Chapter 9

Tests of Hypotheses for a Single Sample

Chapter 9: Tests of Hypotheses for a Single Sample

LEARNING OBJECTIVES

9.1. Hypothesis Testing

9.2. Test on the μ of NORMDIST, σ^2 known

9.3. Test on the μ of NORMDIST, σ^2 unknown

9.5. Test on the *p*: Large-sample

9.1. Hypothesis Testing

Introduction

Definition **Statistical Hypotheses**

A statistical hypothesis is a statement about the parameters of one or more populations.

Example

Suppose that we are interested in the burning rate of a solid propellant used to power aircrew escape systems.

• Burning rate is a random variable

• We are interested in deciding whether or not the mean burning rate is 50 centimeters per second.

Introduction

Two-sided Alternative Hypothesis

$$H_0 : \mu = 50 \; ; \; H_1 : \mu \neq 50$$

Null hypothesis

Alternative hypothesis

One-sided Alternative Hypotheses

$$H_0 : \mu = 50 \; ; \; H_1 : \mu > 50$$

$$H_0 : \mu = 50 \; ; \; H_1 : \mu < 50$$

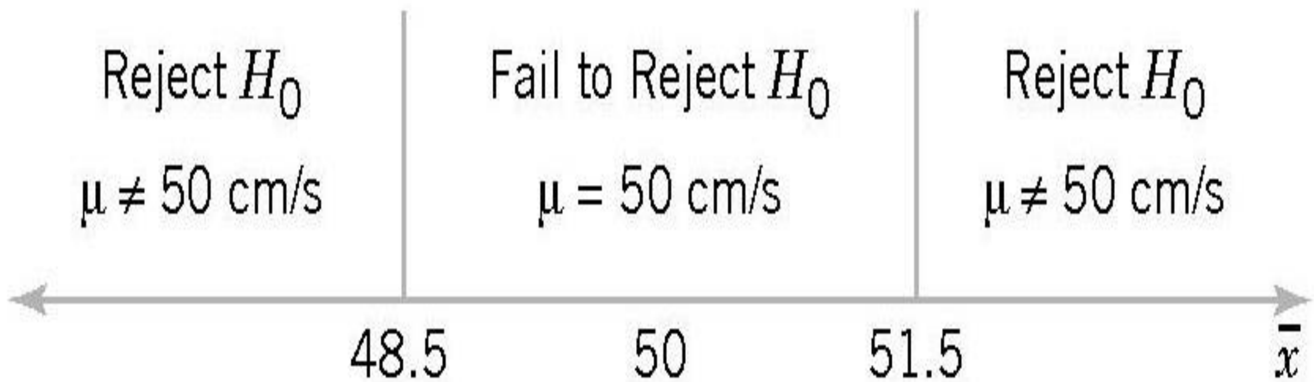
Introduction

Test of a Hypothesis

- A procedure leading to a decision about a particular hypothesis
- Hypothesis-testing procedures rely on using the information in a random sample from the population of interest.
- If this information is *consistent* with the hypothesis, then we will conclude that the hypothesis is true; if this information is *inconsistent* with the hypothesis, we will conclude that the hypothesis is false.

Example 1

$$H_0:\ \mu = 50\ ;\ H_1:\ \mu \neq 50$$



Decision criteria for testing H_0: $\mu = 50$; H_1: $\mu \neq 50$

Introduction

Table 9-1 Decisions in Hypothesis Testing

Decision	H_0 Is True	H_0 Is False
Fail to reject H_0	no error	type II error
Reject H_0	type I error	no error

$$\alpha = P(\text{type I error}) = P(\text{ reject } H_0 \text{ when } H_0 \text{ is true })$$

Significance level

$$\beta = P(\text{type II error}) = P(\text{ fail to reject } H_0 \text{ when } H_0 \text{ is false})$$

