1 = 12.5, s_1 = 7.63$)	Rural Arizona ($\bar{x}_2 = 27.5, s_2 = 15.3$)
Phoenix, 3	Rimrock, 48
Chandler, 7	Goodyear, 44
Gilbert, 25	New River, 40
Glendale, 10	Apachie Junction, 38
Mesa, 15	Buckeye, 33
Paradise Valley, 6	Nogales, 21
Peoria, 12	Black Canyon City, 20
Scottsdale, 25	Sedona, 12
Tempe, 15	Payson, 1
Sun City, 7	Casa Grande, 18

1. The parameters of interest are the mean arsenic concentrations for the two geographic regions, say, μ_1 and μ_2, and we are interested in determining whether $\mu_1 - \mu_2 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, or H_0: $\mu_1 = \mu_2$
3. H_1: $\mu_1 \neq \mu_2$
4. $\alpha = 0.05$ (say)
5. The test statistic is

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6. The degrees of freedom on t_0^* are found from Equation 10-16 as

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left[\frac{(7.63)^2}{10} + \frac{(15.3)^2}{10}\right]^2}{\frac{[(7.63)^2/10]^2}{9} + \frac{[(15.3)^2/10]^2}{9}} = 13.2 \simeq 13$$

Therefore, using $\alpha = 0.05$, we would reject H_0: $\mu_1 = \mu_2$ if $t_0^* > t_{0.025,13} = 2.160$ or if $t_0^* < -t_{0.025,13} = -2.160$

7. Computations: Using the sample data we find

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{12.5 - 27.5}{\sqrt{\frac{(7.63)^2}{10} + \frac{(15.3)^2}{10}}} = -2.77$$

8. Conclusions: Because $t_0^* = -2.77 < t_{0.025,13} = -2.160$, we reject the null hypothesis. Therefore, there is evidence to conclude that mean arsenic concentration in the drinking water in rural Arizona is different from the mean arsenic concentration in metropolitan Phoenix drinking water. Furthermore, the mean arsenic concentration is higher in rural Arizona communities. The *P*-value for this test is approximately $P = 0.016$.

Analysis of Paired Data

- Consider a certain data consisting of *n* independently selected pairs $(X_1,Y_1),\ldots,(X_n,Y_n)$, with $E(X_i) = \mu_1$ and $E(Y_i) = \mu_2$. Let $d_1 = X_1 - Y_1, \ldots, d_n = X_n - Y_n$ be the differences among pairs. Then the d_i's are assumed to be normally distributed with a variance σ_D^2.

The Paired *t*-Test

Null hypothesis $H_0 : \mu_D = \Delta_0$

Test Statistic value: $t_0 = \dfrac{\bar{D} - \Delta_0}{s_D / \sqrt{n}}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \mu_D > \Delta_0$	$t \geq t_{\alpha,n-1}$
$H_1 : \mu_D < \Delta_0$	$t \leq -t_{\alpha,n-1}$
$H_1 : \mu_D \neq \Delta_0$	either $t \geq t_{\alpha/2,n-1}$ or $t \leq -t_{\alpha/2,n-1}$

$\bar{D} \pm t_{\alpha/2,n-1} s_D / \sqrt{n}$

The confidence interval for *d* with a confidence level 100(1 - α)%

- An article in the *Journal of Strain Analysis* (1983, Vol. 18, No. 2) compares several methods for predicting the shear strength for steel plate girders. Data for two of these methods, the Karlsruhe and Lehigh procedures, when applied to nine specific girders, are shown in Table 10-2. We wish to determine whether there is any difference (on the average) between the two methods.

Table 10-2 Strength Predictions for Nine Steel Plate Girders (Predicted Load/Observed Load)

Girder	Karlsruhe Method	Lehigh Method	Difference d_j
S1/1	1.186	1.061	0.119
S2/1	1.151	0.992	0.159
S3/1	1.322	1.063	0.259
S4/1	1.339	1.062	0.277
S5/1	1.200	1.065	0.138
S2/1	1.402	1.178	0.224
S2/2	1.365	1.037	0.328
S2/3	1.537	1.086	0.451
S2/4	1.559	1.052	0.507

