Steps in the Test of a Hypothesis Using the *P*-Value Approach (cont.)

Step 5: Calculate the *P*-value using the test statistic. For the sake of this setup, suppose we are performing a hypothesis test on the mean when σ is known. Thus, the observed value of the test statistic will be

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

The *P*-value will be found as follows:

a. If the alternative hypothesis is H_a: $\mu < \mu_0$, then the *P*-value is calculated as $P(z \le z_0)$.

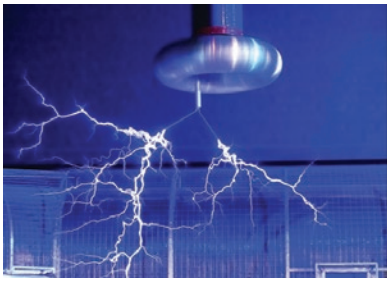
b. If the alternative hypothesis is H_a: $\mu > \mu_0$, then the *P*-value is calculated as $P(z \ge z_0)$.

c. If the alternative hypothesis is H_a: $\mu \ne \mu_0$, then the *P*-value is calculated as $2P(z \ge |z_0|)$.

Note that in this example, we are performing a test on the population mean with σ known. If the parameter that is being tested changes so that the test statistic changes, then we would use the appropriate test statistic in the above probability statements.

Step 6: Make the decision and state the conclusion in terms of the original question. The decision-making process is as follows.

- If the *P*-value is less than or equal to α, reject the null hypothesis in favor of the alternative hypothesis.
- If the *P*-value is greater than α, fail to reject the null hypothesis.



Different *P*-Values for Different Folks

In particle physics, the standard for "discovery" is a *P*-value less than 0.0000003. That is the probability which corresponds to observing a value that is at least 5 standard deviations from the mean for a one-tailed test. Particle physicists consider a *P*-value less than 0.003, which is the probability of observing a value at least 2.75 standard deviations from the mean, "evidence of a particle"—an encouraging result, but not "discovery".

The common significance levels we have used in this book are $\alpha = 0.05$ and $\alpha = 0.01$ which correspond to a value at least 1.645 and 2.33 standard deviations from the mean, respectively. This goes to show you that there is not any one significance level that everyone agrees on. Difference disciplines have different comfort levels with the idea of significance.

10.2 Exercises

Basic Concepts

1. What is the rationale for the *z*-statistic?
2. What are the three key questions to be asked in the hypothesis testing procedure in order to determine which test statistic is appropriate?
3. Describe the distribution of the *z*-test statistic.
4. What are critical values? How do critical values influence the decision rule in the hypothesis testing procedure?

Exercises

5. Determine the critical value(s) of the test statistic for each of the following tests for the population mean when the population standard deviation is known.

   a. Left-tailed test, $\alpha = 0.01$
   b. Right-tailed test, $\alpha = 0.10$
   c. Two-tailed test, $\alpha = 0.05$

6. Determine the critical value(s) of the test statistic for each of the following tests for the population mean when the population standard deviation is known.

   a. Left-tailed test, $\alpha = 0.05$
   b. Right-tailed test, $\alpha = 0.02$
   c. Two-tailed test, $\alpha = 0.08$

7. A random sample of 1000 observations produces a sample mean of 53.5 with a population standard deviation of 5.3. Test the hypothesis that the mean is not equal to 55 at $\alpha = 0.05$.

8. A random sample of 200 observations indicate a sample mean of 4117 with a population standard deviation of 300. Test the hypothesis that the mean is greater than 4100 at $\alpha = 0.01$.

9. The head of the Veterans Administration has been receiving complaints from a Vietnam veterans' organization concerning disability checks. The organization claims that checks are continually late. The checks are supposed to arrive no later than the tenth of each month. The administrator randomly selects 100 disabled veterans and measures the arrival time in relation to the tenth of the month for each check. If the check arrives early, it receives a negative value. For example, if the check arrives on the eighth of the month, it is measured as –2. If the check arrives on the twelfth of the month, it is measured as +2.

- **a.** What statistical measure should you use in your statement of hypothesis?
- **b.** Formulate hypotheses to test the veterans' organization's claim.
- **c.** Suppose in the sample of 100 disabled veterans receiving checks, the average number of days late was 1.2 with a population standard deviation of 1.4. Calculate the test statistic for your hypothesis.
- **d.** If the test is conducted at the 0.05 level, construct the decision rule for the test statistic.
- **e.** Is there overwhelming evidence at the 0.05 level that the checks arrive late?
- **f.** If you are the head of the Veterans Administration, what is your conclusion?

10. Hurricane Andrew swept through southern Florida causing billions of dollars of damage. Because of the severity of the storm and the type of residential construction used in this semitropical area, there was some concern that the average claim size would be greater than the historical average hurricane claim of \$24,000. Several insurance companies collaborated in a data gathering experiment. They randomly selected 84 homes and sent adjusters to settle the claims. In the sample of 84 homes, the average claim was \$27,500 with a population standard deviation of \$2400.

- **a.** What is the population being studied?
- **b.** What statistical measure should you use in your hypothesis?
- **c.** State your hypotheses.
- **d.** Test the hypothesis at the 0.01 level.
- **e.** Is there overwhelming evidence (at the 0.01 level) that home damage is greater than the historical average? Write your conclusion in the context of the original problem.

11. A retail computer store is considering offering a two-year service warranty, instead of its current one-year plan. In order to do this, they must determine the average service costs for their systems in the second year of operation. A committee of technicians, sales, and management staff believe that the average repair cost in the second year should be approximately \$50. Seventy-five customers who purchased machines between two and four years earlier are randomly selected. Tracking the service needs of these customers reveals an average service cost of \$38 with a population standard deviation of \$10.