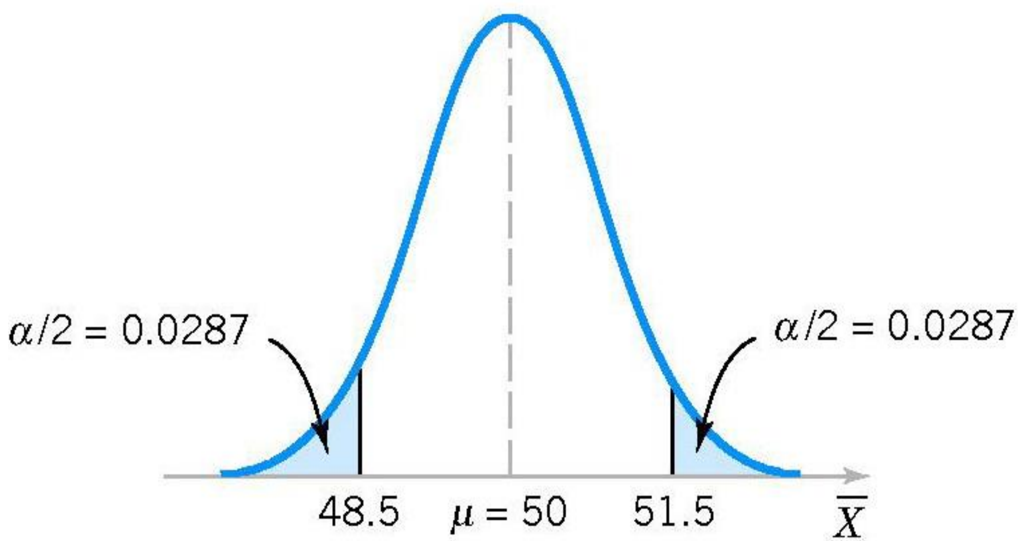
Introduction

To calculate α, we assume that σ = 2.5 and n = 10.

$$\alpha = P(\bar{X} < 48.5 \text{ when } \mu = 50) + P(\bar{X} > 51.5 \text{ when } \mu = 50)$$

$$= P(Z < -1.90) + P(Z > 1.90) \simeq 0.0574$$

$$= \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{X}-50}{2.5/\sqrt{10}}$$



Introduction

$$\beta = P(48.5 < \overline{X} < 51.5 \ \ when \ \ \mu = 52)$$
$$= P(-4.43 < Z < -0.63) \simeq 0.2643$$

$$\beta = P(48.5 < \overline{X} < 51.5 \ \ when \ \ \mu = 50.5)$$
$$= P(-2.53 < Z < 1.27) \simeq 0.8923$$

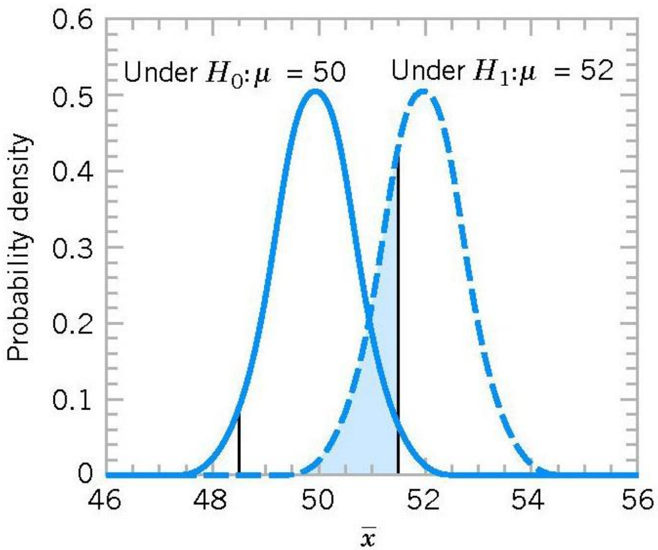


Figure 9-3 The probability of type II error when μ = 52 and n = 10.

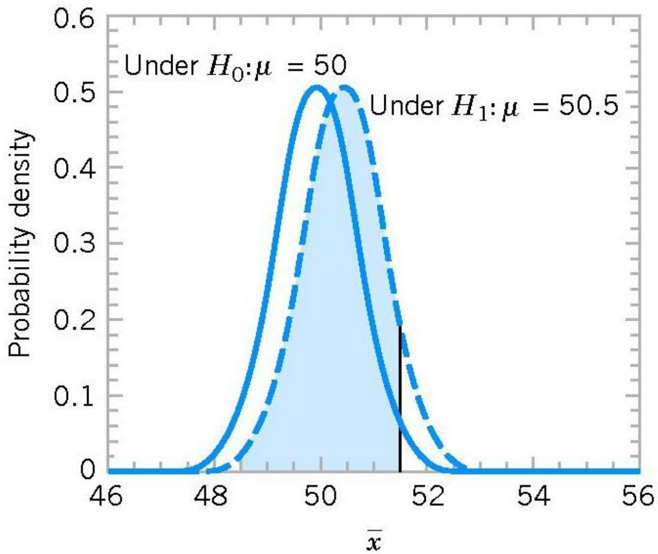


Figure 9-4 The probability of type II error when μ = 50.5 and n = 10.