1. The parameter of interest is the difference in mean shear strength between the two methods, say, $\mu_D = \mu_1 - \mu_2 = 0$.
2. H_0: $\mu_D = 0$
3. H_1: $\mu_D \neq 0$
4. $\alpha = 0.05$
5. The test statistic is

$$t_0 = \frac{\bar{d}}{s_D/\sqrt{n}}$$

6. Reject H_0 if $t_0 > t_{0.025,8} = 2.306$ or if $t_0 < -t_{0.025,8} = -2.306$.
7. Computations: The sample average and standard deviation of the differences d_j are $\bar{d} = 0.2736$ and $s_D = 0.1356$, so the test statistic is

$$t_0 = \frac{\bar{d}}{s_D/\sqrt{n}} = \frac{0.2736}{0.1356/\sqrt{9}} = 6.05$$

8. Conclusions: Since $t_0 = 6.05 > 2.306$, we conclude that the strength prediction methods yield different results. Specifically, the data indicate that the Karlsruhe method produces, on the average, higher strength predictions than does the Lehigh method. The P-value for $t_0 = 6.05$ is $P = 0.0002$, so the test statistic is well into the critical region.

Pooled and Paired Tests

- Whenever there is a positive dependence within pairs, the denominator for the paired *t* statistic should be smaller than that for *t* of the independent samples test.
- This can be shown to be valid using the joint distribution variance

$$V(X \pm Y) = V(X) + V(Y) \pm 2Cov(X,Y)$$

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{V(X) \bullet V(Y)}}$$

$$V(X - Y) = \sigma^2 + \sigma^2 - 2\rho\sigma^2 = 2\sigma^2(1-\rho)$$

$$V(\bar{X} - \bar{Y}) = V(\sum D_i / n) = V(D_i)/n = 2\sigma^2(1-\rho)$$

$\rho = 0$ for two sample t test

Differences Between Population Proportions

Let X~Bin(m,p_1) and Y~Bin(n,p_2)
with X and Y independent variables. Then,

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$

$$V(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{m} + \frac{p_2 q_2}{n}$$

Large Sample with p_1 - p_2 = 0

Null hypothesis $H_0 : p_1 - p_2 = 0$

Test Statistic value: $z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)}}$

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

Alternative hypothesis	Rejection region for level α test
$H_1 : p_1 - p_2 > 0$	$z \ge z_\alpha$
$H_1 : p_1 - p_2 < 0$	$z \le -z_\alpha$
$H_1 : p_1 - p_2 \ne 0$	either $z \ge z_{\alpha/2}$ or $z \le -z_{\alpha/2}$

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{m} + \frac{\hat{p}_2(1-\hat{p}_2)}{n}}$$

- Extracts of St. John's Wort are widely used to treat depression. An article in the April 18, 2001 issue of the *Journal of the American Medical Association* ("Effectiveness of St. John's Wort on Major Depression: A Randomized Controlled Trial") compared the efficacy of a standard extract of St. John's Wort with a placebo in 200 outpatients diagnosed with major depression. Patients were randomly assigned to two groups; one group received the St. John's Wort, and the other received the placebo. After eight weeks, 19 of the placebo-treated patients showed improvement, whereas 27 of those treated with St. John's Wort improved. Is there any reason to believe that St. John's Wort is effective in treating major depression? Use α=0.05.

1. The parameters of interest are p_1 and p_2, the proportion of patients who improve following treatment with St. John's Wort (p_1) or the placebo (p_2).
2. H_0: $p_1 = p_2$
3. H_1: $p_1 \neq p_2$
4. $\alpha = 0.05$
5. The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p}_1 = 27/100 = 0.27$, $\hat{p}_2 = 19/100 = 0.19$, $n_1 = n_2 = 100$, and

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{19 + 27}{100 + 100} = 0.23$$

6. Reject H_0: $p_1 = p_2$ if $z_0 > z_{0.025} = 1.96$ or if $z_0 < -z_{0.025} = -1.96$.
7. Computations: The value of the test statistic is

$$z_0 = \frac{0.27 - 0.19}{\sqrt{0.23(0.77)\left(\frac{1}{100} + \frac{1}{100}\right)}} = 1.35$$

8. Conclusions: Since $z_0 = 1.35$ does not exceed $z_{0.025}$, we cannot reject the null hypothesis. Note that the P-value is $P \simeq 0.177$. There is insufficient evidence to support the claim that St. John's Wort is effective in treating major depression.

