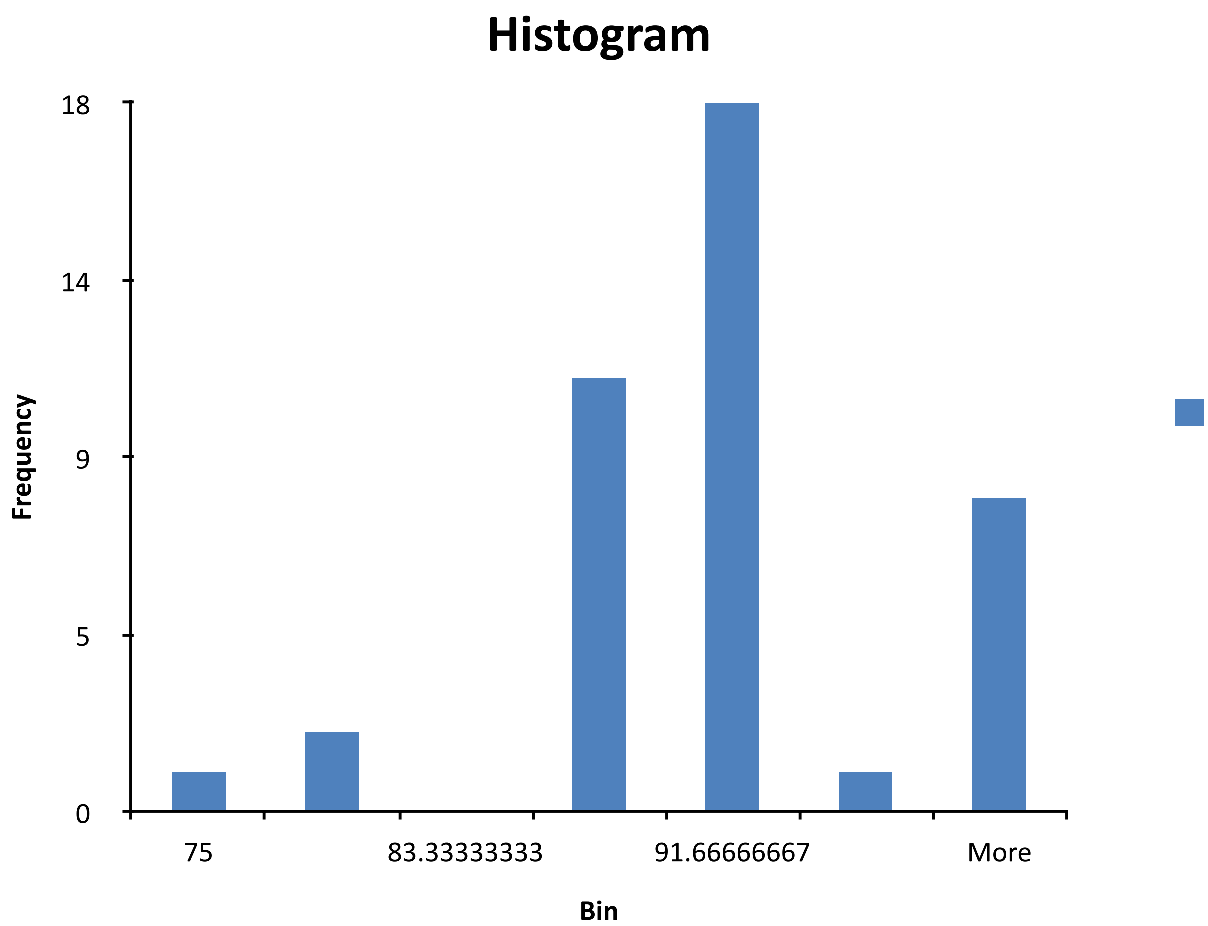
Lecture 9
Two-Sample Test

Fall 2013

Prof. Yao Xie, yao.xie@isye.gatech.edu
H. Milton Stewart School of Industrial Systems & Engineering
Georgia Tech