Introduction

Acceptance Region	Sample Size	α	β at $\mu = 52$	β at $\mu = 50.5$
$48.5 < \bar{x} < 51.5$	10	0.0576	0.2643	0.8923
$48 < \bar{x} < 52$	10	0.0114	$\boxed{0.5000}$	$\boxed{0.9705}$
$48.5 < \bar{x} < 51.5$	16	0.0164	0.2119	$\boxed{0.9445}$
$48 < \bar{x} < 52$	16	$\boxed{0.0014}$	$\boxed{0.5000}$	$\boxed{0.9918}$

Introduction

1. The size of the critical region, and consequently the probability of a type I error, can always be reduced by appropriate selection of the critical values.

2. Type I and type II errors are related: $\beta \downarrow$ then $\alpha \uparrow$ and $\alpha \downarrow$ then $\beta \uparrow$, provided that n fix.

3. An increase in sample size will generally reduce both α and β, provided that the critical values are held constant.

4. When the null H_0 is false, $\beta \uparrow$ as the true value of the parameter approaches the value hypothesized in H_0,$\beta \downarrow$ as the difference between the true mean and the hypothesized value increases.

Introduction

Definition **power of a statistical test**

The **power** of a statistical test is the probability of rejecting the null hypothesis H_0 when the alternative hypothesis is true.

• The power is computed as 1 - beta, and power can be interpreted as *the probability of correctly rejecting a false null hypothesis*.

• We often compare statistical tests by comparing their power properties.

Exercises for Section 9.1

A textile fiber manufacturer is investigating a new drapery yarn, which the company claims has a mean thread elongation of 12 kilograms with a standard deviation of 0.5 kilograms. The company wishes to test the hypothesis H_0: $\mu = 12$ against H_1: $\mu < 12$, using a random sample of four specimens.

a. What is the type I error probability if the critical region is defined as $x < 11.5$ kilograms?

b. Find β for the case in which the true mean elongation is 11.25 kilograms.

c. Find β for the case in which the true mean is 11.5 kilograms.

Exercises for Section 9.1

A consumer products company is formulating a new shampoo and is interested in foam height (in millimeters). Foam height is approximately normally distributed and has a standard deviation of 20 millimeters. The company wishes to test H_0: $\mu = 175$ millimeters versus H_1: $\mu > 175$ millimeters, using the results of $n = 10$ samples.

a. Find the type I error probability α if the critical region is
$x > 185$.

b. What is the probability of type II error if the true mean foam height is 185 millimeters?

c. Find β for the true mean of 195 millimeters.

Exercises for Section 9.1

A manufacturer is interested in the output voltage of a power supply used in a PC. Output voltage is assumed to be normally distributed with standard deviation 0.25 volt, and the manufacturer wishes to test H_0: $\mu = 5$ volts against H_1: $\mu \neq 5$ volts, using $n = 8$ units.

a. The acceptance region is $4.85 \leq x \leq 5.15$. Find the value of α.

b. Find the power of the test for detecting a true mean output voltage of 5.1 volts.