- **a.** What is the population being studied?
- **b.** What variable is being measured in this problem?
- **c.** What level of measurement does the variable possess?
- **d.** Test the committee's claim at the 0.10 level.
- **e.** What concerns might you have about the data that were collected?

12. In preparation for upcoming wage negotiations with the union, the managers for the Bevel Hardware Company want to establish the time required to assemble a kitchen cabinet. A first line supervisor believes that the job should take 45 minutes on average to complete. A random sample of 125 cabinets has an average assembly time of 47 minutes with a population standard deviation of 10 minutes.

a. Is there overwhelming evidence to contradict the first line supervisor's belief at a 0.05 significance level? Make your conclusion using the *P*-value approach.

b. What is the lowest average assembly time that would allow the union to conclude that the supervisor is incorrect?

13. The Better Business Bureau has received several complaints that a flour company is underfilling its five pound bags of flour. The Bureau randomly selects 750 bags of flour and determines the weight of each bag. The sample average weight of the bags is 4.80 pounds with a population standard deviation of 0.15 pounds.

a. Is there overwhelming evidence at the 0.01 level that the bags are underfilled?

b. What is the lowest average bag weight that would allow the Bureau to conclude that the bags are underfilled?

14. A horticulturist working for a large plant nursery is conducting experiments on the growth rate of a new shrub. Based on previous research, the horticulturist feels the average daily growth rate of the new shrub is 1 cm per day. A random sample of 45 shrubs has an average growth of 0.90 cm per day with a population standard deviation of 0.30 cm. Will a test of hypothesis at the 0.05 significance level support the claim that the growth rate is less than 1 cm per day?

15. Del Valley Foods requires that corn supplied for canning must weigh more than 5 ounces per ear. South Valley Farms claims that the corn they supply meets the required specifications. 200 ears of corn are selected at random from a delivery. The sample has a mean of 5.01 ounces and a population standard deviation of 0.30 ounce. Will a test of hypothesis at $\alpha = 0.10$ support South Valley Farms' claim?

16. Government regulations restrict the amount of pollutants that can be released to the atmosphere through industrial smokestacks. To demonstrate that their smokestacks are releasing pollutants below the mandated limit of 5 parts per billion pollutants, REM Industries collects a random sample of 300 readings. The mean pollutant level for the sample is 4.85 parts per billion with a population standard deviation of 0.30 parts per billion. Do the data support the claim that the average pollutants produced by REM Industries are below the mandated level at a 0.01 significance level?

17. The director of the IRS has been flooded with complaints that people must wait more than 45 minutes before seeing an IRS representative. To determine the validity of these complaints, the IRS randomly selects 400 people entering IRS offices across the country and records the times that they must wait before seeing an IRS representative. The average waiting time for the sample is 55 minutes with a population standard deviation of 15 minutes.

a. What is the population being studied?

b. Are the complaints substantiated by the data at $\alpha = 0.10$?

18. The manufacturer of Brand X floor polish is developing a new polish that it hopes will dry faster than the competition's polish. The competition's polish is advertised to have an average drying time of 10 minutes. A random sample of 1000 Brand X polishes has an average drying time of 9.3 minutes with a population standard deviation of 0.5 minute. Based on the data, can the manufacturer conclude that the drying time for Brand X is faster than the competition's brand at a 0.05 significance level?

19. For each of the following combinations of the P-value and α, decide whether you would reject or fail to reject the null hypothesis.

a. P-value $= 0.0935$, $\alpha = 0.10$

b. P-value $= 0.0311$, $\alpha = 0.05$

c. P-value $= 0.0545$, $\alpha = 0.01$

d. P-value $= 0.0489$, $\alpha = 0.05$

20. Consider the following hypothesis tests for the population mean. Compute the P-value for each test and decide whether you would reject or fail to reject the null hypothesis at $\alpha = 0.05$.

a. H_0: $\mu = 15$, H_a: $\mu > 15$, $z = 1.58$

b. H_0: $\mu = 1.9$, H_a: $\mu < 1.9$, $z = -2.25$

c. H_0: $\mu = 100$, H_a: $\mu \neq 100$, $z = 1.90$

21. Consider the following hypothesis tests for the population mean. Compute the P-value for each test and decide whether you would reject or fail to reject the null hypothesis at $\alpha = 0.01$.

a. H_0: $\mu = 10$, H_a: $\mu > 10$, $z = 2.00$

b. H_0: $\mu = 82$, H_a: $\mu < 82$, $z = -2.45$

c. H_0: $\mu = 100$, H_a: $\mu \neq 100$, $z = 2.70$

10.3 Testing a Hypothesis about a Population Mean, σ Unknown

The hypothesis testing strategy from the previous section assumes that the population standard deviation, σ, is known. However, in most instances, the population standard deviation is just as *unknown* as the population mean. Despite the added uncertainty, the general approach to testing a hypothesis is the same provided that it is reasonable to *assume* that the population from which you are sampling is normal. Some of the technical details concerning the distribution of the test statistic change, since the sample standard deviation, s, will be used in place of the population standard deviation, σ. This modification will cause a change in the distribution of the test statistic.

Formula

t-Test Statistic

If the standard deviation of the population is unknown, but the distribution of the population is assumed to be normal, then the test statistic is given by

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

where

n is the sample size,

$\bar{x}$ is the sample mean,

μ_0 is the hypothesized value of the population mean, and

s is the sample standard deviation.

It should be noted that this formula is only valid if $\bar{x}$ is normally distributed.

The test statistic has a t-distribution with $n - 1$ degrees of freedom.