Computer exam 1



mean 89.90
std 6.02
median 90

Midterm 2

- Cover
- Confidence interval
  - One sided and two sided confidence intervals
- Hypothesis testing
  - Two approaches
    - Fixed significance level
    - p-value

- Can bring a 1-page 1-sided cheat sheet

- Make-up lecture on Friday Nov. 8: tentatively noon-1:20pm in the area in front of my office, Groseclose #339

Outline

- Test difference in the mean
  - Known variance
  - Unknown variance
- Test difference in sample proportion
- Test difference in variance

Motivating Example

- Safety of drinking water (Arizona Republic, May 27, 2001)
- Water sampled from 10 communities in Pheonix
- And 10 communities from rural Arizona
- Arsenic concentration (AC): determines water quality, ranges from 3 ppb to 48 ppb
- Is there a difference in AC

between these two areas?

If the difference is large enough?

Formulate into statistical method

- Answered by statistical methods

Pheonix μ_1

rural Arizona μ_2

- Whether or not there is a difference between in mean AC level, μ_1 and μ_2, in these two areas?
- Equivalent to: test whether μ_1-μ_2 is different from 0?

In general: comparing two populations

- Comparing two population means is often the way used to prove one population is different or better than another

  - Competing Companies / Products
  - Treatment vs. No Treatment
  - New method vs. Old method

Test difference in the mean

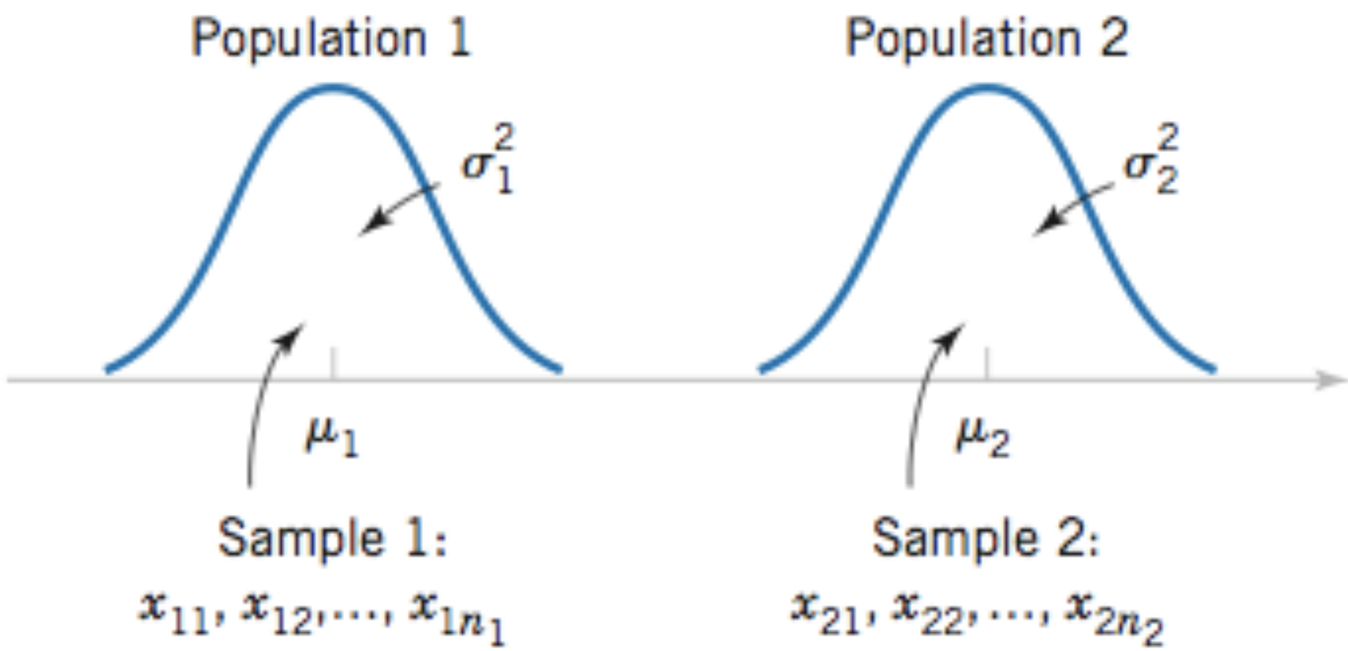


Figure 10-1 Two independent populations.

