Chapter 7

Hypotheses Testing based on Normal Distribution

Overview

- Hypotheses testing based on Normal distribution
- Types I and II Error
- Level of significance
- Hypotheses testing concerning mean
- Critical value approach and p-value approach
- Hypotheses testing concerning difference between two means
- Hypotheses testing concerning variance

7.1 Null and Alternative Hypotheses

7.1.1 Statistical Hypothesis

- A **statistical hypothesis** is an assertion or conjecture concerning one or more populations.
- We shall use the terms accept and reject frequently throughout this chapter.

Null and Alternative Hypotheses (Continued)

7.1.1 Statistical Hypothesis (Continued)

- It is important to understand that the rejection of a hypothesis is to conclude that it is false, while the acceptance of a hypothesis merely implies that we have insufficient evidence to believe otherwise.
- Because of this terminology, the statistician or experimenter will often choose to state the hypothesis in a form that hopefully will be rejected.

Null and Alternative Hypotheses (Continued)

Null hypothesis:

- Hypothesis that we formulate with the hope of rejecting, denoted by H_0.
- A null hypothesis concerning a population parameter will always be stated so as to specify an exact value of the parameter.

Alternative hypothesis:

- The rejection of H_0 leads to the acceptance of an alternative hypothesis, denoted by H_1.
- It allows for the possibility of several values.

Example 1

- We may wish to determine whether the mean IQ of the pupils of a certain school is different from 100.
- Then we have

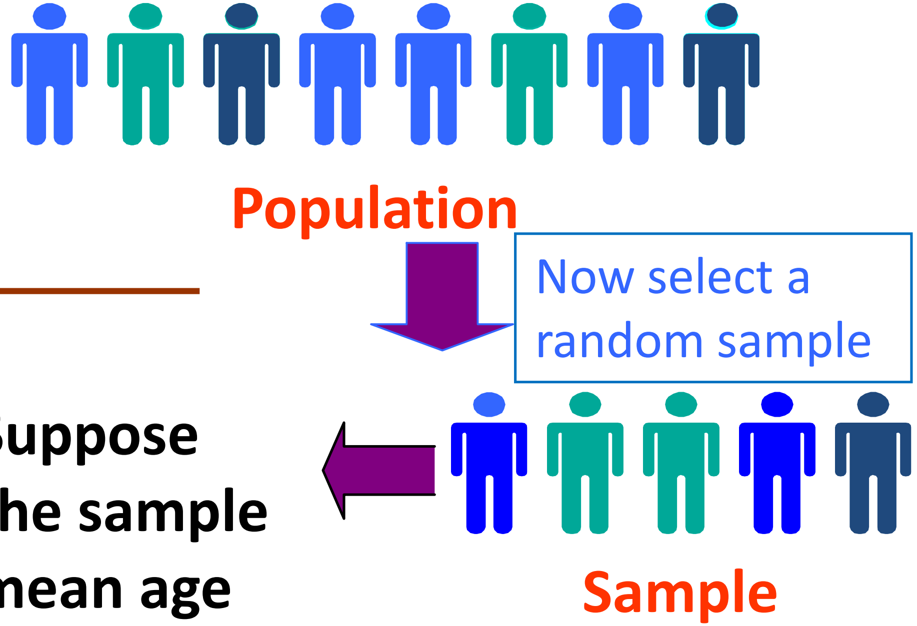
  $\mathrm{H}_0: \mu = 100$ against $\mathrm{H}_1: \mu \neq 100.$
- This is called a two-sided (or two tailed) test.
- We may like to test whether the mean IQ of the pupils is greater than 100 (or less than 100). This is called a one-tailed (or one-sided) test.
- That is,

  $\mathrm{H}_0: \mu = 100$ against $\mathrm{H}_1: \mu > 100$ or

  $\mathrm{H}_0: \mu = 100$ against $\mathrm{H}_1: \mu < 100$.

Example 1

Claim: the population mean age is 50.
(Null hypothesis: $H_0: \mu = 50$)



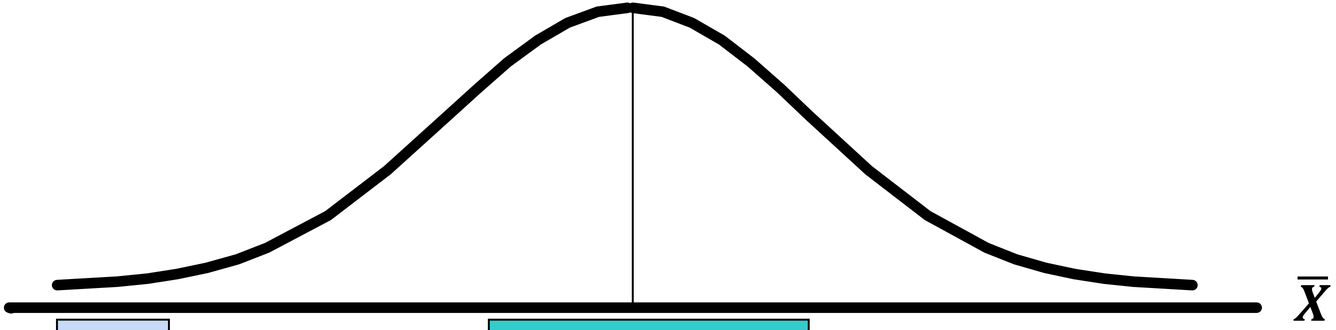
Is $\bar{X} = 20$ likely if $\mu = 50$?

Suppose the sample mean age is 20: $\bar{X} = 20$

If not likely, REJECT Null Hypothesis

Reason for Rejecting H_0

Sampling Distribution of $\bar{X}$



20

$\mu = 50$

If H_0 is true

If it is unlikely that we would get a sample mean of this value ...

... if in fact this were the population mean...

... then we reject the null hypothesis that $\mu = 50$.

7.1.2 Types of Error

- Two types of errors in the hypothesis testing:

	State of Nature	
Decision	H_0 is true	H_0 is false
Reject H_0	**Type I error** Pr(Reject H_0 given that H_0 is true) = α	Correct decision Pr(Reject H_0 given that H_0 is false) = $1 - \beta$
Do not reject H_0	Correct decision Pr(Do not reject H_0 given that H_0 is true) = $1 - \alpha$	**Type II error** Pr(Do not reject H_0 given that H_0 is false) = β

Types I and II Error

Type I error

- Rejection of H_0 when H_0 is true is called a type I error.
- It is considered as a serious type of error

Type II Error

- Not rejecting H_0 when H_0 is false is called a type II error.

