9-1 Hypothesis Testing

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**.

9-1 Hypothesis Testing

9-1.2 Tests of Statistical Hypotheses

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

Probability of Type I and Type II Error

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

Sometimes the type I error probability is called the **significance level**, or the **α-error**, or the **size** of the test.

9-1 Hypothesis Testing

The **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 - β, and power can be interpreted as *the probability of correctly rejecting a false null hypothesis.*

9-1 Hypothesis Testing

9-1.4 *P*-Value

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-value is the **observed significance level.**

9-1 Hypothesis Testing

9-1.6 General Procedure for Hypothesis Tests

1. Identify the parameter of interest.
2. Formulate the null hypothesis, H_0 .
3. Specify an appropriate alternative hypothesis, H_1.
4. Choose a significance level, ⍰.
5. Determine an appropriate test statistic.
6. State the rejection criteria for the statistic.
7. Compute necessary sample quantities for calculating the test statistic.
8. Draw appropriate conclusions.

9-2.1 Hypothesis Tests on the Mean

Consider the two-sided hypothesis test

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

The **test statistic** is:

(9-1)

9-2 Tests on the Mean of a Normal Distribution, Variance Known

9-2.1 Hypothesis Tests on the Mean

Reject H_0 if the observed value of the test statistic z_0 is either:

$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$

Fail to reject H_0 if the observed value of the test statistic z_0 is

$-z_{\alpha/2} < z_0 < z_{\alpha/2}$

EXAMPLE 9-2 Propellant Burning Rate

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 and the standard deviation is □ □ 2 centimeters per second. The significance level of □ □ 0.05 and a random sample of n □ 25 has a sample average burning rate of

centimeters per second. Draw conclusions.

The seven-step procedure is

1. **Parameter of interest:** The parameter of interest is □, the mean burning rate.
2. **Null hypothesis:** H_0: □ □ 50 centimeters per second
3. **Alternative hypothesis:** H_1: □ □ 50 centimeters per second

EXAMPLE 9-2 Propellant Burning Rate

4. Test statistic: The test statistic is

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

5. Reject H_0 if: Reject H_0 if the P-value is less than 0.05. The boundaries of the critical region would be $z_{0.025} = 1.96$ and $-z_{0.025} = -1.96$.

6. Computations: Since $\bar{x} = 51.3$ and $\sigma = 2$,

$$z_0 = \frac{51.3 - 50}{2/\sqrt{25}} = 3.25$$

7. Conclusion: Since $z_0 = 3.25$ and the p-value is $= 2[1 - \Phi(3.25)] = 0.0012$, we reject H_0: $\mu = 50$ at the 0.05 level of significance.

OR

$z_0 = 3.25 > Z_{0.025=1.96}$ Therefore Reject H_0

Interpretation: The mean burning rate differs from 50 centimeters per second, based on a sample of 25 measurements.

Type II Error

Finding the Probability of Type II Error □

Consider the two-sided hypotheses test H_0: $\square\ \square\ \square_0$ and H_1: $\square\ \square\ \square_0$

Suppose the null hypothesis is false and the true value of the mean is $\square\ \square\ \square_0\ \square\ \square$, where $\square\ \square$ 0.

(9-3)

Propellant Burning Rate Type II Error

Consider the rocket propellant problem of Example 1. The true burning rate is 49 centimeters per second. Find β for the two-sided test with $\alpha = 0.05$, $\sigma = 2$, and $n = 25$?

Here $\delta = 1$ and $z_{\alpha/2} = 1.96$. From Equation 9-3,

The probability is about 0.3 that the test will fail to reject the null hypothesis when the true burning rate is 49 centimeters per second.

Interpretation: A sample size of $n = 25$ results in reasonable, but not great power $= 1 - \beta = 1 - 0.3 = 0.70$.

9-3.1 Hypothesis Tests on the Mean

One-Sample *t*-Test

Consider the two-sided hypothesis test

$$H_0: \mu = \mu_0 \quad \text{and} \quad H_1: \mu \neq \mu_0$$

Test statistic:

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

EXAMPLE 9-6 Golf Club Design

An experiment was performed in which 15 drivers produced by a particular club maker were selected at random and their coefficients of restitution measured. It is of interest to determine if there is evidence (with ⍰⍰ 0.05) to support a claim that the mean coefficient of restitution exceeds 0.82.

The observations are:

0.8411	0.8191	0.8182	0.8125	0.8750
0.8580	0.8532	0.8483	0.8276	0.7983
0.8042	0.8730	0.8282	0.8359	0.8660

The sample mean and sample standard deviation are and $s = 0.02456$. The objective of the experimenter is to demonstrate that the mean coefficient of restitution exceeds 0.82, hence a one-sided alternative hypothesis is appropriate.

The seven-step procedure for hypothesis testing is as follows:

1. **Parameter of interest:** The parameter of interest is the mean coefficient of restitution, .

2. **Null hypothesis:** H_0: 0.82

3. **Alternative hypothesis:** H_1: 0.82

EXAMPLE 9-6 Golf Club Design - Continued

4. Test Statistic: The test statistic is

5. Reject H_0 if: Reject H_0 if the *P*-value is less than 0.05.

6. Computations: Since , s = 0.02456, = 0.82, and n 15, we have

7. Conclusions:

= 2.72

Interpretation: There is strong evidence to conclude that the mean coefficient of restitution exceeds 0.82.

