

1. Hypothesis testing for one sample

EXAMPLE 1.1

Canon, Inc., introduced a copying machine that features two-color copying capability in a compact system copier. The average speed of the standard compact system copier is 27 copies per minute (as advertised in national business magazines and elsewhere). Suppose that the company wants to test whether the new two-color copier has the same average speed as its standard compact copier and it conducts a test of 24 runs of the new machines, giving a sample mean $\bar{x} = 24.6$ and sample standard deviation $s = 7.4$ (copies per minute). Using the significance level $\alpha = 0.05$, is there evidence to conclude that the average speed of the new machine is different from that of the standard model?

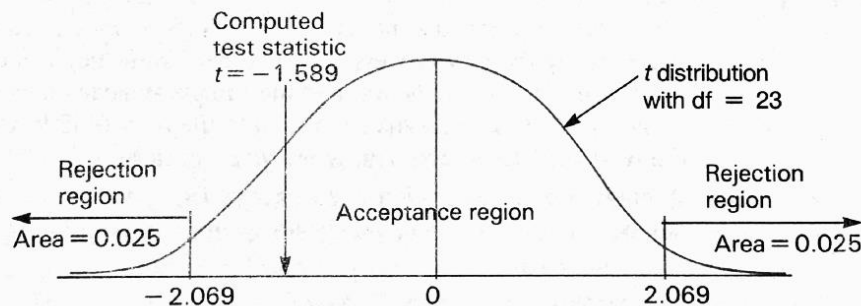
SOLUTION:

We want to test $H_0: \mu = 27$ versus $H_1: \mu \neq 27$. Since the sample size is $n = 24$, we use the t distribution, making the usual implicit assumption about the shape of the population. The critical points for $\alpha = 0.05$, found from Table 3 using $df = 23$, are ± 2.069 . We now compute the value of our test statistic.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{24.6 - 27}{7.4/4.899} = -1.589$$

Is this result significant? No. Since the test statistic falls inside the acceptance region for $\alpha = 0.05$, we accept the null hypothesis. There is no evidence to conclude that the average speed of the new copier is different from the average speed of the standard compact copier. Our statistical hypothesis test is shown in Figure 1

FIGURE 1 Carrying out the Test of Example 1



EXAMPLE 1.2

The Environmental Protection Agency (EPA) sets limits on the concentrations of pollutants emitted by various industries. Suppose that the upper allowable limit on the emission of vinyl chloride is set at an average of 55 parts per million (ppm) within a range of two miles around the plant emitting this chemical. To check compliance with this rule, the EPA collects a random sample of 100 readings at different times and dates within the two-mile range around the plant. The findings are that the sample average concentration is 60 ppm and the sample standard deviation is 20 ppm. Is there evidence to conclude that the plant in question is violating the law?

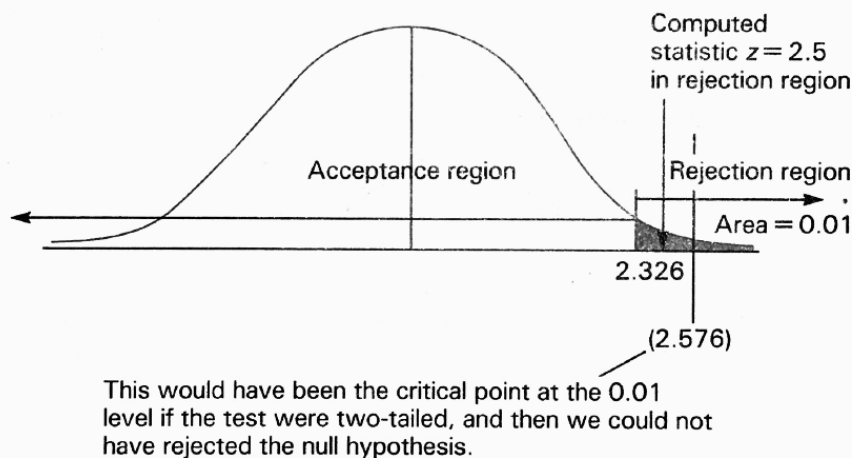
SOLUTION:

