15. (a) $\alpha = 0.01$; H_0: The distributions are the same; H_1: The distributions are different.
(b) Sample $\chi^2 = 3.559$; $d.f. = 8$.
(c) $0.100 < P\text{-value} < 0.900$.
(d) Do not reject H_0.
(e) At the 1% level of significance, the evidence is insufficient to conclude that the distribution of first nonzero digits in the accounting file does not follow Benford's Law.

16. (a) $\alpha = 0.01$; H_0: The distributions are the same; H_1: The distributions are different.
(b) Sample $\chi^2 = 13.44$; $d.f. = 5$.
(c) $0.010 < P\text{-value} < 0.025$.
(d) Do not reject H_0.
(e) At the 1% level of significance, the evidence is insufficient to conclude that the distribution of observed outcomes for the die is different from the expected distribution of a fair die. In other words, at the 1% level of significance, we cannot conclude that the die is not balanced.

17. (a) $P(0) \approx 0.179$; $P(1) \approx 0.308$; $P(2) \approx 0.265$; $P(3) \approx 0.152$; $P(r \geq 4) \approx 0.096$.
(b) For $r = 0$, $E \approx 16.11$; for $r = 1$, $E \approx 27.72$; for $r = 2$, $E \approx 23.85$; for $r = 3$, $E \approx 13.68$; for $r \geq 4$, $E \approx 8.64$.
(c) $\chi^2 \approx 12.55$ with $d.f. = 4$.
(d) $\alpha = 0.01$: H_0: The Poisson distribution fits; H_1: The Poisson distribution does not fit; $0.01 < P\text{-value} < 0.025$; do not reject H_0. At the 1% level of significance, we cannot say that the Poisson distribution does not fit the sample data.

18. (a) $P(0) \approx 0.061$; $P(1) \approx 0.170$; $P(2) \approx 0.238$; $P(3) \approx 0.222$; $P(4) \approx 0.156$; $P(r \geq 5) \approx 0.153$.
(b) For $r = 0$, $E \approx 6.1$; for $r = 1$, $E \approx 17.0$; for $r = 2$, $E \approx 23.8$; for $r = 3$, $E \approx 22.2$; for $r = 4$, $E \approx 15.6$; for $r \geq 5$, $E \approx 15.3$.
(c) $\chi^2 \approx 13.116$ with $d.f. = 5$.

Suppose that $n = 275$ numerical entries were drawn at random from a large accounting file of a major corporation. The first nonzero digits were recorded for the sample.

First nonzero digit	1	2	3	4	5	6	7	8	9
Sample frequency	83	49	32	22	25	18	13	17	16

Use a 1% level of significance to test the claim that the distribution of first nonzero digits in this accounting file follows Benford's law.

16. *Fair Dice: Uniform Distribution* A gambler complained about the dice. They seemed to be loaded! The dice were taken off the table and tested one at a time. One die was rolled 300 times and the following frequencies were recorded.

Outcome	1	2	3	4	5	6
Observed frequency O	62	45	63	32	47	51

Do these data indicate that the die is unbalanced? Use a 1% level of significance. *Hint:* If the die is balanced, all outcomes should have the same expected frequency.

17. *Highway Accidents: Poisson Distribution* A civil engineer has been studying the frequency of vehicle accidents on a certain stretch of interstate highway. Long-term history indicates that there has been an average of 1.72 accidents per day on this section of the interstate. Let r be a random variable that represents number of accidents per day. Let O represent the number of observed accidents per day based on local highway patrol reports. A random sample of 90 days gave the following information.

r	0	1	2	3	4 or more
O	22	21	15	17	15

(a) The civil engineer wants to use a Poisson distribution to represent the probability of r, the number of accidents per day. The Poisson distribution is

$$P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

where $\lambda = 1.72$ is the average number of accidents per day. Compute $P(r)$ for $r = 0, 1, 2, 3$, and 4 or more.

(b) Compute the expected number of accidents $E = 90P(r)$ for $r = 0, 1, 2, 3$, and 4 or more.

