Quick Review
What do samples really tell us about the populations from which they are drawn? Are the results of an experiment meaningful, or are they just sampling error? Statistical inference based on our understanding of sampling models can help answer these questions. Here’s a brief summary of the key concepts and skills:

- **Sampling models** describe the variability of sample statistics using a remarkable result called the Central Limit Theorem.
  - When the number of trials is sufficiently large, proportions found in different samples vary according to an approximately Normal model.
  - When samples are sufficiently large, the means of different samples vary, with an approximately Normal model.
  - The variability of sample statistics decreases as sample size increases.
  - Statistical inference procedures are based on the Central Limit Theorem.
  - No inference procedure is valid unless the underlying assumptions are true. Always check the conditions before proceeding.

- A **confidence interval** uses a sample statistic (such as a proportion) to estimate a range of plausible values for the parameter of a population model.
  - All confidence intervals involve an estimate of the parameter, a margin of error, and a level of confidence.
  - For confidence intervals based on a given sample, the greater the margin of error, the higher the confidence.
  - At a given level of confidence, the larger the sample, the smaller the margin of error.

- A **hypothesis test** proposes a model for the population, then examines the observed statistics to see if that model is plausible.
  - A null hypothesis suggests a parameter value for the population model. Usually, we assume there is nothing interesting, unusual, or different about the sample results.
  - The alternative hypothesis states what we will believe if the sample results turn out to be inconsistent with our null model.
  - We compare the difference between the statistic and the hypothesized value with the standard deviation of the statistic. It’s the sampling distribution of this ratio that gives us a P-value.
  - The P-value of the test is the conditional probability that the null model could produce results at least as extreme as those observed in the sample or the experiment just as a result of sampling error.
  - A low P-value indicates evidence against the null model. If it is sufficiently low, we reject the null model.
  - A high P-value indicates that the sample results are not inconsistent with the null model, so we cannot reject it. However, this does not prove the null model is true.
  - Sometimes we will mistakenly reject the null hypothesis even though it’s actually true—that’s called a Type I error. If we fail to reject a false null hypothesis, we commit a Type II error.
  - The power of a test measures its ability to detect a false null hypothesis.
  - You can lower the risk of a Type I error by requiring a higher standard of proof (lower P-value) before rejecting the null hypothesis. But this will raise the risk of a Type II error and decrease the power of the test.
  - The only way to increase the power of a test while decreasing the chance of committing either error is to design a study based on a larger sample.

And now for some opportunities to review these concepts and skills . . .

1. **Herbal cancer.** A report in the *New England Journal of Medicine* (June 6, 2000) notes growing evidence that the herb *Aristolochia fangchi* can cause urinary tract cancer in those who take it. Suppose you are asked to design an experiment to study this claim. Imagine that you have data on urinary tract cancers in subjects who have used this herb and similar subjects who have not used it and that you can measure incidences of cancer and precancerous lesions in these subjects. State the null and alternative hypotheses you would use in your study.

2. **Colorblind.** Medical literature says that about 8% of males are colorblind. A university’s introductory psychology course is taught in a large lecture hall. Among the students, there are 325 males. Each semester when the
professor discusses visual perception, he shows the class a test for colorblindness. The percentage of males who are colorblind varies from semester to semester.

a) Is the sampling distribution model for the sample proportion likely to be Normal? Explain.

b) What are the mean and standard deviation of this sampling distribution model?

c) Sketch the sampling model, using the 68–95–99.7 Rule.

d) Write a few sentences explaining what the model says about this professor’s class.

3. Birthdays. During a 2-month period in 2002, 72 babies were born at the Tompkins Community Hospital in upstate New York. The table shows how many babies were born on each day of the week.

