Chapter 5
One and Two-Sample Tests of Hypotheses

Department of Statistics and Operations Research

October 21, 2019

Plan

Plan

Consider a population with some unknown parameter Θ. We are interested in testing (confirming or denying) some conjectures about Θ. For example, we might be interested in testing the conjecture that

$$\Theta = \Theta_0 \text{ where } \Theta_0 \text{ is a given value.}$$

Definition

A **statistical hypothesis** is an assertion or conjecture concerning one or more populations.

Plan

We usually test the null hypothesis:

$$H_0 : \Theta = \Theta_0 \quad \text{(Null Hypothesis)}.$$

against one of the following alternative hypotheses:

$$H_1 : \begin{cases} \Theta \neq \Theta_0, \\ \Theta > \Theta_0, \\ \Theta < \Theta_0 \end{cases} \text{(Alternative Hypothesis or Research Hypothesis)}$$

Possible situations in testing a statistical hypothesis:

	H_o is true	H_o is false
Accepting H_o	Correct Decision	Type II error (β)
Rejecting H_o	Type I error (α)	Correct Decision

Definition

Rejection of the null hypothesis when it is true is called a **type I error**.

The probability of committing a type I error, also called the level of significance, is denoted by α.

$$\Pr(\text{type I error}) = \Pr(\text{Rejecting } H_o | H_o \text{is true}) = \alpha$$

Definition

Nonrejection of the null hypothesis when it is false is called a **type II error**.

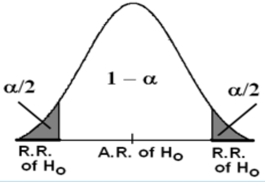
The probability of committing a type II error, denoted by β, is impossible to compute unless we have a specific alternative hypothesis.

$$\Pr(\text{type II error}) = \Pr(\text{Accepting } H_o | H_o \text{is false}) = \beta$$

Plan

Let $X_1, X_2, \ldots, X_n$ representing a random sample from a distribution with mean μ and variance $\sigma^2 > 0$. Consider first the hypothesis

$$H_0 \quad : \quad \mu = \mu_0,$$
$$H_1 \quad : \quad \mu \neq \mu_0.$$



Two-sided alternative

Test statistic (T.S.)

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

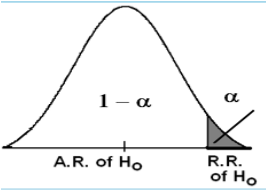
and

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

Theorem Test Procedure for a Single Mean (Variance Known)

If $-z_{\alpha}/2 < z < z_{\alpha}/2$, do not reject H_0. Rejection of H_0, of course, implies acceptance of the alternative hypothesis $\mu \neq \mu_0$. With this definition of the critical region, it should be clear that there will be probability α of rejecting H_0 (falling into the critical region) when, indeed, $\mu = \mu_0$.

$$H_0 : \mu = \mu_0,$$
$$H_1 : \mu > \mu_0.$$



one-sided alternative

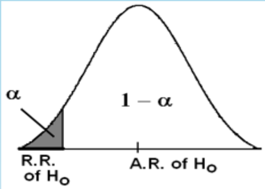
Test statistic (T.S.)

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

and

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

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



one-sided alternative

Test statistic (T.S.)

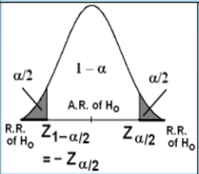
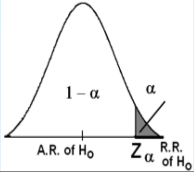
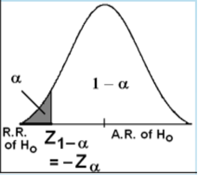
$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

and

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

Conclusion

Test Procedure

Hypotheses	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$
Test Statistic (T.S.)	$Z = \dfrac{\overline{X} - \mu_o}{\sigma / \sqrt{n}} \sim N(0,1)$		
R.R. and A.R. of H_o			
Decision	Reject H_o (and accept H_1) at the significance level α if:		
	$Z > Z_{\alpha/2}$ or $Z < -Z_{\alpha/2}$ Two-Sided Test	$Z > Z_\alpha$ One-Sided Test	$Z < -Z_\alpha$ One-Sided Test

Example

A random sample of 100 recorded deaths in the United States during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

Solution

1)

$$\begin{cases} H_0: \mu = 70 years, \\ H_1: \mu > 70 years. \end{cases}$$

2) $\alpha = 0.05$.