Types I and II Error (Continued)

- α = level of significance

  = Pr(type I error)

  = Pr(reject H_0 when H_0 is true)

  = Pr(reject H_0 | H_0).
- α is set by the researcher in advance
- α is usually set at 5% or 1%

Types I and II Error (Continued)

- β = Pr(type II error)

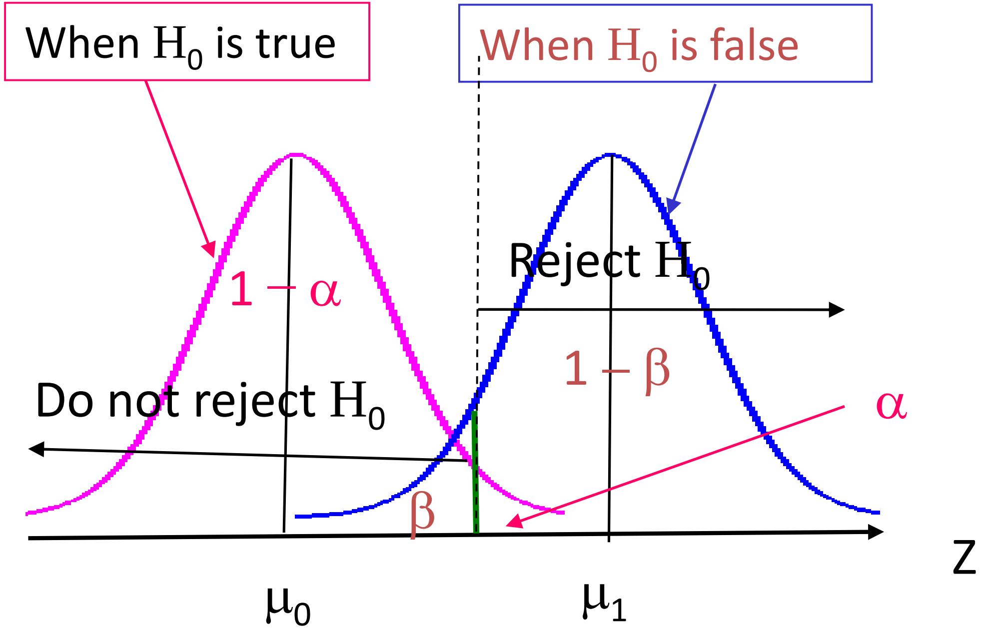
  = Pr(do not reject H_0 when H_0 is false)

  = Pr(do not reject $H_0 \mid H_1$).

- $1 - \beta$ = Power of a test = Pr(reject $H_0 \mid H_1$)

Types I and II Error (Continued)

Test $H_0: \mu = \mu_0$ against $H_1: \mu = \mu_1$



7.1.3 Acceptance and Rejection Regions

- To test a hypothesis about a population parameter, we first select a suitable test statistic for the parameter under the hypothesis.
- Once the significance level, α, is given, a decision rule can be found such that it divides the set of all possible values of the test statistic into two regions,
- one being the **rejection region** (or critical region) and the other the **acceptance region**.

Acceptance and Rejection Regions (Continued)

- Once a sample is taken, the value of the test statistic is obtained.
- If the test statistic assumes a value in the rejection region, the null hypothesis is rejected; otherwise it is not rejected.
- The value that separates the rejection and acceptance regions is called the **critical value**.

Level of Significance and the Rejection Region

Level of significance = α

$H_0: \mu = 3$

$H_1: \mu \neq 3$

Two-tail test

$H_0: \mu = 3$

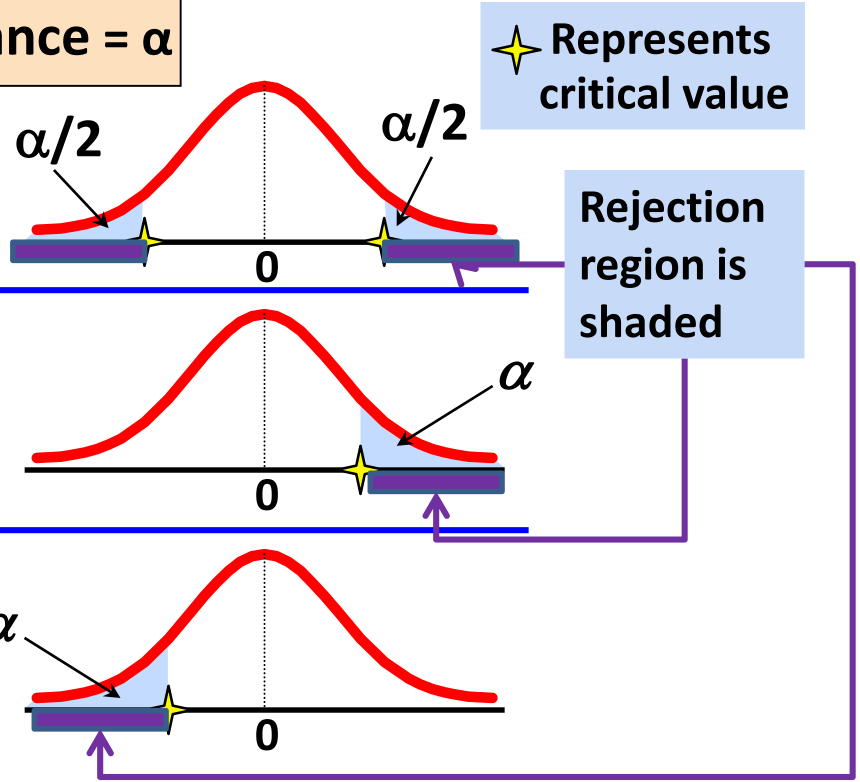
$H_1: \mu > 3$

Upper-tail test

$H_0: \mu = 3$

$H_1: \mu < 3$

Lower-tail test



Example 1

- A certain type of cold vaccine is known to be only 25% effective after a period of 2 years.
- In order to determine if a new and somewhat more expensive vaccine is superior in providing protection against the same virus for a longer period of time.
- 20 people are chosen at random and inoculated with the new vaccine.
- If more than 8 of those receiving the new vaccine surpass the 2-year period without contracting the virus, the new vaccine will be considered superior to the one presently in use.

Example 1 (Continued)

- This is equivalent to testing the hypothesis that the binomial parameter for the probability of a success on a given trial is $p = 1/4$ against the alternative that $p > 1/4$. Or
- $H_0: p = 1/4$ against $H_1: p > 1/4$.

$$X \quad \underbrace{0\ 1\ 2\ 3\ 4 \cdots 7\ 8}_{\text{Acceptance Region}}\ \underbrace{9\ 10\ 11 \cdots 19\ 20}_{\text{Rejection Region}}$$

where X is the number of individuals who remain free of the virus for at least 2 years

Example 1 (Continued)

- The above decision rule has the level of significance given by

$$\begin{aligned}\alpha &= \Pr(\text{Type I error}) \\ &= \Pr(\text{Reject } H_0 \mid H_0) \\ &= \Pr(X > 8 \text{ when } p = 1/4) \\ &= \sum_{i=9}^{20} \binom{20}{i} \left(\frac{1}{4}\right)^{i} \left(\frac{3}{4}\right)^{20-i} \\ &= 0.0409.\end{aligned}$$

Example 1 (Continued)

- The probability of committing a type II error, denoted by β, is impossible to compute unless we have a specific alternative hypothesis.
- Consider testing

  H_0: $p = 1/4$ against H_1: $p = 1/2$ (Note $1/2 > 1/4$).