Test difference in mean, variance known

- Solve the following hypothesis test

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

- Assumptions for two sample inference

1. $X_{11}, X_{12}, \ldots, X_{1n_1}$ is a random sample from population 1.
2. $X_{21}, X_{22}, \ldots, X_{2n_2}$ is a random sample from population 2.
3. The two populations represented by X_1 and X_2 are independent.
4. Both populations are normal.

Test statistics

- A reasonable estimator for $\mu_1 - \mu_2$ is

$$\bar{X}_1 - \bar{X}_2$$

- Under H_0, its mean is Δ
- Its variance is $\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$

- Detection statistic $Z = \dfrac{\bar{X}_1 - \bar{X}_2 - \Delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$

Detection for two sample difference

- For given significance level:
- Reject H_0 when $Z > b$

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - \Delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- And decide threshold b for that given significance level

p-value

- Probability of observing sample difference even more extreme, under H_0

$$P\left(Z > z_0\right) = 1 - \Phi\left(z_0\right)$$

Example: paint drying time

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 $\bar{x}_1 = 121$ minutes and $\bar{x}_2 = 112$ minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using $\alpha = 0.05$?

Solution

- test difference in mean drying time

$$H_0: \mu_1 - \mu_2 = \Delta$$

$$H_1: \mu_1 - \mu_2 > \Delta$$

- $\Delta = 0$

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2$$

- form test statistic

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Fixed significance level approach

- Reject H_0 when $Z = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} > z_\alpha$

- Calculate:

$$\alpha = 0.05 \qquad z_{0.05} = 1.65$$

where $\sigma_1^2 = \sigma_2^2 = (8)^2 = 64$ and $n_1 = n_2 = 10$.

$$\bar{x}_1 = 121 \qquad \bar{x}_2 = 112$$

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{121 - 112}{\sqrt{\dfrac{(8)^2}{10} + \dfrac{(8)^2}{10}}} = 2.52 > 1.65$$

$\Rightarrow$ Reject H_0

Calculate p-value

- Compute p-value:

where $\sigma_1^2 = \sigma_2^2 = (8)^2 = 64$ and $n_1 = n_2 = 10$.

$\bar{x}_1 = 121 \quad \bar{x}_2 = 112$

Value of the statistic from data

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

- p-value:

$$P(Z > z_0) = 1 - \Phi(z_0) = 1 - \Phi(2.52) = 0.0059$$

- Reject H_0 since its value is less than 0.01

Outline

- Test difference in the mean
  - Known variance
  - Unknown variance
- Test difference in sample proportion
- Test difference in variance

Case 2: test difference in mean, variance unknown, true variance equal

- Solve the following hypothesis test

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

- Variances are equal but unknown, so we "***pool***" the samples to estimate the variance

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

- S_1 and S_2 are sample variances

$$\frac{S_p^2(n_1 + n_2 - 2)}{\sigma^2} \sim \chi_{n_1+n_2-2}$$

Use the following as the test statistics

$$\frac{\bar{X}_1 - \bar{X}_2 - \Delta}{S_p\sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}$$

For the following hypothesis test

$$H_0 : \mu_1 - \mu_2 = \Delta$$

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

Reject H_0 when

$$\left|\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{S_p\sqrt{1/n_1 + 1/n_2}}\right| > t_{\alpha/2}$$

Example

$$\alpha = 0.05$$

$$n_1 = 10 \qquad \bar{x}_1 = 28 \qquad S_1^2 = 4$$

$$n_2 = 10 \qquad \bar{x}_2 = 26 \qquad S_2^2 = 5$$

Assume true variance equal

Test Statistic:

$$t = \frac{\bar{x} - \bar{y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_p^{\ 2} = \frac{S_1^2(n_1 - 1) + S_1^2(n_2 - 1)}{n_1 + n_2 - 2} = \frac{4(9) + 5(9)}{18} = 4.5$$