3) Test statistic:
$z = (\overline{x} - \mu_0)/(\sigma/\sqrt{n}) = (71.8 - 70)/(8.9/\sqrt{100}) = 2.02$.

4) Critical region: $z > z_\alpha$, where $z_\alpha = z_{0.05} = 1.645$.

5) Decision: since $z = 2.02 > 1.645$, reject H_0 and conclude that the mean life span today is greater than 70 years.

The P-value corresponding to $z = 2.02$ is given by the area on the right under the density of standard normal distribution. Using Table A.3, we have P-value $= P(Z > 2.02) = 0.0217$. As a result, the evidence in favor of H_1 is even stronger than that suggested by a 0.05 level of significance.

Example

A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5 kilogram. Test the hypothesis that $\mu = 8$ kilograms against the alternative that $\mu \neq 8$ kilograms if a random sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms. Use a 0.01 level of significance.

Solution

1)

$$\begin{cases} H_0: \mu = 8 \text{ kilograms}, \\ H_1: \mu \neq 8 \text{ kilograms}. \end{cases}$$

2) $\alpha = 0.01$.

3) Critical region: $z > z_{\alpha/2}$ and $z < -z_{\alpha/2}$, where $z = (\overline{x} - \mu_0)/(\sigma/\sqrt{n})$ and $z_{\alpha/2} = 2.575$.

4) Computations: $\overline{x} = 7.8$, $\sigma = 0.5$, $n = 50$, hence $z = (\overline{x} - \mu_0)/(\sigma/\sqrt{n}) = (7.8 - 8)/(0.5/\sqrt{50}) = -2.83$

5) Decision: since $z = -2.83 < -2.575$, hence reject H_0 and conclude that the average breaking strength is not equal to 8 but is, in fact, less than 8 kilograms.

Since the test in this example is two tailed, the desired P-value is twice the area of the left of $z = -2.83$. Therefore, using standard normal table, we have

$$\text{P-value} = P(Z > |2.83|) = 2P(Z > 2.83) = 0.0046 < 0.01$$

which allows us to reject the null hypothesis that $\mu = 8$ kilograms at a level of significance smaller than 0.01.

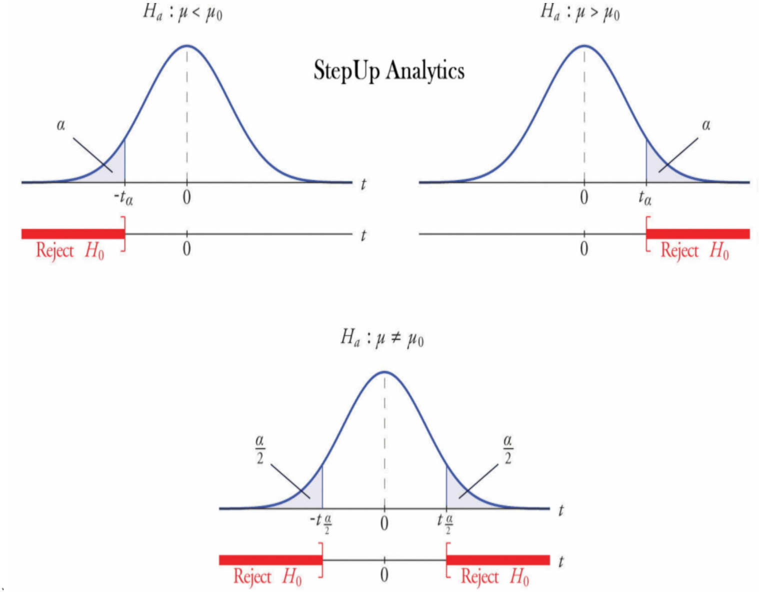
Result:(Single Mean (Variance Unknown) For the two-sided hypothesis

$$\begin{cases} H_0 : \mu = \mu_0, \\ H_1 : \mu \neq \mu_0. \end{cases}$$

We reject H_0 at significance level α when the computed t-statistic

$$T = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

exceeds $t_{\frac{\alpha}{2},n-1}$ or is less than $-t_{\frac{\alpha}{2},n-1}$.
For $H_1 : \mu > \mu_0$, rejection results when $t > t_{\alpha,n-1}$. For $H_1 : \mu < \mu_0$, the critical region is given by $t < -t_{\alpha,n-1}$.