The EPA is interested in determining whether or not the plant operators are violating the law. The EPA will be taking action against the plant operators only if there is enough evidence to conclude that the average concentration of vinyl chloride within the given range of two miles from the plant is *above* the allowed average of 55 ppm. Therefore, the test is a right-hand-tailed test. Since the sample size is 100, we use the normal distribution. If a level of significance $\alpha = 0.01$ is desired, the critical point to be used is +2.326. The test statistic is the same as that used in the two-tailed test and is given by Equation 1

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{60 - 55}{20/10} = 2.5$$

As can be seen from Figure 6-15, the value of the test statistic falls in the rejection region for $\alpha = 0.01$, and the EPA may therefore reject the null hypothesis of “innocence” (nonviolation of the emission control rule) in favor of the alternative hypothesis that the plant operators are indeed violating the law. By using $\alpha = 0.01$, the EPA is willing to take at most a 0.01 chance of concluding that the plant operators are guilty when in reality they are innocent.

FIGURE 2 The One-Tailed Test in Example (h), Compared with a Two-Tailed Test



The elements of a large-sample, two-tailed (standardized) test for the population mean, μ .

The null hypothesis:

$$H_0: \mu = \mu_0 \quad (0)$$

The alternative hypothesis:

$$H_1: \mu \neq \mu_0$$

The significance level of the test:

$$\alpha \text{ (often, } \alpha = 0.05 \text{ or } 0.01)$$

The test statistic:

$$z = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \quad \begin{array}{l} \text{(assuming } \sigma \\ \text{is unknown;} \\ \text{otherwise,} \\ \text{use } \sigma/\sqrt{n} \text{ in} \\ \text{the denominator)} \end{array} \quad (1)$$

The critical points:

These depend on α . They are the bounds $\pm z_{\alpha/2}$ that capture between them an area of $1 - \alpha$. (When $\alpha = 0.05$, the critical points are ± 1.96 ; when $\alpha = 0.01$, the critical points are ± 2.576 . For other values of α , the critical points may be obtained from the standard normal table.)

The decision rule:

Reject the null hypothesis if either $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$, where $z_{\alpha/2}$ and $-z_{\alpha/2}$ are the two critical points. Otherwise, accept the null hypothesis.

2. Hypothesis testing for two samples

EXAMPLE 2.1

Suppose that the makers of Duracell batteries want to prove that their size AA battery lasts an average of at least 45 minutes longer than Duracell's main competitor, the Energizer. Two independent random samples of 100 batteries of each kind are selected, and the batteries are run continuously until they are no longer operational. The sample average life for Duracell is found to be $\bar{x}_1 = 308$ minutes, and the sample standard deviation is $s_1 = 84$ minutes. The results for the Energizer batteries are $\bar{x}_2 = 254$ minutes and $s_2 = 67$ minutes. Is there evidence to substantiate Duracell's claim that their batteries last, on average, at least 45 minutes longer than Energizer batteries of the same size?

SOLUTION:

Our null and alternative hypotheses are:

$$H_0: \mu_1 - \mu_2 \leq 45$$

$$H_1: \mu_1 - \mu_2 > 45$$

The makers of Duracell hope to prove their claim by rejecting the null hypothesis. Recall that accepting a null hypothesis is not a strong conclusion. This is why—in order to prove that Duracell batteries last an average of at least 45 minutes longer—the claim to be proved is stated as the *alternative* hypothesis.

The value of the test statistic in this case is computed as follows:

$$z = \frac{(308 - 254) - 45}{\sqrt{\frac{84^2}{100} + \frac{67^2}{100}}} = 0.838$$

This value falls in the acceptance region of our right-hand-tailed test at any conventional level of significance, α . The p -value is equal to 0.2011. We must conclude that there is insufficient evidence to support Duracell's claim.

Large-sample test statistic for the difference between two population means:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (2)$$

The term $(\mu_1 - \mu_2)_0$ is the difference between μ_1 and μ_2 under the null hypothesis. It is equal to zero in situations I and II and it is equal to the prespecified value D in situation III. The term in the denominator is the standard error of the difference between the two sample means (it relies on the assumption that the two samples are independent).