$S_p^2 = 4.5$

Recall degrees of freedom here is
$n + m - 2 = 18$

Threshold: $t_{18,0.025} = 2.101$

$$t = \left| \frac{28\text{-}26}{\sqrt{4.5}\sqrt{1/10 + 1/10}} \right| = 2.11 > 2.101$$

Weakly reject H_0

Calculate p-value

$$p-value = P(|T| > 2.11) = 2P(T > 2.11) = 2 \times 0.0491 = 0.0982$$

Outline

- Test difference in the mean
  - Known variance
  - Unknown variance
- Test difference in sample proportion
- Test difference in variance

Formulation

- Two binomial parameters of interests
- Two independent random samples are taken from 2 populations
- Estimation of sample proportion

$$X \sim \mathrm{Bin}(n_1, p_1), \quad Y \sim \mathrm{Bin}(n_2, p_2) \quad \Rightarrow \hat{p}_1 = \frac{X}{n_1}, \hat{p}_2 = \frac{Y}{n_2}$$

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$

Test statistics

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

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

Estimate the test statistic: $$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Two-sided test

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}}$$

For two-sided test, $H_0 : p_1 = p_2$

$H_1 : p_1 \neq p_2$

reject H_0 when

$$\left| \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right| > z_{\alpha/2}$$

Test statistics and one-sided test

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 < p_2$$

Reject H_0 when

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} < -z_\alpha$$

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 > p_2$$

Reject H_0 when

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > z_\alpha$$

Comparing 2 population proportions: Example

A new drug is being compared to a standard using 200 clinical trials (100 patients for each group). For the new drug, 83 of 100 patients improved. For the standard, 72 of 100 improved. Is the new drug statistically superior?

Standard drug $X \sim \text{Bin}(100, p_1)$

New drug $Y \sim \text{Bin}(100, p_2)$

Fixed significance level approach

$$H_0 : p_1 = p_2$$
$$H_1 : p_1 < p_2$$

$$X_1 = 72, X_2 = 83$$

$$n_1 = n_2 = 100$$

$$\hat{p}_1 = 0.72, \hat{p}_2 = 0.83$$

$$z_{0.05} = 1.65$$

$$\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = -1.7323 < -1.65$$

Reject H_0

p-value

p-value

$$P(Z < -1.7323) = 0.0418$$

Less than $\alpha = 0.05$, reject H_0

Reject H_0, with p-value 0.0418

Outline

- Test difference in the mean
  - Known variance
  - Unknown variance
- Test difference in sample proportion
- Test difference in variance

Test difference in variance

- two independent normal populations
- means and variances of the two normals are unknown
- test whether or not two variances are the same

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

Test based on sample variance ratio

- Test statistics: ratio of two sample variances

$$F = \frac{S_1^2}{S_2^2}$$

- Need to introduce F distribution

Let W and Y be independent chi-square random variables with u and v degrees of freedom, respectively. Then the ratio

$$F = \frac{W/u}{Y/v} \tag{10-28}$$

is said to follow the F distribution with u degrees of freedom in the numerator v degrees of freedom in the denominator. It is usually abbreviated as $F_{u,v}$.