$H_a : \mu < \mu_0$
StepUp Analytics
$H_a : \mu > \mu_0$
α
$-t_\alpha$
0
t
Reject H_0
α
t_α
Reject H_0
$H_a : \mu \neq \mu_0$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$-t_{\frac{\alpha}{2}}$
$t_{\frac{\alpha}{2}}$
Reject H_0
Reject H_0

Example

The Edison Electric Institute has published figures on the number of kilowatt hours used annually by various home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

Solution

1)

$$\begin{cases} H_0: \mu = 46 \text{ kilowatt hours,} \\ H_1: \mu < 46 \text{ kilowatt hours.} \end{cases}$$

2) $\alpha = 0.05$.

3) Critical region: $t < -t_{\alpha,n-1}$, where $T = (\overline{X} - \mu_0)/(S/\sqrt{n})$ with 11 degrees of freedom and $t_{\alpha,n-1} = 1.796$.

4) Computations: $\overline{X} = 42$ kilowatt hours, $s = 11.9$ kilowatt hours, and $n = 12$.
Hence, $t = (42 - 46)/(11.9/\sqrt{12}) = -1.16$.
5) Since $t > -1.796$, we do not reject H_0 and conclude that the average number of kilowatt hours used annually by home vacuum cleaners is not significantly less than 46.
Also P-value $= \Pr(T < -1.16) = \Pr(T > 1.16) \approx 0.135$.

Plan

The two-sided hypothesis on two means can be written generally as $H_0 : \mu_1 = \mu_2$. For σ_1 and σ_2 known, the test statistic is given by

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

That is, reject H_0 in favor of $H_1 : \mu_1 \neq \mu_2$ if $z > z_{\alpha/2}$, and $z < -z_{\alpha/2}$.
One-tailed critical regions are used in the case of the one-sided alternatives. The reader should, as before, study the test statistic and be satisfied that for, say, $H_1 : \mu_1 > \mu_2$, the signal favoring H_1 comes from large values of z. Thus, the upper-tailed critical region applies.

Unknown But Equal Variances

If we assume that both distributions are normal and that $\sigma_1 = \sigma_2 = \sigma$, the two-sample t-test may be used. The test statistic is given by the following test procedure.

Result:(Two-Sample Pooled t-Test)

For the two-sided hypothesis

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2, \end{aligned}$$

we reject H_0 at significance level α when the computed t-statistic

$$T = \frac{\overline{X}_1 - \overline{X}_2}{S_p\sqrt{1/n_1 + 1/n_2}}$$

where

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

exceeds $t_{\alpha/2, n_1+n_2-2}$ or is less than $-t_{\alpha/2, n_1+n_2-2}$.

For $H_1 : \mu_1 > \mu_2$, reject $H_0 : \mu_1 = \mu_2$ when $t > t_{\alpha, n_1+n_2-2}$. For $H_1 : \mu_1 < \mu_2$, reject $H_0 : \mu_1 = \mu_2$ when $t < -t_{\alpha, n_1+n_2-2}$.

Example

An experiment was performed to compare the abrasive wear of two materials. Twelve pieces of material 1 were tested and ten pieces of material 2 were similarly tested. The samples of material 1 gave an average (coded) wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 83 with a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2? Assume the populations to be approximately normal with equal variances.

Solution

Let μ_1 and μ_2 represent the population means of the abrasive wear for material 1 and material 2, respectively. 1)

$$\begin{cases} H_0 : \mu_1 = \mu_2, \\ H_1 : \mu_1 > \mu_2. \end{cases}$$

2) $\alpha = 0.05$.
3) Critical region: $t > t_{\alpha, n_1+n_2-2}$, where

$$T = \frac{\overline{X}_1 - \overline{X}_2}{S_p\sqrt{1/n_1 + 1/n_2}}$$

with $\nu = 20$ degrees of freedom and $t_{\alpha, n_1+n_2-2} = t_{0.05,20} = 1.725$.
4) Computations:

$$\begin{array}{lll} \overline{X}_1 = 58 & S_1 = 4 & n_1 = 12 \\ \overline{X}_2 = 81 & S_2 = 5 & n_2 = 10 \end{array}$$