Example 1 (Continued)

- Then

$$\beta = \Pr(\text{Type II error}) = \Pr(\text{Accept } H_0 \mid H_1)$$

$$= \Pr(X \le 8 \text{ when } p = 1/2)$$

$$= \sum_{i=0}^{8} \binom{20}{i} \left(\frac{1}{2}\right)^{i} \left(\frac{1}{2}\right)^{20-i} = 1 - \sum_{i=9}^{20} \binom{20}{i} \left(\frac{1}{2}\right)^{20}$$

$$= 1 - 0.7483 = 0.2517.$$

7.2 Hypotheses Testing Concerning Mean

7.2.1 Hypo. Testing on Mean with Known Variance

Consider the problem of testing the hypothesis concerning the mean, μ, of a population with

1. **Variance, $\boldsymbol{\sigma^2}$, known and**
2. **Underlying distribution is normal or $\boldsymbol{n}$ is sufficiently large (say $\boldsymbol{n > 30}$)**

   Refer to Section 6.3.1

7.2.1.1 Two sided-test

- Test $\mathrm{H}_0: \mu = \mu_0$ against $\mathrm{H}_1: \mu \neq \mu_0$.
- When the population is normal or the sample size is large (then by the Central Limit Theorem), we can expect that

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

- Hence under $\mathrm{H}_0: \mu = \mu_0$, we have

$$\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right).$$

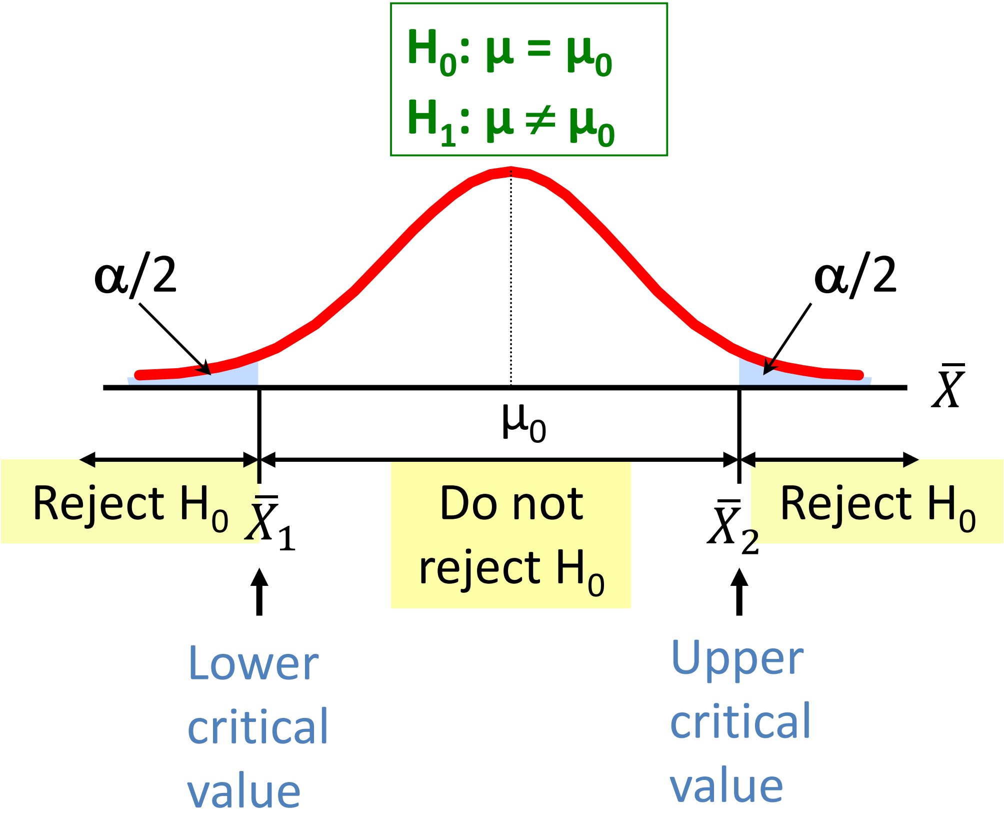
Two sided Test (Continued)

Critical Value Approach

- By using a significance level of α, it is possible to find two critical values $\bar{x}_1$ and $\bar{x}_2$ such that
- the interval $\bar{x}_1 < \bar{X} < \bar{x}_2$ defines the acceptance region and
- the two tails of the distribution, $\bar{X} < \bar{x}_1$ and $\bar{X} > \bar{x}_2$ constitute the critical (or rejection) region.

Two-sided Test (Continued)

- There are two cutoff values (critical values), defining the regions of rejection



Finding critical values

The critical region can be given in terms of z values by means of the transformation

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Note: μ_0 is the value of μ under H_0.

Finding critical values (Continued)

Therefore

$$\Pr\left(-z_{\alpha/2} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

or

$$\Pr\left(\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Hence $\bar{x}_1 = \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and $\bar{x}_2 = \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

Hypothesis testing process

- From the population we select a random sample of size n and compute the sample mean.
- If $\bar{X}$ falls in the acceptance region $\bar{x}_1 < \bar{X} < \bar{x}_2$, we conclude that $\mu = \mu_0$; otherwise we reject H_0 and accept the H_1: $\mu \neq \mu_0$.
- Since $Z = (\bar{X} - \mu_0)/(\sigma/\sqrt{n})$, therefore $\bar{x}_1 < \bar{X} < \bar{x}_2$ is equivalent to $-z_{\alpha/2} < Z < z_{\alpha/2}$.
- The critical region is usually stated in terms of Z rather then $\bar{X}$.

Example 1

- The director of a factory wants to determine if a new machine A is producing cloths with a breaking strength of 35 kg with a standard deviation of 1.5 kg.
- A random sample of 49 pieces of cloths is tested and found to have a mean breaking strength of 34.5 kg.
- Is there evidence that the machine is not meeting the specifications for mean breaking strength?
- (Use $\alpha = 0.05$)

Solution to Example 1

Step 1

- Let μ be the mean breaking strength of cloths manufactured by the new machine.
- Test $\mathrm{H}_0: \mu = 35$ kg vs $\mathrm{H}_1: \mu \neq 35$ kg. (Why?)

Step 2

- Set $\alpha = 0.05$.

Solution to Example 1 (Continued)

Step 3

- Since σ is known, the test statistic

$$Z = \frac{(\bar{X} - \mu_0)}{\sigma/\sqrt{n}}$$

  is used.
- $z_{\alpha/2} = z_{0.025} = 1.96.$
- Critical region $z < -1.96$ or $z > 1.96$, where

$$z = \frac{(\bar{x} - \mu_0)}{\sigma/\sqrt{n}}$$

Solution to Example 1

Step 4

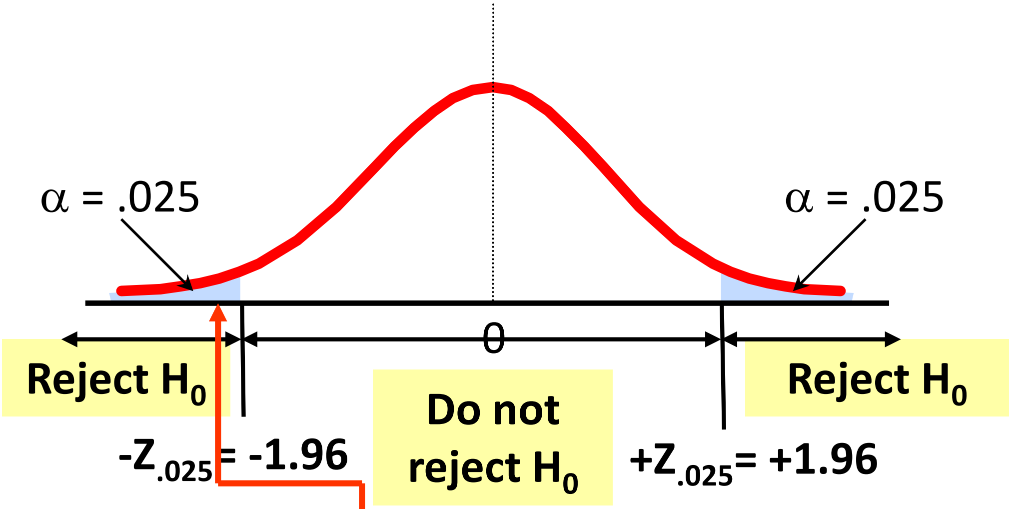
- Computations: $\bar{x} = 34.5$ kg, $n = 49$, and hence

$$z = \frac{34.5 - 35}{1.5/\sqrt{49}} = -2.3333$$

Step 5

- Conclusion: Since the observed z value = –2.3333 falls inside the critical region (i.e. $z < z_{0.025} = -1.96$), hence H_0: $\mu = 35$ kg is rejected at the 5% level of significance.