<table>
<thead>
<tr>
<th>Day</th>
<th>Births</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mon.</td>
<td>7</td>
</tr>
<tr>
<td>Tues.</td>
<td>17</td>
</tr>
<tr>
<td>Wed.</td>
<td>8</td>
</tr>
<tr>
<td>Thurs.</td>
<td>12</td>
</tr>
<tr>
<td>Fri.</td>
<td>9</td>
</tr>
<tr>
<td>Sat.</td>
<td>10</td>
</tr>
<tr>
<td>Sun.</td>
<td>9</td>
</tr>
</tbody>
</table>

a) If births are uniformly distributed across all days of the week, how many would you expect on each day?

b) Only 7 births occurred on a Monday. Does this indicate that women might be less likely to give birth on a Monday? Explain.

c) Are the 17 births on Tuesdays unusually high? Explain.

d) Can you think of any reasons why births may not occur completely at random?

4. Polling 2004. In the 2004 U.S. presidential election, the official results showed that George W. Bush received 50.7% of the vote and John Kerry received 48.3%. Ralph Nader, running as a third-party candidate, picked up only 0.4%. After the election, there was much discussion about exit polls, which had initially indicated a different result. Suppose you had taken a random sample of 1000 voters in an exit poll and asked them for whom they had voted.

a) Would you always get 507 votes for Bush and 483 for Kerry? Explain.

b) In 95% of such polls, your sample proportion of voters for Bush should be between what two values?

c) In 95% of such polls, your sample proportion of voters for Nader should be between what two numbers?

5. Leaky gas tanks. Nationwide, it is estimated that 40% of service stations have gas tanks that leak to some extent. A new program in California is designed to lessen the prevalence of these leaks. We want to assess the effectiveness of the program by seeing if the percentage of service stations whose tanks leak has decreased. To do this, we randomly sample 27 service stations in California and determine whether there is any evidence of leakage. In our sample, only 7 of the stations exhibit any leakage. Is there evidence that the new program is effective?

a) What are the null and alternative hypotheses?

b) Check the assumptions necessary for inference.

c) Test the null hypothesis.

d) What do you conclude (in plain English)?

e) If the program actually works, have you made an error? What kind?

f) What two things could you do to decrease the probability of making this kind of error?

g) What are the advantages and disadvantages of taking those two courses of action?

6. Surgery and germs. Joseph Lister (for whom Listerine is named!) was a British physician who was interested in the role of bacteria in human infections. He suspected that germs were involved in transmitting infection, so he tried using carbolic acid as an operating room disinfectant. In 75 amputations, he used carbolic acid 40 times. Of the 40 amputations using carbolic acid, 34 of the patients lived. Of the 35 amputations without carbolic acid, 19 patients lived. The question of interest is whether carbolic acid is effective in increasing the chances of surviving an amputation.

a) What kind of a study is this?

b) What do you conclude? Support your conclusion by testing an appropriate hypothesis.

c) What reservations do you have about the design of the study?

7. Scrabble. Using a computer to play many simulated games of Scrabble, researcher Charles Robinove found that the letter “A” occurred in 54% of the hands. This study had a margin of error of ±10%. (Chance, 15, no. 1 [2002])

a) Explain what the margin of error means in this context.

b) Why might the margin of error be so large?

c) Probability theory predicts that the letter “A” should appear in 63% of the hands. Does this make you concerned that the simulation might be faulty? Explain.

8. Dice. When one die is rolled, the number of spots showing has a mean of 3.5 and a standard deviation of 1.7. Suppose you roll 10 dice. What’s the approximate probability that your total is between 30 and 40 (that is, the average for the 10 dice is between 3 and 4)? Specify the model you use and the assumptions and conditions that justify your approach.