Hence

$$\begin{aligned} S_p &= \sqrt{\frac{(11)(16)+(9)(25)}{12+10-2}} = 4.478 \\ t &= \frac{(85-83)}{4.478\sqrt{1/12+1/10}} = 1.04 < 1.725 \\ \text{P-value} &= \Pr(T > 1.04) \approx 0.16. (\text{See Table A.4.}) \end{aligned}$$

5) Decision: Do not reject H_0. We are unable to conclude that the abrasive wear of material 1 exceeds that of material 2.

Testing of two means can be accomplished when data are in the form of paired observations, as discussed in Chapter 4. The statistical test for two means μ_1 and μ_2 in the situation with paired observations is based on the random variable

$$T = \frac{\overline{D} - \mu_D}{S_D/\sqrt{n}}$$

where $\overline{D}$ and S_D are random variables representing the sample mean and standard deviation of the differences of the observations in the experimental units. As in the case of the pooled t-test, the assumption is that the observations from each population are normal. This two-sample problem is essentially reduced to a one-sample problem by using the computed differences $d_1, d_2, \ldots, d_n$. Critical regions are constructed using the t-distribution with $n-1$ degrees of freedom.

Example **Blood Sample Data**

A study was conducted to examine the influence of the drug succinylcholine on the circulation levels of androgens in the blood. Blood samples were taken from wild, free-ranging deer immediately after they had received an intramuscular injection of succinylcholine administered using darts and a capture gun. A second blood sample was obtained from each deer 30 minutes after the first sample, after which the deer was released. The levels of androgens at time of capture and 30 minutes later, measured in nanograms per milliliter (ng/mL), for 15 deer are given in Table 6.2. Assuming that the populations of androgen levels at time of injection and 30 minutes later are normally distributed, test at the 0.05 level of significance whether the androgen concentrations are altered after 30 minutes.

	Androgen (ng/mL)		
Deer	At Time of Injection	30 Minutes after Injection	d_i
1	2.76	7.02	4.26
2	5.18	3.10	−2.08
3	2.68	5.44	2.76
4	3.05	3.99	0.94
5	4.10	5.21	1.11
6	7.05	10.26	3.21
7	6.60	13.91	7.31
8	4.79	18.53	13.74
9	7.39	7.91	0.52
10	7.30	4.85	−2.45
11	11.78	11.10	−0.68
12	3.90	3.74	−0.16
13	26.00	94.03	68.03
14	67.48	94.03	26.55
15	17.04	41.70	24.66

Solution

Let μ_1 and μ_2 be the average androgen concentration at the time of injection and 30 minutes later, respectively. We proceed as follows:

1. $H_0 : \mu_1 = \mu_2$ or $\mu_D = \mu_1 - \mu_2 = 0$.
2. $H_1 : \mu_1 \neq \mu_2$ or $\mu_D = \mu_1 - \mu_2 \neq 0$.
3. $\alpha = 0.05$.
4. Critical region: $t < -t_{\alpha/2,n-1}$ and $t > t_{\alpha/2,n-1}$, where $t = \frac{\overline{D}}{S_D/\sqrt{n}}$ with $\nu = 14$ degrees of freedom and $t_{\alpha/2,n-1} = 2.145$.
5. Computations: The sample mean and standard deviation for the d_i are

$$\overline{D} = 9.848 \quad \text{and} \quad S_D = 18.474$$

Therefore

$$t = \frac{9.848}{18.474/\sqrt{15}} = 2.06$$

6. Hence $-2.145 < t = 2.06 < 2.145$. Though the t-statistic is not significant at the 0.05 level, from Table A.4,

$$P = P(|T| > 2.06) \approx 0.06$$

As a result, there is no evidence that there is a difference in mean circulating levels of androgen.