9-4 Hypothesis Tests on the Variance and Standard Deviation of a Normal Distribution

9-4.1 Hypothesis Test on the Variance

Suppose that we wish to test the hypothesis that the variance of a normal population σ^2 equals a specified value, say , or equivalently, that the standard deviation σ is equal to σ_0. Let X_1, X_2,... ,Xn be a random sample of *n* observations from this population. To test

(9-6)

we will use the test statistic:

(9-7)

9-4.1 Hypothesis Test on the Variance

If the null hypothesis is true, the test statistic defined in Equation 9-7 follows the chi-square distribution with *n* 1 degrees of freedom. This is the reference distribution for this test procedure.

Therefore, we calculate , the value of the test statistic , and the null hypothesis would be rejected if

where and are the upper and lower 100 /2 percentage points of the chi-square distribution with *n* 1 degrees of freedom, respectively. Figure 9-17(a) shows the critical region.

9-4.1 Hypothesis Test on the Variance

The same test statistic is used for one-sided alternative hypotheses. For the one-sided hypotheses.

(9-8)

we would reject H_0 if , whereas for the other one-sided hypotheses

(9-9)

we would reject H_0 if . The one-sided critical regions are shown in Fig. 9-17(b) and (c).

9-4.1 Hypothesis Tests on the Variance

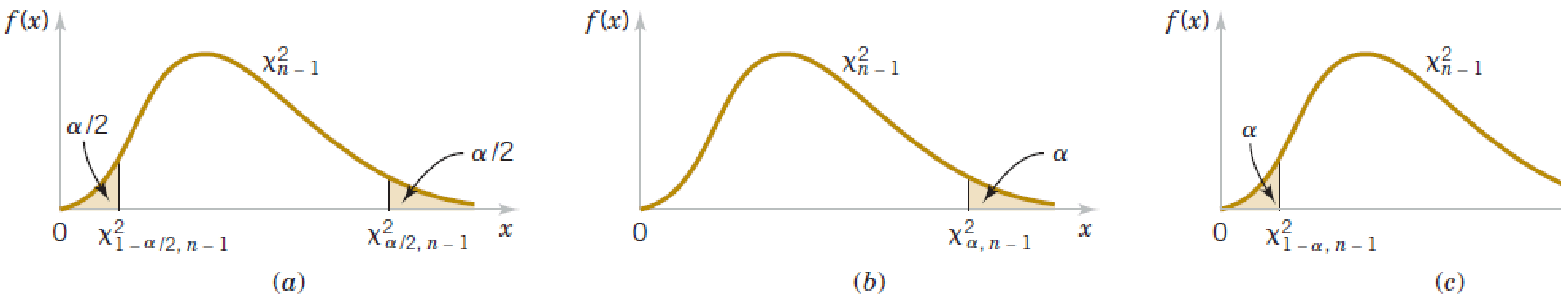

FIGURE 9-17 Reference distribution for the test of $H_0\colon \sigma^2 = \sigma_0^2$ with critical region values for (a) $H_1\colon \sigma^2 \neq \sigma_0^2$. (b) $H_1\colon \sigma^2 > \sigma_0^2$. (c) $H_1\colon \sigma^2 < \sigma_0^2$.

EXAMPLE 9-8 **Automated Filling**

An automated filling machine is used to fill bottles with liquid detergent. A random sample of 20 bottles results in a sample variance of fill volume of $s^2 = 0.0153$ (fluid ounces)^2. If the variance of fill volume exceeds 0.01 (fluid ounces)^2, an unacceptable proportion of bottles will be underfilled or overfilled. Is there evidence in the sample data to suggest that the manufacturer has a problem with underfilled or overfilled bottles? Use $\alpha = 0.05$, and assume that fill volume has a normal distribution.

Using the seven-step procedure results in the following:

1. **Parameter of Interest:** The parameter of interest is the population variance σ^2.

2. **Null hypothesis:** H_0: $\sigma^2 = 0.01$

3. **Alternative hypothesis:** H_1: $\sigma^2 > 0.01$

9-4 Hypothesis Tests on the Variance and Standard Deviation of a Normal Distribution

Example 9-8

1. **Test statistic:** The test statistic is

2. **Reject H_0:** Use α = 0.05, and reject H_0 if .

3. **Computations:**

4. **Conclusions:** , Therefore Accept

we conclude that there is no strong evidence that the variance of fill volume exceeds 0.01 (fluid ounces)^2.

9-5 Tests on a Population Proportion

9-5.1 Large-Sample Tests on a Proportion

Many engineering decision problems include hypothesis testing about *p*.

An appropriate **test statistic** is

$$Z_0 = \frac{\hat{P} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

9-5 Tests on a Population Proportion

EXAMPLE 9-10 **Automobile Engine Controller** 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?

We may solve this problem using the seven-step hypothesis-testing procedure as follows:

1. **Parameter of Interest:** The parameter of interest is the process fraction defective *p.*

2. **Null hypothesis:** $H_0\!: p = 0.05$

3. **Alternative hypothesis:** $H_1\!: p < 0.05$

9-5 Tests on a Population Proportion

Example 9-10

1. **The test statistic is:**

$$Z_0 = \frac{\hat{P} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

where x = 4, n = 200, and p_0 = 0.05.   ==0.02

2. **Reject H_0 if:** Reject H_0: p = 0.05 if the

3. **Computations:** The test statistic is

=-1.95

7. **Conclusions:**

Therefore the defective rate is less than 0.05