5. Before agreeing to purchase a large order of polyethylene sheaths for a particular type of high-pressure oil-filled submarine power cable, a company wants to see conclusive evidence that the true standard deviation of sheath thickness is less than .05 mm. What hypotheses should be tested, and why? In this context, what are the type I and type II errors?

6. Many older homes have electrical systems that use fuses rather than circuit breakers. A manufacturer of 40-amp fuses wants to make sure that the mean amperage at which its fuses burn out is in fact 40. If the mean amperage is lower than 40, customers will complain because the fuses require replacement too often. If the mean amperage is higher than 40, the manufacturer might be liable for damage to an electrical system due to fuse malfunction. To verify the amperage of the fuses, a sample of fuses is to be selected and inspected. If a hypothesis test were to be performed on the resulting data, what null and alternative hypotheses would be of interest to the manufacturer? Describe type I and type II errors in the context of this problem situation.

7. Water samples are taken from water used for cooling as it is being discharged from a power plant into a river. It has been determined that as long as the mean temperature of the discharged water is at most 150°F, there will be no negative effects on the river's ecosystem. To investigate whether the plant is in compliance with regulations that prohibit a mean discharge water temperature above 150°C, 50 water samples will be taken at randomly selected times and the temperature of each sample recorded. The resulting data will be used to test the hypotheses $H_0: \mu = 150^\circ$ versus $H_1: \mu > 150^\circ$. In the context of this situation, describe type I and type II errors. Which type of error would you consider more serious? Explain.

8. A regular type of laminate is currently being used by a manufacturer of circuit boards. A special laminate has been developed to reduce warpage. The regular laminate will be used on one sample of specimens and the special laminate on another sample, and the amount of warpage will then be determined for each specimen. The manufacturer will then switch to the special laminate only if it can be demonstrated that the true average amount of warpage for that laminate is less than for the regular laminate. State the relevant hypotheses, and describe the type I and type II errors in this context of this situation.

9. Two different companies have applied to provide cable television service in a certain region. Let $p$ denote the proportion of all potential subscribers who favor the first company over the second. Consider testing $H_0: p = .5$ versus $H_1: p \neq .5$ based on a random sample of 25 individuals. Let $X$ denote the number in the sample who favor the first company and $x$ represent the observed value of $X$.

   a. Which of the following rejection regions is most appropriate and why?
      
      $R_1 = \{x: x \leq 7 \text{ or } x \geq 18\}$, $R_2 = \{x: x \leq 8\}$, $R_3 = \{x: x \geq 17\}$

   b. In the context of this problem situation, describe what the type I and type II errors are.

   c. What is the probability distribution of the test statistic $X$ when $H_0$ is true? Use it to compute the probability of a type I error.

   d. Compute the probability of a type II error for the selected region when $p = .3$, again when $p = .4$, and also for both $p = .6$ and $p = .7$.

   e. Using the selected region, what would you conclude if 6 of the 25 queried favored company 1?

10. A mixture of pulverized fuel ash and Portland cement to be used for grouting should have a compressive strength of more than 1300 KN/m$^2$. The mixture will not be used unless experimental evidence indicates conclusively that the strength specification has been met. Suppose compressive strength for specimens of this mixture is normally distributed with $\sigma = 60$. Let $\mu$ denote the true average compressive strength.

   a. What are the appropriate null and alternative hypotheses?

   b. Let $\bar{X}$ denote the sample average compressive strength for $n = 20$ randomly selected specimens. Consider the test procedure with test statistic $\bar{X}$ and rejection region $\bar{X} \geq 1331.26$. What is the probability distribution of the test statistic when $H_0$ is true? What is the probability of a type I error for the test procedure?

   c. What is the probability distribution of the test statistic when $\mu = 1350$? Using the test procedure of part (b), what is the probability that the mixture will be judged unsatisfactory when in fact $\mu = 1350$ (a type II error)?

   d. How would you change the test procedure of part (b) to obtain a test with significance level .05? What impact would this change have on the error probability of part (c)?

   e. Consider the standardized test statistic $Z = (\bar{X} - 1300)/(\sigma/\sqrt{n}) = (\bar{X} - 1300)/13.42$. What are the values of $Z$ corresponding to the rejection region of part (b)?