Plan

Tests of hypotheses concerning proportions are required in many areas. We now consider the problem of testing the hypothesis that the proportion of successes in a binomial experiment equals some specified value. That is, we are testing the null hypothesis H_0 that $p = p_0$, where p is the parameter of the binomial distribution. The alternative hypothesis may be one of the usual one-sided or two-sided alternatives:

$$p < p_0 \quad p > p_0 \quad \text{or} \quad p \neq p_0$$

We know that if $np_0 \geq 5$ and $n(1 - p_0) \geq 5$, then the random variable $\widehat{P}$ is approximately a normal distribution with mean p_0 and standard deviation $\sigma_{\widehat{P}} = \sqrt{p_0(1 - p_0)/n}$.

The **z-value for testing** $\mathbf{p} = \mathbf{p}_0$ is given by

$$z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}$$

Hence, for a two-tailed test at the α-level of significance, the critical region is $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$. For the one-sided alternative $p < p_0$, the critical region is $z < -z_{\alpha}$, and for the alternative $p > p_0$, the critical region is $z > z_{\alpha}$.

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.

Solution

1)

$$\begin{cases} H_0: \ p = 0.6, \\ H_1: \ p > 0.6. \end{cases}$$

2) $\alpha = 0.05$.
3) Critical region: $Z > z_\alpha$, where $z_\alpha = 1.645$. Then, the critical region: $z > 1.645$.
4) Computations: $x = 70, n = 100$, $\widehat{p} = 70/100 = 0.7$, and

$$\begin{aligned} z &= \frac{0.7 - 0.6}{\sqrt{\frac{(0.6)(0.4)}{100}}} = 2.04 \\ z &= 2.04 > 1.645 \\ P\text{-value} &= \Pr(Z > 2.04) < 0.0207. \end{aligned}$$

5) Decision: Reject H_0 and conclude that the new drug is superior.

Plan

Two Samples: Tests on Two Proportions

Situations often arise where we wish to test the hypothesis that two proportions are equal. That is, we are testing $p_1 = p_2$ against one of the alternatives $p_1 < p_2$, $p_1 > p_2$, or $p_1 = p_2$. The statistic on which we base our decision is the random variable $\widehat{P}_1 - \widehat{P}_2$. When $H_0 : p_1 = p_2 \ (= p)$ is true, we know that

$$Z = \frac{\widehat{P}_1 - \widehat{P}_2}{\sqrt{pq(1/n_1 + 1/n_2)}}$$

To compute a value of Z, however, we must estimate the parameters p and q that appear in the radical. Under H_0, both $\widehat{P}_1$ and $\widehat{P}_2$ are estimators of p.

We use the pooled estimate of the proportion p, which is

$$\widehat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

where x_1 and x_2 are the numbers of successes in each of the two samples. Substituting $\widehat{p}$ for p and $\widehat{q} = 1 - \widehat{p}$ for q, the z-value for testing $p_1 = p_2$ is determined from the formula

$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}\widehat{q}(1/n_1 + 1/n_2)}}$$

The critical regions for the appropriate alternative hypotheses are set up as before, using critical points of the standard normal curve.

Example

A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. To determine if there is a significant difference in the proportions of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county residents favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use an $\alpha = 0.05$ level of significance.

Solution

Let p_1 and p_2 be the true proportions of voters in the town and county, respectively, favoring the proposal.
$\widehat{p}_1 = x_1/n_1 = 120/200 = 0.6$, $\widehat{p}_2 = x_2/n_2 = 240/500 = 0.48$, and the pooled estimate
$\widehat{p} = (x_1 + x_2)/(n_1 + n_2) = (120 + 240)/(200 + 500) = 0.51.$
1)

$$\begin{cases} H_0 : \; p_1 = p_2 \\ H_1 : \; p_1 > p_2. \end{cases}$$

2) $\alpha = 0.05.$
3) The test statistic

$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}\widehat{q}(1/n_1 + 1/n_2)}} = \frac{0.60 - 0.48}{(0.51)(0.49)(1/200 + 1/500)} = 2.9$$

4) Critical region: $z > 1.645$. P-value $= P(Z > 2.9) = 0.0019.$
5) Decision: Reject H_0 and agree that the proportion of town voters favouring the proposal is higher than the proportion of county voters.