9. News sources. In May of 2000, the Pew Research Foundation sampled 1593 respondents and asked how they obtain news. In Pew’s report, 33% of respondents say that they now obtain news from the Internet at least once a week.

a) Pew reports a margin of error of ±3% for this result. Explain what the margin of error means.

b) Pew also asked about investment information, and 21% of respondents reported that the Internet is their main source of this information. When limited to the 780 respondents who identified themselves as investors, the percent who rely on the Internet rose to 28%. How would you expect the margin of error for this statistic to change in comparison with the margin of error for the percentage of all respondents?

c) When restricted to the 239 active traders in the sample, Pew reports that 45% rely on the Internet for investment information. Find a confidence interval for this statistic.

d) How does the margin of error for your confidence interval compare with the values in parts a and b? Explain why.
10. **Death penalty 2006.** In May of 2006, the Gallup Organization asked a random sample of 537 American adults this question:

   *If you could choose between the following two approaches, which do you think is the better penalty for murder, the death penalty or life imprisonment, with absolutely no possibility of parole?*

   Of those polled, 47% chose the death penalty, the lowest percentage in the 21 years that Gallup has asked this question.

   a) Create a 95% confidence interval for the percentage of all American adults who favor the death penalty.

   b) Based on your confidence interval, is it clear that the death penalty no longer has majority support? Explain.

   c) If pollsters wanted to follow up on this poll with another survey that could determine the level of support for the death penalty to within 2% with 98% confidence, how many people should they poll?

11. **Bimodal.** We are sampling randomly from a distribution known to be bimodal.

   a) As our sample size increases, what’s the expected shape of the sample’s distribution?

   b) What’s the expected value of our sample’s mean? Does the size of the sample matter?

   c) How is the variability of sample means related to the standard deviation of the population? Does the size of the sample matter?

   d) How is the shape of the sampling distribution model affected by the sample size?

12. **Vitamin D.** In July 2002 the *American Journal of Clinical Nutrition* reported that 42% of 1546 African-American women studied had vitamin D deficiency. The data came from a national nutrition study conducted by the Centers for Disease Control and Prevention in Atlanta.

   a) Do these data meet the assumptions necessary for inference? What would you like to know that you don’t?

   b) Create a 95% confidence interval.

   c) Interpret the interval in this context.

   d) Explain in this context what “95% confidence” means.

13. **Archery.** A champion archer can generally hit the bull’s-eye 80% of the time. Suppose she shoots 200 arrows during competition. Let \( \hat{p} \) represent the percentage of bull’s-eyes she gets (the sample proportion).

   a) What are the mean and standard deviation of the sampling distribution model for \( \hat{p} \)?

   b) Is a Normal model appropriate here? Explain.

   c) Sketch the sampling model, using the 68–95–99.7 Rule.

   d) What’s the probability that she gets at least 85% bull’s-eyes?

14. **Free throws 2007.** During the 2006–2007 NBA season, Kyle Korver led the league by making 191 of 209 free throws, for a success rate of 91.39%. But Matt Carroll was close behind, with 188 of 208 (90.39%).

   a) Find a 95% confidence interval for the difference in their free throw percentages.

   b) Based on your confidence interval, is it certain that Korver is better than Carroll at making free throws?

15. **Twins.** There is some indication in medical literature that doctors may have become more aggressive in inducing labor or doing preterm cesarean sections when a woman is carrying twins. Records at a large hospital show that, of the 43 sets of twins born in 1990, 20 were delivered before the 37th week of pregnancy. In 2000, 26 of 48 sets of twins were born preterm. Does this indicate an increase in the incidence of early births of twins? Test an appropriate hypothesis and state your conclusion.

16. **Eclampsia.** It’s estimated that 50,000 pregnant women worldwide die each year of eclampsia, a condition involving elevated blood pressure and seizures. A research team from 175 hospitals in 33 countries investigated the effectiveness of magnesium sulfate in preventing the occurrence of eclampsia in at-risk patients. Results are summarized below. (*Lancet*, June 1, 2002)

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Total Subjects</th>
<th>Reported side effects</th>
<th>Developed eclampsia</th>
<th>Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnesium sulfate</td>
<td>4999</td>
<td>1201</td>
<td>40</td>
<td>11</td>
</tr>
<tr>
<td>Placebo</td>
<td>4993</td>
<td>228</td>
<td>96</td>
<td>20</td>
</tr>
</tbody>
</table>

   a) Write a 95% confidence interval for the increase in the proportion of women who may develop side effects from this treatment. Interpret your interval.

   b) Is there evidence that the treatment may be effective in preventing the development of eclampsia? Test an appropriate hypothesis and state your conclusion.