(c) Compute the sample statistic $\chi^2 = \Sigma \frac{(O - E)^2}{E}$ and the degrees of freedom.

(d) Test the statement that the Poisson distribution fits the sample data. Use a 1% level of significance.

18. *Bacteria Colonies: Poisson Distribution* A pathologist has been studying the frequency of bacterial colonies within the field of a microscope using samples of throat cultures from healthy adults. Long-term history indicates that there is an average of 2.80 bacteria colonies per field. Let r be a random variable that represents the number of bacteria colonies per field. Let O represent the number of observed bacteria colonies per field for throat cultures from healthy adults. A random sample of 100 healthy adults gave the following information.

r	0	1	2	3	4	5 or more
O	12	15	29	18	19	7

(d) $\alpha = 0.05$; H_0: The Poisson distribution fits; H_1: The Poisson distribution does not fit; $0.010 < P\text{-value} < 0.025$; reject H_0. At the 5% level of significance, the evidence is sufficient to say that the Poisson distribution does not fit the sample data.

(a) The pathologist wants to use a Poisson distribution to represent the probability of r, the number of bacteria colonies per field. The Poisson distribution is

$$P(r) = \frac{e^{-\lambda}\lambda^r}{r!}$$

where $\lambda = 2.80$ is the average number of bacteria colonies per field. Compute $P(r)$ for $r = 0, 1, 2, 3, 4$, and 5 or more.

(b) Compute the expected number of colonies $E = 100P(r)$ for $r = 0, 1, 2, 3, 4$, and 5 or more.

(c) Compute the sample statistic $\chi^2 = \Sigma\frac{(O - E)^2}{E}$ and the degrees of freedom.

(d) Test the statement that the Poisson distribution fits the sample data. Use a 5% level of significance.

SECTION 10.3

Testing and Estimating a Single Variance or Standard Deviation

FOCUS POINTS

- Set up a test for a single variance σ^2.
- Compute the sample χ^2 statistic.
- Use the χ^2 distribution to estimate a P-value and conclude the test.
- Compute confidence intervals for σ^2 or σ.

Testing σ^2

This section can be presented by itself as long as the overview of the χ^2 distribution at the beginning of Part I is used.

Many problems arise that require us to make decisions about variability. In this section, we will study two kinds of problems: (1) we will test hypotheses about the variance (or standard deviation) of a population, and (2) we will find confidence intervals for the variance (or standard deviation) of a population. It is customary to talk about variance instead of standard deviation because our techniques employ the sample variance rather than the standard deviation. Of course, the standard deviation is just the square root of the variance, so any discussion about variance is easily converted to a similar discussion about standard deviation.

Test of variance

Let us consider a specific example in which we might wish to test a hypothesis about the variance. Almost everyone has had to wait in line. In a grocery store, bank, post office, or registration center, there are usually several checkout or service areas. Frequently, each service area has its own independent line. However, many businesses and government offices are adopting a "single-line" procedure.

In this section, we present more applications of the chi-square distribution. Linking Concepts, Problem 1, can be a good topic for a class discussion in which students summarize various methods and applications they have learned in Sections 10.1, 10.2, and 10.3.

In a single-line procedure, there is only one waiting line for everyone. As any service area becomes available, the next person in line gets served. The old, independent-lines procedure has a line at each service center. An incoming customer simply picks the shortest line and hopes it will move quickly. In either procedure, the number of clerks and the rate at which they work is the same, so the average waiting time is the *same*. What is the advantage of the single-line procedure? The difference is in the *attitudes* of people who wait in the lines. A lengthy waiting line will be more acceptable, even though the average waiting time is the same, if the variability of waiting times is smaller. When the variability is small, the inconvenience of waiting (although it might not be reduced) does become more predictable. This means impatience is reduced and people are happier.

To test the hypothesis that variability is less in a single-line process, we use the chi-square distribution. The next theorem tells us how to use the sample and population variance to compute values of χ^2.