Plan

In this section, we are concerned with testing hypotheses concerning population variances or standard deviations. Let us first consider the problem of testing the null hypothesis H_0 that the population variance σ^2 equals a specified value σ_0^2 against one of the usual alternatives $\sigma^2 < \sigma_0^2$, $\sigma^2 > \sigma_0^2$, or $\sigma^2 \neq \sigma_0^2$. If we assume that the distribution of the population being sampled is normal, the chi-squared value for testing $\sigma^2 = \sigma_0^2$ is given by

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

where n is the sample size, s^2 is the sample variance, and σ_0^2 is the value of σ^2 given by the null hypothesis. If H_0 is true, χ^2 is a value of the chi-squared distribution with $\nu = n - 1$ degrees of freedom.

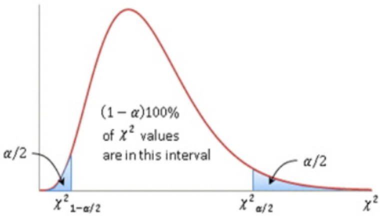


Figure 5-2: Figure 5.2

Hence, for a two-tailed test at the α-level of significance, the critical region is $\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$ (see figure 5.2). For the one-sided alternative $\sigma^2 < \sigma_0^2$, the critical region is $\chi^2 < \chi^2_{1-\alpha}$, and for the one-sided alternative $\sigma^2 > \sigma_0^2$, the critical region is $\chi^2 > \chi^2_{\alpha}$.

Example

A manufacturer of car batteries claims that the life of the company's 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

1)

$$\begin{cases} H_0 : \sigma^2 = 0.81, \\ H_1 : \sigma^2 > 0.81. \end{cases}$$

2) $\alpha = 0.05$.

3) Critical region: The null hypothesis is rejected when $\chi^2 > 16.919$, where $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$, with $\nu = 9$ degrees of freedom.

4) Computations: $s^2 = 1.44$, $n = 10$, and

$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16.0, \; P \approx 0.07.$$

5) Decision: The χ^2-statistic is not significant at the 0.05 level. However, based on the P-value 0.07, there is evidence that $\sigma > 0.9$.

Now let us consider the problem of testing the equality of the variances σ_1^2 and σ_2^2 of two populations. That is, we shall test the null hypothesis H_0 that $\sigma_1^2 = \sigma_2^2$ against one of the usual alternatives $\sigma_1^2 < \sigma_2^2$, $\sigma_1^2 > \sigma_2^2$, or $\sigma_1^2 \neq \sigma_2^2$. For independent random samples of sizes n_1 and n_2, respectively, from the two populations, the f-value for testing $\sigma_1^2 = \sigma_2^2$ is the ratio

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

where s_1^2 and s_2^2 are the variances computed from the two samples. If the two populations are approximately normally distributed and the null hypothesis is true, then the ratio $f = s_1^2/s_2^2$ is a value of the F-distribution with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom. Therefore, the critical regions of size α corresponding to the one-sided alternatives $\sigma_1^2 < \sigma_2^2$ and $\sigma_1^2 > \sigma_2^2$ are, respectively, $f < f_{1-\alpha}(\nu 1, \nu 2)$ and $f > f_{\alpha}(\nu 1, \nu 2)$. For the two-sided alternative $\sigma_1^2 \neq \sigma_2^2$, the critical region is $f < f_{1-\alpha/2}(\nu 1, \nu 2)$ or $f > f_{\alpha/2}(\nu 1, \nu 2)$.

Example

In testing for the difference in the abrasive wear of the two materials in Example 166, we assumed that the two unknown population variances were equal. Were we justified in making this assumption? Use a 0.10 level of significance.

Solution

Let σ_1^2 and σ_2^2 be the population variances for the abrasive wear of material 1 and material 2, respectively.

1)

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2, \\ H_1 : \sigma_1^2 \neq \sigma_2^2. \end{cases}$$

2) $\alpha = 0.10$.

3) Critical region: We have $f_{0.05}(11, 9) = 3.11$, and, by using Theorem 99, we find $f_{0.95}(11, 9) = \frac{1}{f0.05(9,11)} = 0.34$. Therefore, the null hypothesis is rejected when $f < 0.34$ or $f > 3.11$, where $f = s_1^2/s_2^2$ with $\nu_1 = 11$ and $\nu_2 = 9$ degrees of freedom.

4) Computations: $s_1^2 = 16$, $s_2^2 = 25$, hence $f = 16/25 = 0.64$.

5) Decision: Do not reject H_0. Conclude that there is insufficient evidence that the variances differ.