17. **Eclampsia.** Refer again to the research summarized in Exercise 16. Is there any evidence that when eclampsia does occur, the magnesium sulfate treatment may help prevent the woman’s death?

   a) Write an appropriate hypothesis.

   b) Check the assumptions and conditions.

   c) Find the P-value of the test.

   d) What do you conclude about the magnesium sulfate treatment?

   e) If your conclusion is wrong, which type of error have you made?

   f) Name two things you could do to increase the power of this test.

   g) What are the advantages and disadvantages of those two options?

18. **Eggs.** The ISA Babcock Company supplies poultry farmers with hens, advertising that a mature B300 Layer produces eggs with a mean weight of 60.7 grams. Suppose that egg weights follow a Normal model with standard deviation 3.1 grams.

   a) What fraction of the eggs produced by these hens weigh more than 62 grams?

   b) What’s the probability that a dozen randomly selected eggs average more than 62 grams?

   c) Using the 68–95–99.7 Rule, sketch a model of the total weights of a dozen eggs.
19. **Polling disclaimer.** A newspaper article that reported the results of an election poll included the following explanation:

The Associated Press poll on the 2000 presidential campaign is based on telephone interviews with 798 randomly selected registered voters from all states except Alaska and Hawaii. The interviews were conducted June 21–25 by ICR of Media, Pa.

The results were weighted to represent the population by demographic factors such as age, sex, region, and education.

No more than 1 time in 20 should chance variations in the sample cause the results to vary by more than 4 percentage points from the answers that would be obtained if all Americans were polled.

The margin of sampling error is larger for responses of subgroups, such as income categories or those in political parties. There are other sources of potential error in polls, including the wording and order of questions.

a) Did they describe the 5 W's well?
b) What kind of sampling design could take into account the several demographic factors listed?
c) What was the margin of error of this poll?
d) What was the confidence level?
e) Why is the margin of error larger for subgroups?
f) Which kinds of potential bias did they caution readers about?

20. **Enough eggs?** One of the important issues for poultry farmers is the production rate—the percentage of days on which a given hen actually lays an egg. Ideally, that would be 100% (an egg every day), but realistically, hens tend to lay eggs on about 3 of every 4 days. ISA Babcock wants to advertise the production rate for the B300 Layer (see Exercise 18) as a 95% confidence interval with a margin of error of ±2%. How many hens must they collect data on?

21. **Teen deaths.** Traffic accidents are the leading cause of death among people aged 15 to 20. In May 2002, the National Highway Traffic Safety Administration reported that even though only 6.8% of licensed drivers are between 15 and 20 years old, they were involved in 14.3% of all fatal crashes. Insurance companies have long known that teenage boys were high risks, but what about teenage girls? One insurance company found that the driver was a teenage girl in 44 of the 388 fatal accidents they investigated. Is this strong evidence that the accident rate is lower for girls than for teens in general?

a) Test an appropriate hypothesis and state your conclusion.
b) Explain what your P-value means in this context.

c) Explain in this context what the P-value means.
d) The researchers claimed that the data prove that genetic differences between the two populations cause a difference in the frequency of occurrence of perfect pitch. Do you agree? Why or why not?

23. **Largemouth bass.** Organizers of a fishing tournament believe that the lake holds a sizable population of largemouth bass. They assume that the weights of these fish have a model that is skewed to the right with a mean of 3.5 pounds and a standard deviation of 2.2 pounds.

a) Explain why a skewed model makes sense here.
b) Explain why you cannot determine the probability that a largemouth bass randomly selected (“caught”) from the lake weighs over 3 pounds.
c) Each fisherman in the contest catches 5 fish each day. Can you determine the probability that someone’s catch averages over 3 pounds? Explain.
d) The 12 fishermen competing each caught the limit of 5 fish. What’s the probability that the total catch of 60 fish averaged more than 3 pounds?