Solution to Example 1 (Continued)



Reject H_0 if $Z < -1.96$ or $Z > 1.96$; otherwise do not reject H_0

Here, $Z = -2.3333 < -1.96$, so the test statistic is in the rejection region

Relationship between two-sided test and confidence interval

- The two-sided test procedure just described is equivalent to finding a $(1-\alpha)100\%$ confidence interval for μ
- H_0 is accepted if the confidence interval covers μ_0.
- If the C.I. does not cover μ_0, we reject $\mu = \mu_0$ in favour of the alternative H_1: $\mu \neq \mu_0$ since

$$\Pr\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

$$\Leftrightarrow \quad \Pr\left(\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Example 1 (Continued)

- For $\bar{x} = 34.5$, $\sigma = 1.5$ and $n = 49$, the 95% confidence interval is:

$$34.5 - (1.96)\frac{1.5}{\sqrt{49}} < \mu < 34.5 + (1.96)\frac{1.5}{\sqrt{49}}$$

$$34.08 \leq \mu \leq 34.92$$

- Since this interval does not contain the hypothesized mean, μ_0 (= 35), we reject the null hypothesis at $\alpha = 0.05$.

p-Value Approach to Testing

- p-value: Probability of obtaining a test statistic more extreme ($\leq$ or $\geq$) than the observed sample value given H_0 is true
  - Also called **observed level of significance**

p-Value Approach to Testing (Continued)

- Convert a sample statistic (e.g., $\bar{X}$) to a test statistic (e.g., Z statistic)
- Obtain the p-value
- Compare the p-value with $\boldsymbol{\alpha}$
  - If p-value $< \alpha$, reject H_0
  - If p-value $\geq \alpha$, do not reject H_0

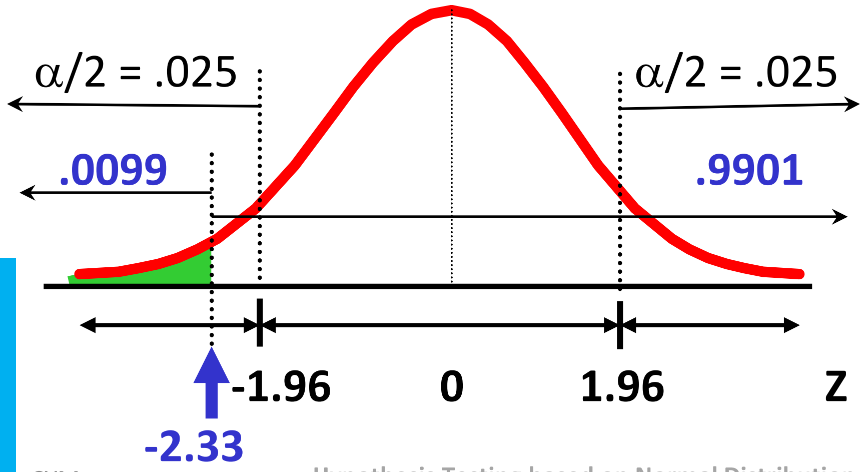
Example 1 (Continued)

- How likely is it to see a sample mean of 34.5 (or something further from the mean, in either direction) if the true mean is 35? ($\sigma = 1.5$ and $n = 49$)

$\bar{X}$ = 34.5 is translated to a Z score of Z = -2.33

Pr(*Z* < -2.33) = 0.0099
Pr(*Z* > -2.33) = 0.9901

***p*-value**
= 2 min{Pr(*Z* < -2.33), Pr(*Z* >-2.33) }
= 2(0.0099) = .0198



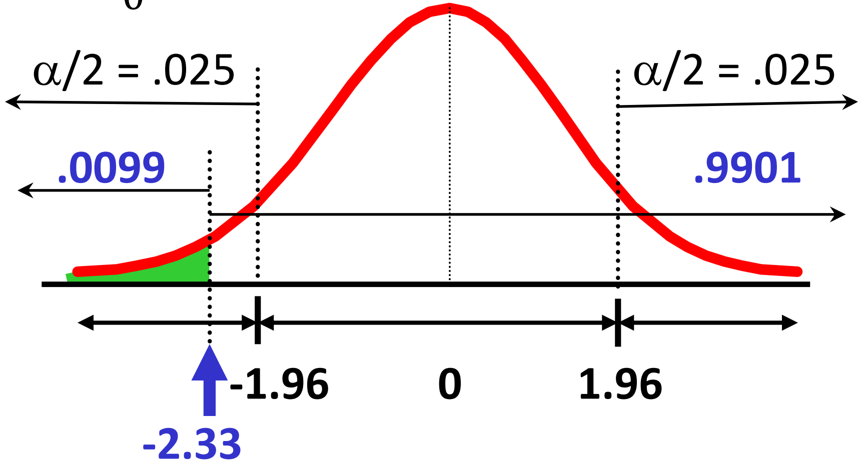
p-Value Approach to Testing (Continued)

- Compare the *p*-value with α
  - If p-value $< \alpha$, reject H_0
  - If p-value $\geq \alpha$, do not reject H_0

Here: *p*-value = .0198

α = .05

Since .0198 < .05, we reject the null hypothesis



7.2.1.2 One sided test

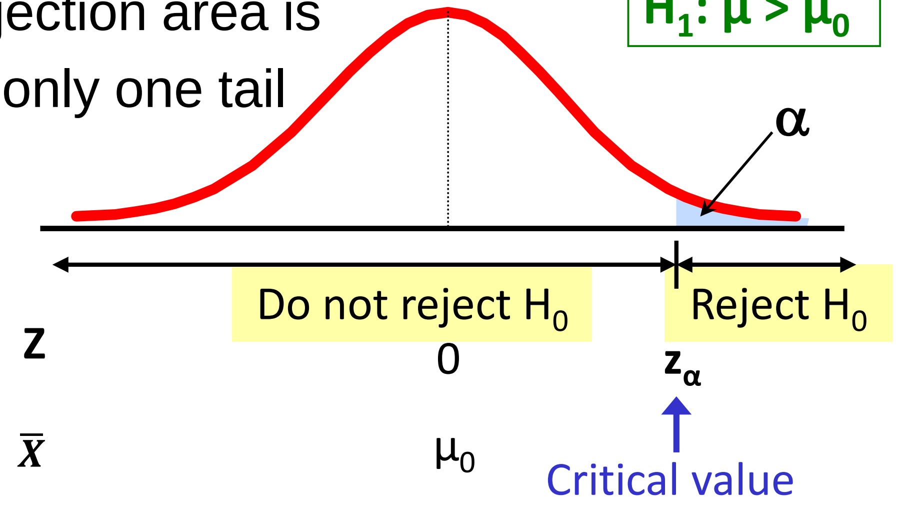
(a) Test $H_0: \mu = \mu_0$ against
$H_1: \mu > \mu_0$.