9.2. Test on the μ of NORMDIST, σ^2 known

σ^2 known

Hypothesis-testing problems

Two-sided test: H_0: $\mu = \mu_0$; H_1: $\mu \neq \mu_0$

One-sided test: H_0: $\mu = \mu_0$; H_1: $\mu > \mu_0$

H_0: $\mu = \mu_0$; H_1: $\mu < \mu_0$

σ^2 known

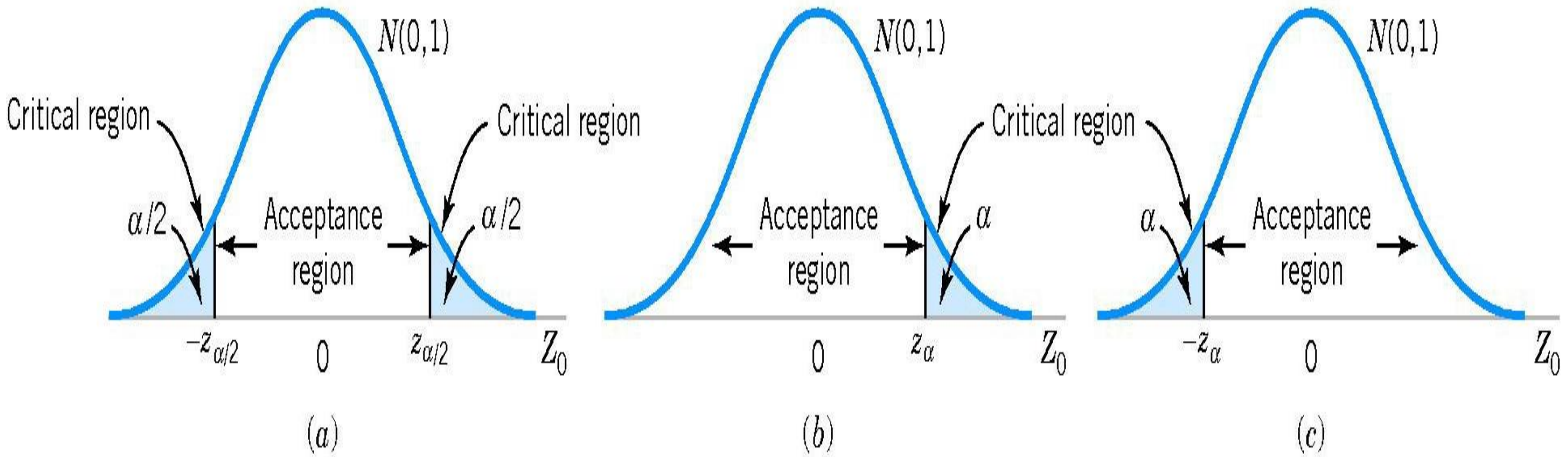


Figure 9-6 The distribution of Z_0 when H_0: $\mu = \mu_0$ is true, with critical region for (a) the two-sided alternative $H_1: \mu \neq \mu_0$, (b) the one-sided alternative $H_1: \mu > \mu_0$, and (c) the one-sided alternative $H_1: \mu < \mu_0$.

σ^2 known

Hypothesis Tests on the Mean	*z*-test

Null hypothesis: H_0: $\mu = \mu_0$

Test statistic: $$Z_0 = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Alternative hypothesis	Rejection criteria
H_1: $\mu \neq \mu_0$	$\lvert z_0 \rvert > z_{\alpha/2}$
H_1: $\mu > \mu_0$	$z_0 > z_\alpha$
H_1: $\mu < \mu_0$	$z_0 < -z_\alpha$

σ^2 known

Hypothesis Tests on the Mean *z*-test

1. Parameter of interest: From the problem context, identify the parameter of interest.

2. Null hypothesis: H_0: $\mu = \mu_0$

Alternative hypothesis: H_1:

$\mu \neq \mu_0$ (1) $\mu > \mu_0$ (2) $\mu < \mu_0$ (3)

3. Test statistic:

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

4. Reject H_0 if

(1) $|z_0| > z_{\alpha/2}$ (2) $z_0 > z_\alpha$ (3) $z_0 < -z_\alpha$

Example 2 σ^2 known

A melting point test of n = 10 samples of a binder used in manufacturing a rocket propellant resulted in $\bar{x} = 154.2^0$. Assume that melting point is normally distributed with $\sigma = 1.5^0$.

Test H_0: $\mu = 155$; H_1: $\mu \neq 155$ using $\alpha = 0.05$.

Solution: We have $z_{\alpha/2} = z_{0.025} = 1.96$.

Test statistic:

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{154.2 - 155}{1.5 / \sqrt{10}} \simeq -1.69$$

$| z_0 | < z_{\alpha/2}$.

=> Fail to reject H_0 at $\alpha = 0.05$

σ^2 known

P-Values in Hypothesis Tests

Definition

The ***P*-value** is the smallest level of significance that would lead to rejection of the null hypothesis H_0 with the given data.

$$P = \begin{cases} 2[1 - \Phi(|z_0|)] & \text{for a two-tailed test: } H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0 \\ 1 - \Phi(z_0) & \text{for a upper-tailed test: } H_0: \mu = \mu_0 \quad H_1: \mu > \mu_0 \\ \Phi(z_0) & \text{for a lower-tailed test: } H_0: \mu = \mu_0 \quad H_1: \mu < \mu_0 \end{cases}$$

σ^2 known

Exercise 2

Air crew escape systems are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specifications require that the mean burning rate must be 50 centimeters per second. We know that the standard deviation of burning rate is $\sigma = 2$ centimeters per second. The experimenter decides to specify a type I error probability or significance level of $\alpha = 0.05$ and selects a random sample of $n = 25$ and obtains a sample average burning rate of bar(x) = 51.3 centimeters per second. What conclusions should be drawn?