F distribution

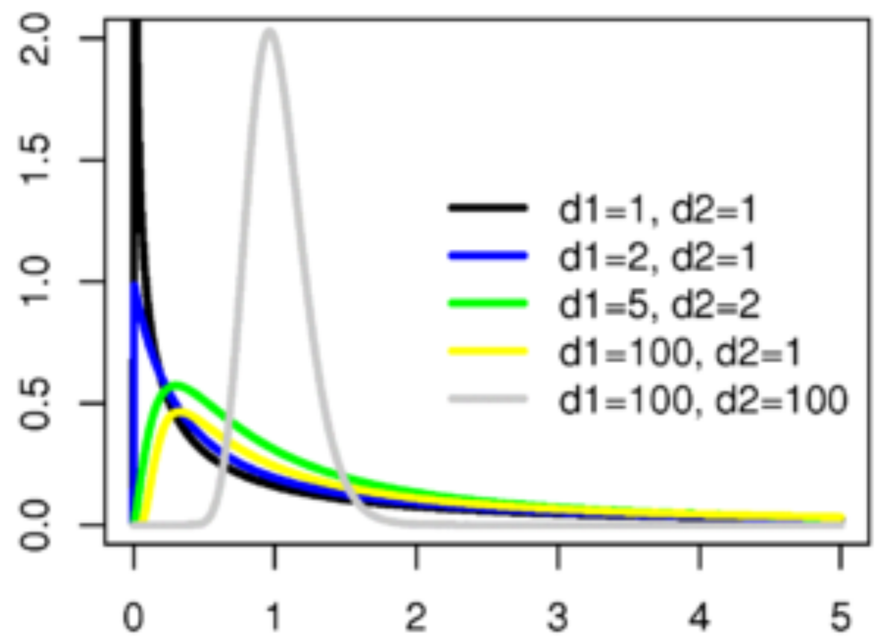
- A continuous distribution

$$f(x; d_1, d_2) = \frac{\sqrt{\frac{(d_1\,x)^{d_1}\,\,d_2^{d_2}}{(d_1\,x+d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)}$$

$$= \frac{1}{\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)} \left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}} x^{\frac{d_1}{2}-1} \left(1+\frac{d_1}{d_2}\,x\right)^{-\frac{d_1+d_2}{2}}$$

- mean = $\dfrac{d_2}{d_2 - 2}$

- we should reject H_0 when the statistic is large



Sample distribution

- Under H_0 the detection statistic

$$F = \frac{S_1^2}{S_2^2} = \frac{\overbrace{\left[(n_1 - 1)S_1^2 / \sigma_1^2\right]}^{\chi^2_{n_1-1}} / (n_1 - 1)}{\underbrace{\left[(n_2 - 1)S_2^2 / \sigma_2^2\right]}_{\chi^2_{n_2-1}} / (n_2 - 1)} \qquad \left(\sigma_1^2 = \sigma_2^2\right)$$

- has F_{n_1-1,n_2-1} distribution

Form of test

Null hypothesis:	$H_0\colon \sigma_1^2 = \sigma_2^2$	
Test statistic:	$F_0 = \dfrac{S_1^2}{S_2^2}$	(10-31)

Alternative Hypotheses	Rejection Criterion
$H_1\colon \sigma_1^2 \neq \sigma_2^2$	$f_0 > f_{\alpha/2, n_1-1, n_2-1}$ or $f_0 < f_{1-\alpha/2, n_1-1, n_2-1}$
$H_1\colon \sigma_1^2 > \sigma_2^2$	$f_0 > f_{\alpha, n_1-1, n_2-1}$
$H_1\colon \sigma_1^2 < \sigma_2^2$	$f_0 < f_{1-\alpha, n_1-1, n_2-1}$