- Let $Z = \dfrac{(\bar{X} - \mu_0)}{\sigma/\sqrt{n}}$.
- Then H_0 is rejected if the observed values of Z, say z, is greater than z_α.

There is only one critical value, since the rejection area is in only one tail

$H_0: \mu = \mu_0$

$H_1: \mu > \mu_0$

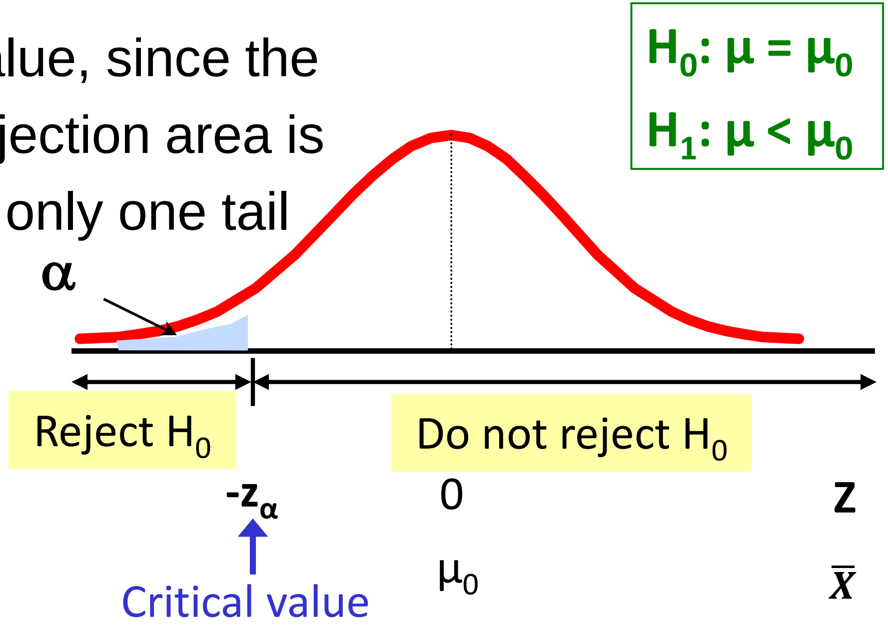


7.2.1.2 One sided test

(b) Test $H_0: \mu = \mu_0$ against $H_1: \mu < \mu_0$.

- Let $Z = \frac{(\bar{X} - \mu_0)}{\sigma/\sqrt{n}}$.
- Then H_0 is rejected if the observed values of Z, say z, is less than $-z_\alpha$.

There is only one critical value, since the rejection area is in only one tail



Example 2

- A manufacturer of sports equipment has developed a new synthetic fishing line that he claims has a mean breaking strength better than the market average strength of 8 kilograms.
- Suppose the breaking strength of this type of fishing lines has a standard deviation of 0.5 kg.
- A random sample of 50 lines is tested and found to have a mean breaking strength of 8.2 kg.
- Test the manufacturer’s claim.
- Use a 0.01 level of significance.

Solution to Example 2

Step 1

- Let μ be the mean breaking strength of the new type of fishing lines.
- Test $\mathrm{H}_0: \mu = 8$ against $\mathrm{H}_1: \mu > 8$. (Why?)

Step 2

- Set $\alpha = 0.01$.

Solution to Example 2 (Continued)

Step 3

- Since σ is known, the test statistic

$$Z = \frac{(\bar{X} - 8)}{0.5/\sqrt{50}}$$

  is used.

- $z_\alpha = z_{0.01} = 2.326.$
- Critical region $z > 2.326$, where

$$z = \frac{(\bar{x} - 8)}{0.5/\sqrt{50}}$$

Solution to Example 2 (Continued)

Step 4

- Computations: $\bar{x} = 8.2$, hence

$$z = \frac{8.2 - 8}{0.5/\sqrt{50}} = 2.828.$$

- p-value $= \Pr(Z > 2.828) \approx 0.00233.$

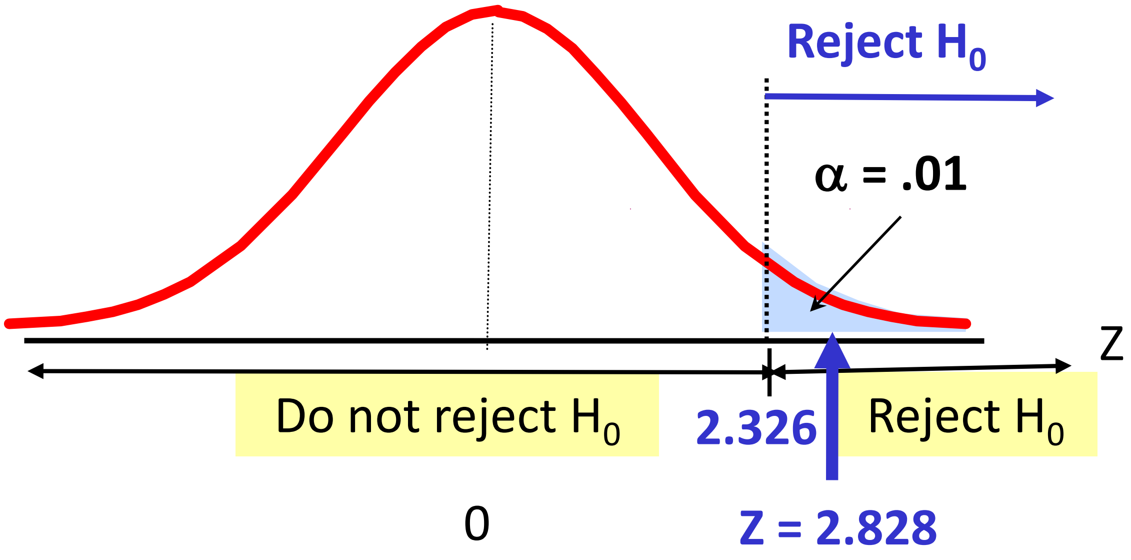
Solution to Example 2 (Continued)

Step 5

- Conclusion: Since the observed z value $= 2.828$ falls inside the critical region (i.e. $z > z_{0.01} = 2.326$), hence H_0: $\mu = 8$ kg is rejected at the 1% level of significance.
- Conclusion based on p-value: Since p-value ≈ 0.00233 is less than 0.01, hence H_0 is rejected at the 1% level of significance.