11. The calibration of a scale is to be checked by weighing a 10-kg test specimen 25 times. Suppose that the results of different weighings are independent of one another and that the weight on each trial is normally distributed with $\sigma = 200$ kg. Let $\mu$ denote the true average weight reading on the scale.

   a. What hypotheses should be tested?

   b. Suppose the scale is to be recalibrated if either $\bar{x} \geq 10.1032$ or $\bar{x} \leq 9.8968$. What is the probability that recalibration is carried out when it is actually unnecessary?

   c. What is the probability that recalibration is judged unnecessary when in fact $\mu = 10.17$ when $\mu = 9.86$?

   d. Let $z = (\bar{x} - 10)/(\sigma/\sqrt{n})$. For what value of $c$ is the rejection region of part (b) equivalent to the "two-tailed" region of either $z \geq c$ or $z \leq -c$?

   e. If the sample size were only 10 rather than 25, how should the procedure of part (d) be altered so that $\alpha = .05$?

   f. Using the test of part (e), what would you conclude from the following sample data?

   
16. Let the test statistic be the t distribution when H₀ is true. Give the significance level for each of the following situations:
   a. H₀: μ > μ₀, df = 15, rejection region t ≥ 3.733
   b. H₀: μ < μ₀, n = 24, rejection region t ≤ −2.500
   c. H₀: μ ≠ μ₀, n = 31, rejection region t ≥ 1.697 or t ≤ −1.697

17. Answer the following questions for the tire problem in Example 8.7.
   a. If x = 30.960 and a level α = .01 test is used, what is the decision?
   b. If a level .01 test is used, what is β(30,500)?
   c. If a level .01 test is used and it is also required that β(30,500) = .05, what sample size n is necessary?
   d. If x = 30.960, what is the smallest α at which H₀ can be rejected (based on n = 16)?

18. Reconsider the paint-drying situation of Example 8.2, in which drying time for a test specimen is normally distributed with σ = 9. The hypotheses H₀: μ = 75 versus H₁: μ < 75 are to be tested using a random sample of n = 25 observations.
   a. How many standard deviations (of X̄) below the null value is X̄ = 72.3?
   b. If X̄ = 72.3, what is the conclusion using α = .01?
   c. What is α for the test procedure that rejects H₀ when z = −2.88?
   d. For the test procedure of part (c), what is β(70)?
   e. If the test procedure of part (c) is used, what n is necessary to ensure that β(70) = .01?
   f. If a level .01 test is used with n = 100, what is the probability of a type II error when μ = 76?

19. The melting point of each of 16 samples of a certain brand of hydrogenated vegetable oil was determined, resulting in X̄ = 94.32. Assume that the distribution of the melting point is normal with σ = 1.20.
   a. Test H₀: μ = 95 versus H₁: μ ≠ 95 using a two-tailed level .01 test.
   b. If a level .01 test is used, what is β(94), the probability of a type II error when μ = 94?
   c. What value of n is necessary to ensure that β(94) = .1 when α = .01?

20. Lightbulbs of a certain type are advertised as having an average lifetime of 750 hours. The price of these bulbs is very favorable, so a potential customer has decided to go ahead with a purchase arrangement unless it can be conclusively demonstrated that the true average lifetime is smaller than what is advertised. A random sample of 50 bulbs was selected, the lifetime of each bulb determined, and the appropriate hypotheses were tested using Minitab, resulting in the accompanying output.

```
Variable  N  Mean  StdDev  SEMean  Z  P-Value
lifetime  50  738.44  38.20  5.40  -2.14  0.031
```

What conclusion would be appropriate for a significance level of .05? A significance level of .01? What significance level and conclusion would you recommend?