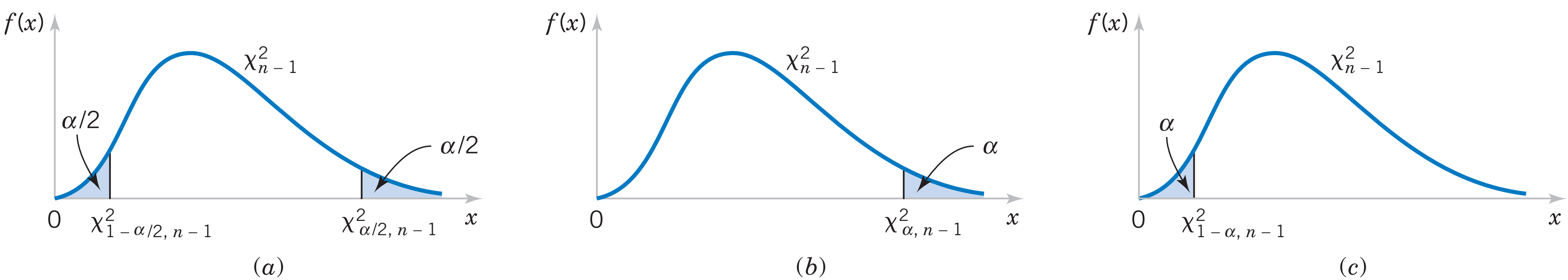


Figure 10-6 The F distribution for the test of $H_0\colon \sigma_1^2 = \sigma_2^2$ with critical region values for (a) $H_1\colon \sigma_1^2 \neq \sigma_2^2$, (b) $H_1\colon \sigma_1^2 > \sigma_2^2$, and (c) $H_1\colon \sigma_1^2 < \sigma_2^2$.

Example: Semiconductor etch variability

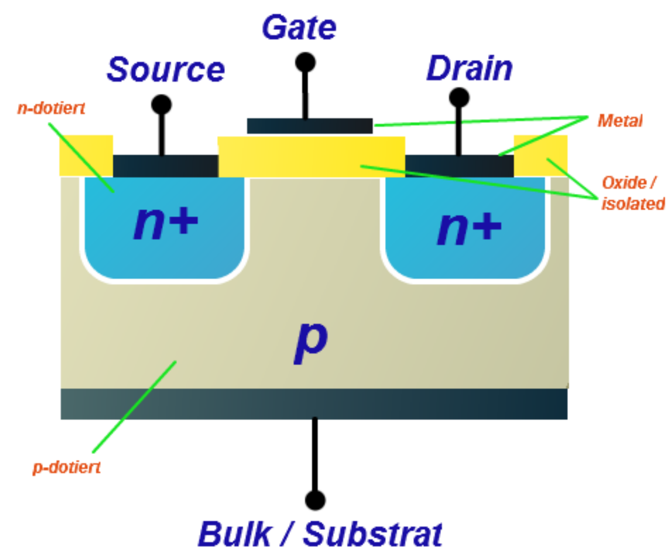
- variability in oxide layer of semiconductor is a critical characteristic of the semiconductor
- two kind of semiconductors, sample standard deviation

$$s_1 = 1.96$$

$$s_2 = 2.13$$

$$n_1 = n_2 = 16$$

$$\alpha = 0.05$$



- test: whether or not their variances are the same

1. **Parameter of interest:** The parameter 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. **Null hypothesis:** H_0: $\sigma_1^2 = \sigma_2^2$
3. **Alternative hypothesis:** H_1: $\sigma_1^2 \neq \sigma_2^2$
4. **Test statistic:** The test statistic is given by

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

6. **Reject $\boldsymbol{H_0}$ if :** Because $n_1 = n_2 = 16$ and $\alpha = 0.05$, we will reject H_0: $\sigma_1^2 = \sigma_2^2$ if $f_0 > f_{0.025,15,15} = 2.86$ or if $f_0 < f_{0.975,15,15} = 1/f_{0.025,15,15} = 1/2.86 = 0.35$.

7. Computations: Because $s_1^2 = (1.96)^2 = 3.84$ 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: Because $f_{0.975,15,15} = 0.35 < 0.85 < f_{0.025,15,15} = 2.86$, we cannot reject the null hypothesis H_0: $\sigma_1^2 = \sigma_2^2$ at the 0.05 level of significance.

p-value

- Observe “test statistic” more extreme than what we got

Alternative Hypotheses	p-value
$H_1: \sigma_1^2 \neq \sigma_2^2$	$2P(F > f_0)$ or $2P(F < f_0)$, depends on f_0 fall in upper or lower tail
$H_1: \sigma_1^2 > \sigma_2^2$	$P(F > f_0)$
$H_1: \sigma_1^2 < \sigma_2^2$	$P(F < f_0)$

- calculate using R command

```
p <- pf(x,d1,d2)
```

Back to semiconductor example

computed value of the test statistic in this example is $f_0 = 0.85$

$$P(F_{15,15} \leq 0.85) = 0.3785$$

p-value $2(0.3785) = 0.7570.$

- calculate using R command

```
> p <- pf(0.85,15,15)
> p
[1] 0.3785271
>
```