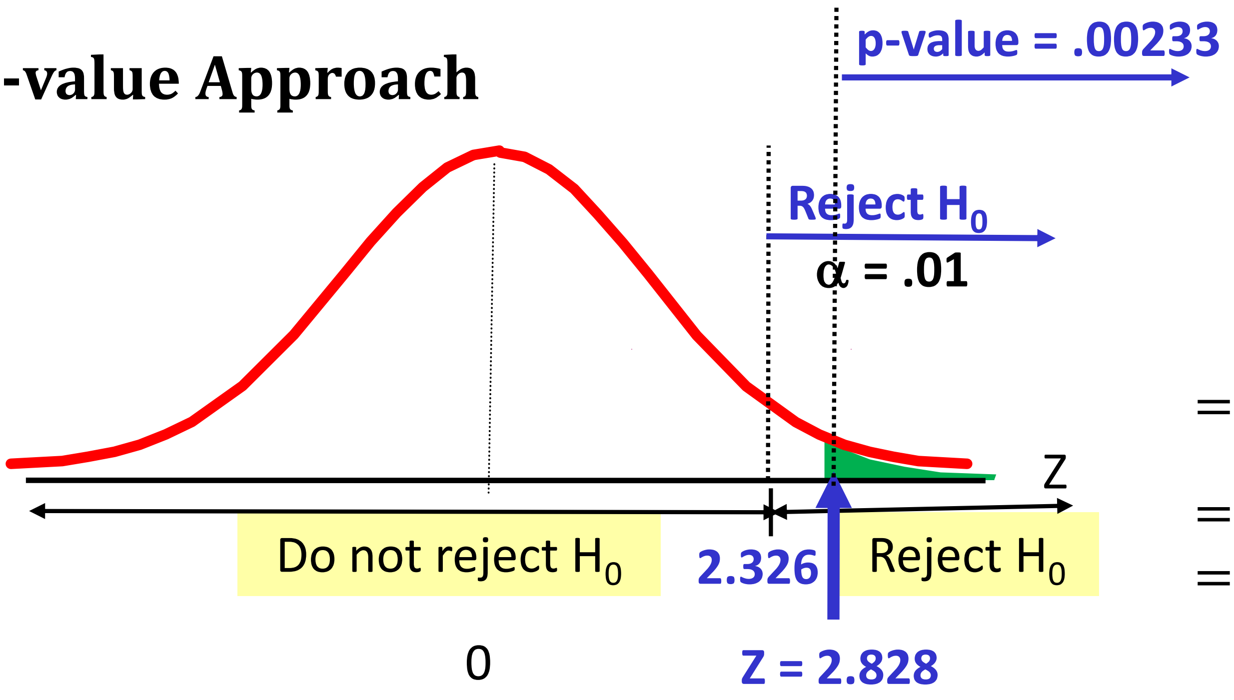
Solution to Example 2 (Continued)

Critical Value Approach



Solution to Example 2 (Continued)

p-value Approach



$$\Pr(\bar{X} > 8.2)$$
$$= \Pr\left(Z \geq \frac{8.2 - 8}{0.5/\sqrt{50}}\right)$$
$$= \Pr(Z \geq 2.828)$$
$$= 0.00233$$

Reject H_0 since *p*-value = .00233 < α = .01

7.2.2 Hypothesis Testing on Mean with Variance Unknown

Consider the problem of testing the hypothesis concerning the mean, μ, of a population with

1. **Variance unknown and**
2. **Underlying distribution is normal**

   Refer to Section 6.3.2

Test for mean with unknown variance

(1) Two sided test

- Test $\mathrm{H}_0: \mu = \mu_0$ against $\mathrm{H}_1: \mu \neq \mu_0$.
- Let

$$T = \frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}$$

where S^2 is the sample variance.

- Then H_0 is rejected if the observed value of T, say t, $> t_{n-1;\alpha/2}$ or $< -t_{n-1;\alpha/2}$.

Test for mean with unknown variance (Continued)

(2) One sided test

- Test H_0: $\mu = \mu_0$ against H_1: $\mu > \mu_0$.
- Then H_0 is rejected if $t > t_{n-1;\alpha}$.
- Test H_0: $\mu = \mu_0$ against H_1: $\mu < \mu_0$.
- Then H_0 is rejected if $t < -t_{n-1;\alpha}$.

Example 3

- The average length of time for students to register for summer classes at a certain college has been 50 minutes.
- A new registration procedure is being tried.
- A random sample of 12 students had an average registration time of 42 minutes with a standard deviation of 11.9 minutes under the new system.
- Test the hypothesis that the population mean is now less than 50, using a level of significance of 0.05.
- Assume the population of times to be normal.

Solution to Example 3

Step 1

- Let μ be the mean registration time.
- Test $\mathrm{H}_0: \mu = 50$ against $\mathrm{H}_1: \mu < 50$. (Why?)

Step 2

- Set $\alpha = 0.05$.

Solution to Example 3 (Continued)

Step 3

- Since σ is unknown, the test statistic

$$T = \frac{(\bar{X} - \mu_0)}{S/\sqrt{n}}$$

is used.

- $n = 12$ implies that $t_{11;0.05} = 1.796$
- **Critical region** $t < -1.796$, where

$$t = \frac{(\bar{x} - \mu_0)}{s/\sqrt{n}}$$

Solution to Example 3 (Continued)

Step 4

- **Computations:** $\bar{x} = 42$, $s = 11.9$, $n = 12$, and hence
$$t = \frac{42 - 50}{11.9/\sqrt{12}} = -2.329$$
- p-value $= \Pr(T < -2.329) = 0.0199$.

  [or p-value is between 0.025 and 0.01 since 2.329 is between $t_{11;0.025} = 2.201$ and $t_{11;0.01} = 2.718$ if statistical table is used.]

Solution to Example 2 (Continued)

Step 5

- **Conclusion:** Since the observed $t = -2.329$ falls inside the critical region (i.e. $t < t_{0.05} = -1.796$), hence H_0: $\mu = 50$ minutes is rejected at the 5% level of significance and we conclude that the true mean is likely to be less than 50 minutes.
- **Conclusion based on *p*-value:** Since *p*-value = 0.0199 is less than 0.05, hence H_0 is rejected at the 5% level of significance and we conclude that the true mean is likely to be less than 50 minutes.

7.3 Hypotheses Testing Concerning Difference Between Two Means

7.3.1 Known Variances

1. Variances σ_1^2 and σ_2^2 are known and

2. Underlying distribution is normal or both n_1 and n_2 are sufficiently large

(say $n_1 \geq 30, n_2 \geq 30$)

Refer to Section 6.4.1

Example 1

- Analysis of a random sample consisting of $n_1 = 20$ specimens of cold-rolled steel to determine yield strengths resulted in a sample average strength of $\bar{x}_1 = 29.8$ ksi.
- A second random sample of $n_2 = 25$ two-side galvanized steel specimens gave a sample average strength of $\overline{x}_2 = 34.7$ ksi.
- Assuming that the two yield strength distributions are normal with $\sigma_1 = 4.0$ and $\sigma_2 = 5.0$,
- does the data indicate that the corresponding true average yield strengths μ_1 and μ_2 are different?
- Use $\alpha = 0.01$.

Solution to Example 1

Step 1

- Let μ_1 and μ_2 be the mean strength of cold-rolled steel and two-side galvanized steel respectively.
- Test $\mathrm{H}_0: \mu_1 - \mu_2 = 0$ against $\mathrm{H}_1: \mu_1 - \mu_2 \neq 0$.

Step 2

- Set $\alpha = 0.01$.

Solution to Example 1 (Continued)

Step 3

- Since σ_1^2 and σ_2^2 are known, therefore the test statistic

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

  is used.

- $\alpha = 0.01$ implies $z_{\alpha/2} = z_{0.005} = 2.5728$.

Solution to Example 1 (Continued)

Step 3 (Continued)

- Critical region: $z < -2.5758$ or $z > 2.5758$, where

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

Solution to Example 1 (Continued)

Step 4

- **Computations:** $\bar{x}_1 = 29.8, \bar{x}_2 = 34.7, \sigma_1^2 = 16, \sigma_2^2 = 25,$ $n_1 = 20$ and $n_2 = 25$, so

$$z = \frac{[(29.8 - 34.7) - 0]}{\sqrt{16/20 + 25/25}} = -3.652.$$

- ***p*-value** $= 2 \times \min\{\Pr(Z > -3.652), \Pr(Z < -3.652)\} = 2(0.00013) = 0.00026.$

Solution to Example 1 (Continued)

Step 5

- Conclusion: Since $z = -3.652$ falls inside the critical region, hence $\mathrm{H_0}$: $\mu_1 = \mu_2$ is rejected at the 1% level of significance and conclude that the sample data strongly suggest that the true average yield strength for cold-rolled steel differs from that for galvanized steel.
- Conclusion based on p-value: Since p-value = 0.00026 is less than the level of significance 0.01, hence $\mathrm{H_0}$ is rejected at the 1% level of significance.