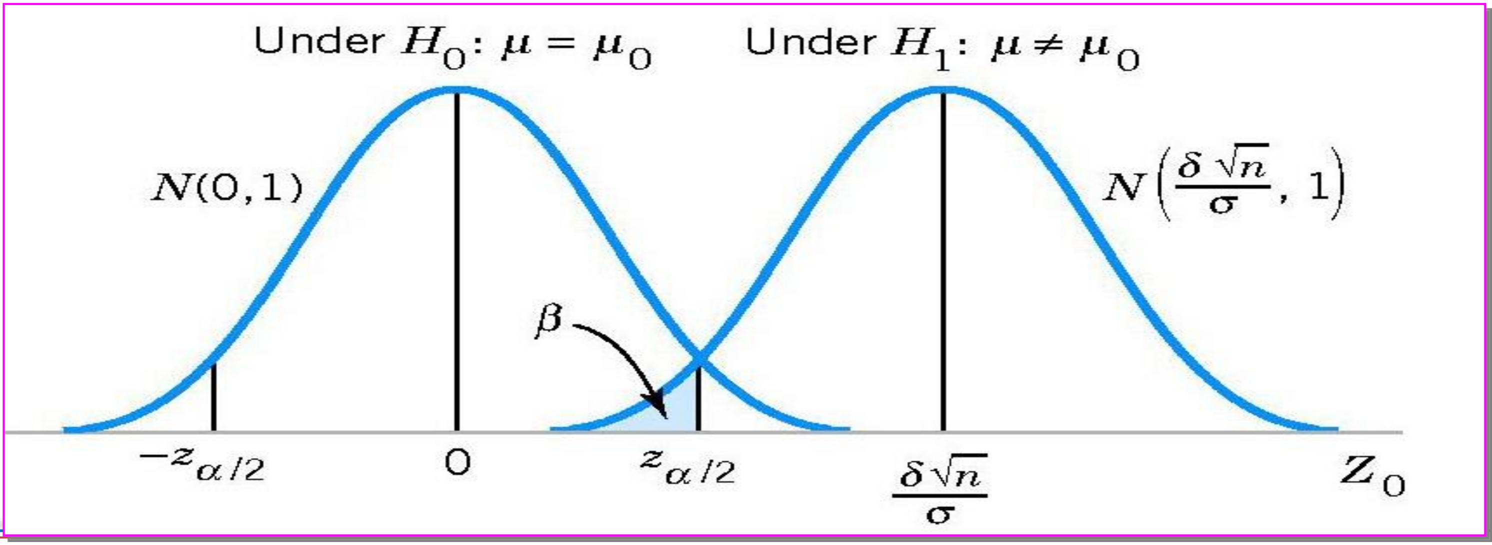
σ^2 known

Type II Error and Choice of Sample Size

Probability of type II error β for a two-sided test

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

where $\delta = \mu - \mu_0$



σ^2 known

Type II Error and Choice of Sample Size

Sample Size Formulas

Two-sided test

$$n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2 \sigma^2}{\delta^2}$$

One-sided test

$$n \simeq \frac{(z_{\alpha} + z_{\beta})^2 \sigma^2}{\delta^2}$$

where $\delta = \mu - \mu_0$.

Exercise 4

Consider the rocket propellant problem in Example 3. Suppose that the true burning rate is 49 centimeters per second. What is β for the two-sided test with $\alpha = 0.05$, $\sigma = 2$, and $n = 25$?

Exercise

The life in hours of a battery is known to be approximately normally distributed, with standard deviation σ =1.25 hours. A random sample of 10 batteries has a mean life of $\bar{X} = 40.5$ hours.

(a) Is there evidence to support the claim that battery life exceeds 40 hours? Use $\alpha = 0.05$.

(b) What is the *P*-value for the test in part (a)?

(c) What is the β-error for the test in part (a) if the true mean life is 42 hours?

(d) What sample size would be required to ensure that does not exceed 0.10 if the true mean life is 44 hours?