24. **Cheating.** A Rutgers University study released in 2002 found that many high school students cheat on tests. The researchers surveyed a random sample of 4500 high school students nationwide; 74% of them said they had cheated at least once.

a) Create a 90% confidence interval for the level of cheating among high school students. Don’t forget to check the appropriate conditions.
b) Interpret your interval.
c) Explain what “90% confidence” means.
d) Would a 95% confidence interval be wider or narrower? Explain without actually calculating the interval.

25. **Language.** Neurological research has shown that in about 80% of people language abilities reside in the brain’s left side. Another 10% display right-brain language centers, and the remaining 10% have two-sided language control. (The latter two groups are mainly left-handers.) (Science News, 161, no. 24 [2002])

a) We select 60 people at random. Is it reasonable to use a Normal model to describe the possible distribution of the proportion of the group that has left-brain language control? Explain.
b) What’s the probability that our group has at least 75% left-brainers?
c) If the group had consisted of 100 people, would that probability be higher, lower, or about the same? Explain why, without actually calculating the probability.
d) How large a group would almost certainly guarantee at least 75% left-brainers? Explain.

26. **Cigarettes 2006.** In 1999 the Centers for Disease Control and Prevention estimated that about 34.8% of high school students smoked cigarettes. They established a national health goal of reducing that figure to 16% by the year 2010. To that end, they hoped to achieve a reduction to 20% by 2006. In 2006 they released a research study in which 23% of a random sample of 1815 high school students said they were current smokers. Is this evidence that progress toward the goal is off track?

a) Write appropriate hypotheses.
b) Verify that the appropriate assumptions are satisfied.
c) Find the P-value of this test.

d) Explain what the P-value means in this context.

e) State an appropriate conclusion.

f) Of course, your conclusion may be incorrect. If so, which kind of error did you commit?

27. Crohn’s disease. In 2002 the medical journal The Lancet reported that 335 of 573 patients suffering from Crohn’s disease responded positively to injections of the arthritis-fighting drug infliximab.

a) Create a 95% confidence interval for the effectiveness of this drug.

b) Interpret your interval in context.

c) Explain carefully what “95% confidence” means in this context.

28. Teen smoking 2006. The Centers for Disease Control and Prevention say that about 23% of teenagers smoke tobacco (down from a high of 38% in 1997). A college has 522 students in its freshman class. Is it likely that more than 30% of them are smokers? Explain.

29. Alcohol abuse. Growing concern about binge drinking among college students has prompted one large state university to conduct a survey to assess the size of the problem on its campus. The university plans to randomly select students and ask how many have been drunk during the past week. If the school hopes to estimate the true proportion among all its students with 90% confidence and a margin of error of 4%, how many students must be surveyed?

30. Errors. An auto parts company advertises that its special oil additive will make the engine “run smoother, cleaner, longer, with fewer repairs.” An independent laboratory decides to test part of this claim. It arranges to use a taxicab company’s fleet of cars. The cars are randomly divided into two groups. The company’s mechanics will use the additive in one group of cars but not in the other. At the end of a year the laboratory will compare the percentage of cars in each group that required engine repairs.

a) What kind of a study is this?

b) Will they do a one-tailed or a two-tailed test?

c) Explain in this context what a Type I error would be.

d) Explain in this context what a Type II error would be.

e) Which type of error would the additive manufacturer consider more serious?

f) If the cabs with the additive do indeed run significantly better, can the company conclude it is an effect of the additive? Can they generalize this result and recommend the additive for all cars? Explain.