7.3.2 Large Sample Testing with Unknown Variances

1. **Variances $\boldsymbol{\sigma_1^2}$ and $\boldsymbol{\sigma_2^2}$ are unknown and**
2. **both $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ are sufficiently large (say $\boldsymbol{n_1 \geq 30, n_2 \geq 30}$)**

   Refer to Section 6.4.2

Example 2

- In selecting a sulfur concrete for roadway construction in regions that experience heavy frost,
- it is important that the chosen concrete have a low value of thermal conductivity in order to minimize subsequent damage due to changing temperatures.
- Suppose two types of concrete, a graded aggregate and a no-fines aggregate, are being considered for a certain road.

Example 2 (Continued)

- The following table summarizes data from an experiment carried out to compare the two types of concrete.

Type	Sample size	Sample average conductivity	Sample s.d.
Graded	35	0.497	0.187
No-fines	35	0.359	0.158

- Does this information suggest that the true conductivity for the graded concrete exceeds that for the no-fines concrete?
- Use $\alpha = 0.01$.

Solution to Example 2

Step 1

- Let μ_1 and μ_2 be the mean conductivity of graded and no-fines concretes respectively.
- Test $\mathrm{H}_0: \mu_1 - \mu_2 = 0$ against $\mathrm{H}_1: \mu_1 - \mu_2 > 0$.

Step 2

- Set $\alpha = 0.01$.

Solution to Example 2 (Continued)

Step 3

- Since σ_1^2 and σ_2^2 are unknown and the sample sizes are large, therefore the test statistic

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$$

  is used.
- $\alpha = 0.01$ implies $z_\alpha = z_{0.01} = 2.3263$.

Solution to Example 2 (Continued)

Step 3 (Continued)

- **Critical region:** $z > 2.3263$, where

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Solution to Example 2 (Continued)

Step 4

- **Computations:** $\bar{x}_1 = 0.497, \bar{x}_2 = 0.359, s_1^2 = 0.187^2, s_2^2 = 0.158^2, n_1 = n_2 = 35$, so

$$z = \frac{[(0.497 - 0.359) - 0]}{\sqrt{0.187^2/35 + 0.158^2/35}} = 3.335.$$

- p-value $= \Pr(Z > 3.335) = 0.00043.$

Solution to Example 2 (Continued)

Step 5

- **Conclusion:** Since $z = 3.335$ falls inside the critical region, hence H_0: $\mu_1 = \mu_2$ is rejected at the 1% level of significance and conclude that the sample data argue strongly that the true average thermal conductivity for the graded concrete does exceed that for the no-fines concrete.
- **Conclusion based on *p*-value**: Since *p*-value = 0.00043 is less than the level of significance 0.01, hence H_0 is rejected at the 1% level of significance.

7.3.3 Unknown but Equal Variances

1. **σ_1^2 and σ_2^2 are unknown but equal and**
2. **the populations are normal**
3. **Small sample sizes (say $n_1 \leq 30, n_2 \leq 30$)**

   Refer to Section 6.4.3

Example 3

- A course in mathematics is taught to 12 students by the conventional classroom procedure.
- A second group of 10 students was given the same course by means of programmed materials.
- At the end of the semester the same examination was given to each group.
- The 12 students meeting in the classroom made an average grade of 85 with a standard deviation of 4,

Example 3 (Continued)

- while the 10 students using programmed materials made an average of 81 with a standard deviation of 5.
- Test the hypothesis that the two methods of learning are equal using a 0.10 level of significance.
- Assume the populations to be approximately normal with equal variances.

Solution to Example 3

Step 1

- Let μ_1 and μ_2 be the average grades students taking this course by the classroom and programmed presentations, respectively.
- Test $\mathrm{H}_0: \mu_1 - \mu_2 = 0$ against $\mathrm{H}_1: \mu_1 - \mu_2 \neq 0$.

Step 2

- Set $\alpha = 0.1$.

Solution to Example 3 (Continued)

Step 3

- $n_1 = 12$ and $n_2 = 10$ implies $t_{n_1+n_2-2;\alpha} = t_{20;0.05} = 1.725.$
- **Critical region :** $t < -1.725$ or $t > 1.725$, where

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

with

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

Solution to Example 3 (Continued)

Step 4

- **Computations:**

$\bar{x}_1 = 85$ and $\bar{x}_2 = 81$,

$s_1^2 = 16, s_2^2 = 25$,

$n_1 = 12$ and $n_2 = 10$, so

$$s_p^2 = \frac{[11(16) + 9(25)]}{(12 + 10 - 2)} = 20.05$$

and $s_p = 4.478$.

Solution to Example 3 (Continued)

Step 4 (Continued)

- Hence

$$t = \frac{[(85-81)-0]}{\sqrt{20.05(1/12+1/10)}} = 2.086$$

- p-value $= 2 \times \min\{\Pr(T_{20} > 2.086)\,, \Pr(T_{20} < 2.806)\}$

$$= \ 2(0.025) \ = \ 0.05.$$

Solution to Example 2 (Continued)

Step 5

- **Conclusion**: Since the observed t-value = 2.086 which falls inside the critical region, hence H_0: $\mu_1 = \mu_2$ is rejected at the 10% level of significance

  and conclude that the two methods of learning are not equal.
- Since p-value = 0.05 is less than 0.10, therefore we reject H_0 at the 10% level of significance and conclude that the two methods of learning are not equal.

7.3.4 Paired Data

Refer to Section 6.4.4

Example 4

- We wish to compare two methods for determining the percentage of iron ore in ore samples.
- Each of 12 ore samples was split into two parts.
- One-half of each sample was randomly selected and subjected to Method 1;
- The other half was subject to Method 2.
- The results are given in next slide.

Example 4 (Continued)

Sample	1	2	3	4	5	6
Method 1	38.25	31.68	26.24	41.29	44.81	46.37
Method 2	38.27	31.71	26.22	41.33	44.80	46.39
$d_i = X_1 - X_2$	–0.02	–0.03	0.02	–0.04	0.01	–0.02

Sample	7	8	9	10	11	12
Method 1	35.42	38.41	42.68	46.71	29.20	30.76
Method 2	35.46	38.39	42.72	46.76	29.18	30.79
$d_i = X_1 - X_2$	–0.04	0.02	–0.04	–0.05	0.02	–0.03

Example 4 (Continued)

- Do the data provide sufficient evidence that Method 2 yields a higher average percentage than Method 1?
- Assume the differences are normally distributed.
- Use $\alpha = 0.05$.

Solution to Example 4

Step 1

- Let μ_d be the average difference in percentage between methods 1 and 2.
- Test $\mathrm{H}_0: \mu_d = 0$ against $\mathrm{H}_1: \mu_d < 0$. (why?)

Step 2

- Set $\alpha = 0.05$.