9.3. Test on the μ of NORMDIST, σ^2 Unknown

σ^2 unknown

Hypothesis Tests on the Mean — *t*-test

Null hypothesis: H_0: $\mu = \mu_0$

Test statistic: $$T_0 = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

Alternative hypothesis	Rejection criteria
H_1: $\mu \neq \mu_0$	$\lvert t_0 \rvert > t_{\alpha/2,\, n-1}$
H_1: $\mu > \mu_0$	$t_0 > t_{\alpha,\, n-1}$
H_1: $\mu < \mu_0$	$t_0 < -t_{\alpha,\, n-1}$

σ^2 unknown

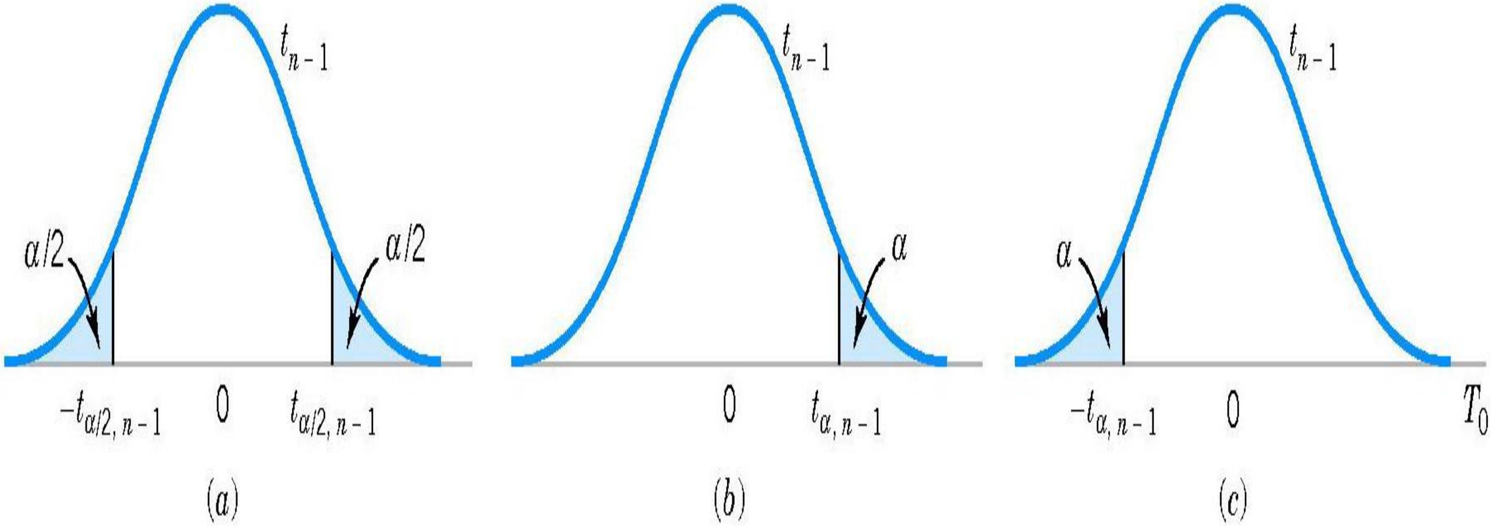


Figure 9-8 The reference distribution for H_0: $\mu = \mu_0$ with critical region for (a) H_1: $\mu \neq \mu_0$, (b) H_1: $\mu > \mu_0$, and (c) H_1: $\mu < \mu_0$.

σ^2 unknown

Hypothesis Tests on the Mean *t*-test

Null hypothesis: $H_0: \mu = \mu_0$

Test statistic: $$T_0 = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

Alternative hypothesis	P_value
$H_1: \mu \neq \mu_0$	$P = 2P(T_{n-1} > \lvert t_0 \rvert)$
$H_1: \mu > \mu_0$	$P = P(T_{n-1} > t_0)$
$H_1: \mu < \mu_0$	$P = P(T_{n-1} < t_0)$

Type II Error and Choice of Sample Size

$$\beta = P\left(-t_{\alpha/2,n-1} \le T_0 \le t_{\alpha/2,n-1} \mid \delta \ne 0\right)$$
$$= P\left(-t_{\alpha/2,n-1} \le T_0^{'} \le t_{\alpha/2,n-1}\right)$$

TEST ON THE μ OF NORMDIST: σ^2 UNKNOWN

Exercise

Consider the dissolved oxygen concentration at TVA dams. The observations are (in milligrams per liter): 5.0, 3.4, 3.9, 1.3, 0.2, 0.9, 2.7, 3.7, 3.8, 4.1, 1.0, 1.0, 0.8, 0.4, 3.8, 4.5, 5.3, 6.1, 6.9, and 6.5.