21. The true average diameter of ball bearings of a certain type is supposed to be .5 in. A one-sample t test will be carried out to see whether this is the case. What conclusion is appropriate in each of the following situations?
   a. n = 13, t = 1.6, α = .05
   b. n = 13, t = −1.6, α = .05
   c. n = 25, t = −2.6, α = .01
   d. n = 25, t = −3.9

22. The article "The Foreman's View of Quality Control" (Quality Eng., 1990: 257–280) described an investigation into the coating weights for large pipes resulting from a galvanized coating process. Production standards call for a true average weight of 200 lb per pipe. The accompanying descriptive summary and boxplot are from Minitab.

```
Variable  N  Mean  Median  TrMean  StDev  SEMean  ctg wt
          30  206.73  206.00  206.81  6.35  1.16
```

```
Variable  N  Min  Max  Q1  Q3
          30  193.00  218.00  202.75  212.00
```

Coating weight

```
190  200  210  220
```

a. What does the boxplot suggest about the status of the specification for true average coating weight?
   b. A normal probability plot of the data was quite straight. Use the descriptive output to test the appropriate hypotheses.

23. Exercise 36 in Chapter 1 gave n = 26 observations on escape time (sec) for oil workers in a simulated exercise, from which the sample mean and sample standard deviation are 370.69 and 24.36, respectively. Suppose the investigators had believed a priori that true average escape time would be at most 6 min. Does the data contradict this prior belief? Assuming normality, test the appropriate hypotheses using a significance level of .05.

24. Reconsider the sample observations on stabilized viscosity of asphalt specimens introduced in Exercise 46 in Chapter 1 (2781, 2900, 3013, 2856, and 2888). Suppose that for a particular application it is required that true average viscosity be 3000. Does this requirement appear to have been satisfied? State and test the appropriate hypotheses.

25. The desired percentage of SiO₂ in a certain type of aluminous cement is 5.5. To test whether the true average percentage is 5.5 for a particular production facility, 16 independently obtained samples are analyzed. Suppose that the percentage of SiO₂ in a sample is normally distributed with σ = .3 and that X̄ = 5.25.
   a. Does this indicate conclusively that the true average percentage differs from 5.5? Carry out the analysis using the sequence of steps suggested in the text.
   b. If the true average percentage is μ = 5.6 and a level .01 test based on n = 10 is used, what is the probability of detecting this departure from H₀?
   c. What value of n is required to satisfy α = .01 and β(5.6) = .01?
be based on a random sample of size $n$ from a normal population distribution. What conclusion is appropriate in each of the following situations?

- a. $n = 15, t = 3.2, \alpha = .05$
- b. $n = 9, t = 1.8, \alpha = .01$
- c. $n = 24, t = -.2$

53. Let $\mu$ denote true average serum receptor concentration for all pregnant women. The average for all women is known to be 5.63. The article “Serum Transferrin Receptor for the Detection of Iron Deficiency in Pregnancy” (Amer. J. of Clinical Nutr., 1991: 1077–1081) reports that $P$-value > .10 for a test of $H_0: \mu = 5.63$ versus $H_1: \mu \neq 5.63$ based on $n = 176$ pregnant women. Using a significance level of .01, what would you conclude?

54. The article “Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice?” (Chance, Summer 2005, pp. 9–15) reported on an experiment to investigate whether wine tasters could distinguish between more expensive reserve wines and their regular counterparts. Wine was presented to tasters in four containers labeled A, B, C, and D, with two of these containing the reserve wine and the other two the regular wine. Each taster randomly selected three of the containers, tasted the selected wines, and indicated which of the three he/she believed was different from the other two. Of the $n = 855$ tasting trials, 346 resulted in correct distinctions (either the one reserve that differed from the two regular wines or the one regular wine that differed from the two reserves). Does this provide compelling evidence for concluding that tasters of this type have some ability to distinguish between reserve and regular wines? State and test the relevant hypotheses using the $P$-value approach. Are you particularly impressed with the ability of tasters to distinguish between the two types of wine?