Solution to Example 4 (Continued)

Step 3

- $n = 12$ implies $t_{n-1;\alpha} = t_{11;0.05} = 1.796$.
- Critical region $t < -1.796$, where

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}}, \qquad \text{with } d_i = X_{1i} - X_{2i}$$

Solution to Example 4 (Continued)

Step 4

- **Computations:** From the data, we have $\sum_i d_i = -0.2$ and $\sum_i d_i^2 = 0.0112$. Hence $\bar{d} = -0.0167$ and $s_d^2 = [0.0112 - 12(-0.0167)^2]/11 = 0.00072$.
- Therefore

$$t = [(-0.0167) - 0]/\sqrt{0.00072/12} = -2.156.$$

- p-value $= \Pr(\boldsymbol{T_{11}} < -2.156) = 0.027$.

  [or p-value is between 0.05 and 0.025 since 2.156 is between $t_{11;0.05} = 1.796$ and $t_{11;0.025} = 2.201$ if statistical table is used.]

Solution to Example 4 (Continued)

Step 5

- Since the observed t-value = –2.156 falls in the critical region, hence H_0: $\mu_d = 0$ is rejected at the 5% level of significance and conclude that sufficient evidence exists to permit us to conclude that Method 2 yields a higher average percentage than does Method 1.
- Since p-value = 0.027 is less than 0.05, therefore we reject H_0 and conclude that sufficient evidence exists to permit us to conclude that Method 2 yields a higher average percentage than does Method 1.

7.4 Hypotheses Testing Concerning Variance

7.4.1 One Variance case

Assumption: Underlying distribution is normal

- Let $X_1, X_2, \cdots, X_n$ be a random sample of size n from a (approximate) $N(\mu, \sigma^2)$ distribution, where σ^2 is unknown.
- We wish to test null hypothesis

$$\mathrm{H}_0: \sigma^2 = \sigma_0^2.$$

- We know that

$$\chi^2 = \frac{(n_1 - 1)S^2}{\sigma_0^2} \sim \chi^2(n-1).$$

Hypothesis Testing for σ^2 (Continuous)

Hence

H_0	Test Statistic
$\sigma^2 = \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

Hypothesis Testing for σ^2 (Continuous)

- $\mathrm{H}_0: \sigma^2 = \sigma_0^2$ is rejected if the observed χ^2-value

$\mathbf{H_1}$	**Critical Region**
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{(n-1;\alpha)}$
$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{(n-1;1-\alpha)}$
$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{(n-1;1-\alpha/2)}$ or $\chi^2 > \chi^2_{(n-1;\alpha/2)}$

where $\Pr\left(W > \chi^2_{n-1;\alpha}\right) = \alpha$ with $W \sim \chi^2(n-1)$

Example 1

- A manufacturer of car batteries claims that the life of his batteries is approximately normally distributed with a standard deviation equal to 0.9 year.
- If a random sample of 10 of these batteries has a standard deviation of 1.2 years,
- do you think that $\sigma > 0.9$ year?
- Use a 0.05 level of significance.

Solution to Example 1

Step 1

- Let σ^2 be the variance of the battery life.
- Test $\mathrm{H}_0: \sigma^2 = 0.81$. $\mathrm{H}_1: \sigma^2 > 0.81$.

Step 2

- Set $\alpha = 0.05$.

Solution to Example 1 (Continued)

Step 3

- $n = 10$ implies $\chi^2_{n-1;\,\alpha} = \chi^2_{9;0.05} = 16.919$.
- Critical region $\chi^2 > 16.919$, where

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2},$$

with $n = 10$ and $\sigma_0^2 = 0.81$.

Solution to Example 1 (Continued)

Step 4

- **Computations:**

  $s^2 = 11.44$, and $n = 10$, so

$$\chi^2 = \frac{9(1.44)}{0.81} = 16.0.$$

- p-value $= \Pr(\chi_9^2 > 16) = 0.0669$. [or it is between 0.05 and 0.10 from the statistical table]

Solution to Example 1 (Continued)

Step 5

- **Conclusion:** Since the observed χ^2-value = 16, which falls outside the critical region, hence H_0: $\sigma^2 = 0.81$ is not rejected at the 5% level of significance and conclude that there is no reason to doubt that the standard deviation is 0.9 year. Or
- Since p-value is greater than 0.05, we do not reject H_0.

7.4.2 H.T. Concerning Ratio of Variances

Assumption:

1. **Underlying distributions is normal**
2. **Means are unknown**

H.T. Concerning Ratio of Variances (Continued)

Examples

- When we are comparing the precision of one measuring device with that of another,
- the variability in grading practices of one teacher with that of another, and
- the consistency of one production process with that of another,
- we are testing about the difference between two population variances (or standard deviations).

H.T. Concerning Ratio of Variances (Continued)

- We know that when two independent samples of sizes n_1 and n_2 are randomly selected from two normal populations then

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1).$$

- Under H_0: $\sigma_1^2 = \sigma_2^2$,

$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1).$$

H.T. Concerning Ratio of Variances (Continued)

- Hence

H_0	Test Statistic
$\sigma_1^2 = \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2}$

H.T. Concerning Ratio of Variances (Continued)

- $\text{H}_0: \sigma_1^2 = \sigma_2^2$ is rejected if the observed F-value falls in the critical region

H_1	**Critical Region**
$\sigma_1^2 > \sigma_2^2$	$F > F_{(n_1-1, n_2-1\,; \alpha)}$
$\sigma_1^2 > \sigma_2^2$	$F < F_{(n_1-1, n_2-1\,; 1-\alpha)}$
$\sigma_1^2 \neq \sigma_2^2$	$F < F_{(n_1-1, n_2-1\,; 1-\alpha/2)}$ or $F > F_{(n_1-1, n_2-1\,; \alpha/2)}$

where $\Pr\left(W > F_{v_1, v_2; \alpha}\right) = \alpha$ with $W \sim F(v_1, v_2)$

Example 2

- An experiment was performed to compare the abrasive wear of two different laminated materials.
- Eleven pieces of Material 1 were tested, by exposing each piece to a machine measuring wear.
- Nine pieces of Material 2 were similarly tested.
- In each case, the depth of wear was observed.
- The samples of Material 1 gave an average (coded) wear of 85 units with a standard deviation of 4,

Example 2 (Continued)

- while the samples of Material 2 gave an average of 81 and a standard deviation of 5.
- Assume that the two unknown populations to be approximately normal,
- test the two variances are equal.
- Use a 0.10 level of significance.

Solution to Example 2

Step 1

- Let σ_1^2 and σ_2^2 be the variances of the abrasive wear made from Materials 1 and 2 respectively.
- Test: $\sigma_1^2 = \sigma_2^2$ against H_1: $\sigma_1^2 \neq \sigma_2^2$.

Step 2

- Set $\alpha = 0.1$.

Solution to Example 2 (Continued)

Step 3

- $n_1 = 11, n_2 = 9$ implies $F_{n_1-1\,,n_2-1\,;\alpha/2} = F_{10,8;0.05} = 3.35$ and
- $F_{n_1-1,n_2-1;1-\alpha/2} = F_{10,8;0.95} = 1/F_{8,10;0.05} = 1/3.07 = 0.326.$
- Critical region: $F > 3.35$ or $F < 0.326$, where $F = s_1^2/S_2^2$

Solution to Example 2 (Continued)

Step 4

- **Computations:**

  $s_1^2 = 16, s_2^2 = 25$, so $F = 16/25 = 0.64.$

Step 5

- **Conclusion:** Since the observed F-value = 0.64 which falls outside the critical region, hence H_0: $\sigma_1^2 = \sigma_2^2$ is not rejected at the 10% level of significance and we conclude that we were justified in assuming the unknown variances equal.