(a) Test the hypotheses H_0: $\mu = 4$; H_1: $\mu \neq 4$. Use $\alpha = 0.01$.

(b) What is the P-value in part (a)?

9.5 Tests on a Population Proportion

Test on the *p*

Hypothesis Tests on the Mean

Null hypothesis: $H_0: p = p_0$

Test statistic: $$Z_0 = \frac{\hat{P} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$$

Alternative hypothesis	Rejection criteria
$H_1: p \neq p_0$	$\lvert z_0 \rvert > z_{\alpha/2}$
$H_1: p > p_0$	$z_0 > z_\alpha$
$H_1: p < p_0$	$z_0 < - z_\alpha$

Test on the *p*

Example

In a random sample of 85 automobile engine crankshaft bearings, 10 have a surface finish roughness that exceeds the specifications. Does this data present strong evidence that the proportion of crankshaft bearings exhibiting excess surface roughness exceeds 0.10? State and test the appropriate hypotheses using $\alpha = 0.05$.

1. We have $n = 85,\ p = \dfrac{10}{85} = 0.12, z_\alpha = 1.645$

2. Null hypothesis: H_0: $p = 0.1$

Alternative hypothesis: H_1: $p > 0.1$

3. Test statistic:

$$z_0 = \frac{p - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} = \frac{0.12 - 0.1}{\sqrt{\dfrac{0.1(1-0.1)}{85}}} = 0.61$$

4. Therefore, because $z_0 < z_\alpha$, we fail to reject H_0.

Example

A commonly prescribed drug for relieving nervous tension is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance.

Summary of Approximate Tests on a Binomial Proportion

Testing Hypotheses on a Binomial Proportion

Null hypotheses: $H_0: p = p_0$

Test statistic: $$Z_0 = \frac{X - np_0}{\sqrt{np_0(1 - p_0)}}$$

Alternative Hypotheses	*P*-Value	Rejection Criterion for Fixed-Level Tests
$H_1: p \neq p_0$	Probability above $\|z_0\|$ and probability below $-\|z_0\|$, $P = 2[1 - \Phi(z_0)]$	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$
$H_1: p > p_0$	Probability above z_0, $P = 1 - \Phi(z_0)$	$z_0 > z_\alpha$
$H_1: p < p_0$	Probability below z_0, $P = \Phi(z_0)$	$z_0 < -z_\alpha$

Example

A semiconductor manufacturer produces controllers used in automobile engine applications. The customer requires that the process fallout or fraction defective at a critical manufacturing step not exceed 0.05 and that the manufacturer demonstrate process capability at this level of quality using $\alpha = 0.05$. The semiconductor manufacturer takes a random sample of 200 devices and finds that four of them are defective. Can the manufacturer demonstrate process capability for the customer?

Test on the p

Type II Error and Choice of Sample Size

Probability of type II error β

Two-sided test

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right) - \Phi\left(\frac{p_0 - p - z_{\alpha/2}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

One-sided test: $p > p_0$

$$\beta = \Phi\left(\frac{p_0 - p + z_{\alpha}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

One-sided test: $p < p_0$

$$\beta = 1 - \Phi\left(\frac{p_0 - p - z_{\alpha}\sqrt{p_0(1-p_0)/n}}{\sqrt{p(1-p)/n}}\right)$$

Test on the *p*

Type II Error and Choice of Sample Size

Sample Size Formulas

Two-sided test

$$n = \left[\frac{z_{\alpha/2}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p-p_0} \right]^2$$

One-sided test

$$n = \left[\frac{z_{\alpha}\sqrt{p_0(1-p_0)} + z_\beta\sqrt{p(1-p)}}{p-p_0} \right]^2$$

Example

A random sample of 500 registered voters in Phoenix is asked if they favor the use of oxygenated fuels year-round to reduce air pollution. If more than 315 voters respond positively, we will conclude that at least 60% of the voters favor the use of these fuels.

a. Find the probability of type I error if exactly 60% of the voters favor the use of these fuels.

b. What is the type II error probability β if 75% of the voters favor this action?