55. An aspirin manufacturer fills bottles by weight rather than by count. Since each bottle should contain 100 tablets, the average weight per tablet should be 5 grains. Each of 100 tablets taken from a very large lot is weighed, resulting in a sample average weight per tablet of 4.87 grains and a sample standard deviation of .35 grain. Does this information provide strong evidence for concluding that the company is not filling its bottles as advertised? Test the appropriate hypotheses using $\alpha = .01$ by first computing the $P$-value and then comparing it to the specified significance level.

56. Because of variability in the manufacturing process, the actual yielding point of a sample of mild steel subjected to increasing stress will usually differ from the theoretical yielding point. Let $p$ denote the true proportion of samples that yield before their theoretical yielding point. If on the basis of a sample it can be concluded that more than 20% of all specimens yield before the theoretical point, the production process will have to be modified.

- a. If 15 of 60 specimens yield before the theoretical point, what is the $P$-value when the appropriate test is used, and what would you advise the company to do?

- b. If the true percentage of “early yields” is actually 50% (so that the theoretical point is the median of the yield distribution) and a level .01 test is used, what is the probability that the company concludes a modification of the process is necessary?

57. The article “Heavy Drinking and Polydrug Use Among College Students” (J. of Drug Issues, 2008: 445–466) stated that 51 of the 462 college students in a sample had a lifetime abstinence from alcohol. Does this provide strong evidence for concluding that more than 10% of the population sampled had completely abstained from alcohol use? Test the appropriate hypotheses using the $P$-value method. [Note: The article used more advanced statistical methods to study the use of various drugs among students characterized as light, moderate, and heavy drinkers.]

58. A random sample of soil specimens was obtained, and the amount of organic matter (%) in the soil was determined for each specimen, resulting in the accompanying data (from “Engineering Properties of Soil,” Soil Science, 1998: 93–102).

| 1.10 | 5.09 | 0.97 | 1.59 | 4.60 | 0.32 | 0.55 | 1.45 |
| 0.14 | 4.47 | 1.20 | 3.50 | 5.02 | 4.67 | 5.22 | 2.69 |
| 3.98 | 3.17 | 3.03 | 2.21 | 0.69 | 4.47 | 3.31 | 1.17 |
| 0.76 | 1.17 | 1.57 | 2.62 | 1.66 | 2.05 |

The values of the sample mean, sample standard deviation, and (estimated) standard error of the mean are 2.481, 1.616, and .295, respectively. Does this data suggest that the true average percentage of organic matter in such soil is something other than 3%? Carry out a test of the appropriate hypotheses at significance level .10 by first determining the $P$-value. Would your conclusion be different if $\alpha = .05$ had been used? [Note: A normal probability plot of the data shows an acceptable pattern in light of the reasonably large sample size.]

59. The accompanying data on cube compressive strength (MPa) of concrete specimens appeared in the article “Experimental Study of Recycled Rubber-Filled High-Strength Concrete” (Magazine of Concrete Res., 2009: 549–556):

| 112.3 | 97.0 | 92.7 | 86.0 | 102.0 |
| 99.2 | 95.8 | 103.5 | 89.0 | 86.7 |

- a. Is it plausible that the compressive strength for this type of concrete is normally distributed?
- b. Suppose the concrete will be used for a particular application unless there is strong evidence that true average strength is less than 100 MPa. Should the concrete be used? Carry out a test of appropriate hypotheses using the $P$-value method.

60. A certain pen has been designed so that true average writing lifetime under controlled conditions (involving the use of a writing machine) is at least 10 hours. A random sample of 18 pens is selected, the writing lifetime of each is determined, and a normal probability plot of the resulting data supports the use of a one-sample $t$ test.