Two Population Variances: *F* Distribution

- The F distribution has two parameters, denoted by ν_1 and ν_2. The parameter ν_1 is called the number of **numerator** degrees of freedom, and ν_2 is called the number of **denominator** degrees of freedom. Both are positive integers.
- Tow chi-squared random variables that are divided by their respective degrees of freedom have an *F* distribution

$$F = \frac{\chi_1 / \nu_1}{\chi_2 / \nu_2}$$

$$F_{1-\alpha,\nu_1,\nu_2} = 1 / F_{\alpha,\nu_2,\nu_1}$$

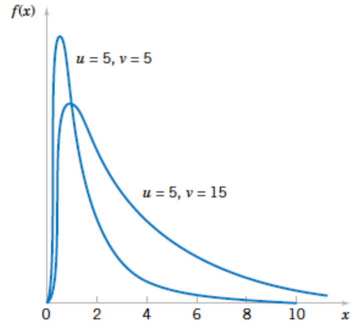


Figure 10-4 Probability density functions of two *F* distributions.

Testing Variances

- Use the ratio of variances as test statistic
- Let $X_1,\ldots,X_m$ be a random sample from a normal distribution with a variance σ_1^2, let $Y_1,\ldots,Y_n$ be a random sample from a normal distribution with a variance σ_2^2, *X* and *Y* are independent. The sample variances are denoted as s_1^2 and s_2^2. Then the rv

$$F = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2}$$

Has an *F* distribution with $\nu_1 = m - 1$ and $\nu_2 = n - 1$

Test Procedure

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

Test Statistic value: $f_0 = \dfrac{s_1^2}{s_2^2}$

Alternative hypothesis	Rejection region for level α test
$H_1 : \sigma_1^2 > \sigma_2^2$	$f_0 \geq F_{\alpha, n_1-1, n_2-1}$
$H_1 : \sigma_1^2 < \sigma_2^2$	$f_0 \leq F_{1-\alpha, n_1-1, n_2-1}$
$H_1 : \sigma_1^2 \neq \sigma_2^2$	either $f_0 \geq F_{\alpha/2, n_1-1, n_2-1}$ or $f_0 \leq F_{1-\alpha/2, n_1-1, n_2-1}$

- Oxide layers on semiconductor wafers are etched in a mixture of gases to achieve the proper thickness. The variability in the thickness of these oxide layers is a critical characteristic of the wafer, and low variability is desirable for subsequent processing steps. Two different mixtures of gases are being studied to determine whether one is superior in reducing the variability of the oxide thickness. Twenty wafers are etched in each gas. The sample standard deviations of oxide thickness are1.96 and 2.13 angstroms, respectively. Is there any evidence to indicate that either gas is preferable? Use α=0.05.

1. The parameters of interest are the variances of oxide thickness σ_1^2 and σ_2^2. We will assume that oxide thickness is a normal random variable for both gas mixtures.
2. H_0: $\sigma_1^2 = \sigma_2^2$
3. H_1: $\sigma_1^2 \neq \sigma_2^2$
4. $\alpha = 0.05$
5. The test statistic is given by Equation 10-29:

$$f_0 = \frac{s_1^2}{s_2^2}$$

6. Since $n_1 = n_2 = 20$, we will reject H_0: $\sigma_1^2 = \sigma_2^2$ if $f_0 > f_{0.025,19,19} = 2.53$ or if $f_0 < f_{0.975,19,19} = 1/f_{0.025,19,19} = 1/2.53 = 0.40$.
7. Computations: Since $s_1^2 = (1.96)^2 = 3.84$ and $s_2^2 = (2.13)^2 = 4.54$, the test statistic is

$$f_0 = \frac{s_1^2}{s_2^2} = \frac{3.84}{4.54} = 0.85$$

8. Conclusions: Since $f_{0.975,19,19} = 0.40 < f_0 = 0.85 < f_{0.025,19,19} = 2.53$, we cannot reject the null hypothesis H_0: $\sigma_1^2 = \sigma_2^2$ at the 0.05 level of significance. Therefore, there is no strong evidence to indicate that either gas results in a smaller variance of oxide thickness.