31. Premies. Among 242 Cleveland-area children born prematurely at low birth weights between 1977 and 1979, only 74% graduated from high school. Among a comparison group of 233 children of normal birth weight, 83% were high school graduates. (“Outcomes in Young Adulthood for Very-Low-Birth-Weight Infants,” New England Journal of Medicine, 346, no. 3 [2002])

a) Create a 95% confidence interval for the difference in graduation rates between children of normal and children of very low birth weights. Be sure to check the appropriate assumptions and conditions.

b) Does this provide evidence that premature birth may be a risk factor for not finishing high school? Use your confidence interval to test an appropriate hypothesis.

c) Suppose your conclusion is incorrect. Which type of error did you make?

32. Safety. Observers in Texas watched children at play in eight communities. Of the 814 children seen biking, roller skating, or skateboarding, only 14% wore a helmet.

a) Create and interpret a 95% confidence interval.

b) What concerns do you have about this study that might make your confidence interval unreliable?

c) Suppose we want to do this study again, picking various communities and locations at random, and hope to end up with a 98% confidence interval having a margin of error of ±4%. How many children must we observe?

33. Fried PCs. A computer company recently experienced a disastrous fire that ruined some of its inventory. Unfortunately, during the panic of the fire, some of the damaged computers were sent to another warehouse, where they were mixed with undamaged computers. The engineer responsible for quality control would like to check out each computer in order to decide whether it’s undamaged or damaged. Each computer undergoes a series of 100 tests. The number of tests it fails will be used to make the decision. If it fails more than a certain number, it will be classified as damaged and then scrapped. From past history, the distribution of the number of tests failed is known for both undamaged and damaged computers.

<table>
<thead>
<tr>
<th>Number of tests failed</th>
<th>Undamaged (%)</th>
<th>Damaged (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>&gt;5</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

The table indicates, for example, that 80% of the undamaged computers have no failures, while 70% of the damaged computers have 2 failures.

a) To the engineers, this is a hypothesis-testing situation. State the null and alternative hypotheses.

b) Someone suggests classifying a computer as damaged if it fails any of the tests. Discuss the advantages and disadvantages of this test plan.

c) What number of tests would a computer have to fail in order to be classified as damaged if the engineers want to have the probability of a Type I error equal to 5%?

d) What’s the power of the test plan in part c?

e) A colleague points out that by increasing α just 2%, the power can be increased substantially. Explain.

34. Power. We are replicating an experiment. How will each of the following changes affect the power of our test? Indicate whether it will increase, decrease, or remain the same, assuming that all other aspects of the situation remain unchanged.

a) We increase the number of subjects from 40 to 100.

b) We require a higher standard of proof, changing from $\alpha = 0.05$ to $\alpha = 0.01$. 


35. **Approval 2007.** Of all the post–World War II presidents, Richard Nixon had the highest disapproval rating near the end of his presidency. His disapproval rating peaked at 66% in July 1974, just before he resigned. In May 2007, George W. Bush’s disapproval rating was 63%, according to a Gallup poll of 1000 voters. Pundits started discussing whether his rating was still discernibly better than Nixon’s. What do you think?

36. **Grade inflation.** In 1996, 20% of the students at a major university had an overall grade point average of 3.5 or higher (on a scale of 4.0). In 2000, a random sample of 1100 student records found that 25% had a GPA of 3.5 or higher. Is this evidence of grade inflation?

37. **Name recognition.** An advertising agency won’t sign an athlete to do product endorsements unless it is sure the person is known to more than 25% of its target audience. The agency always conducts a poll of 500 people to investigate the athlete’s name recognition before offering a contract. Then it tests \( H_0: p = 0.25 \) against \( H_A: p > 0.25 \) at a 5% level of significance.
   a) Why does the company use upper tail tests in this situation?
   b) Explain what Type I and Type II errors would represent in this context, and describe the risk that each error poses to the company.
   c) The company is thinking of changing its test to use a 10% level of significance. How would this change the company’s exposure to each type of risk?

38. **Name recognition, part II.** The advertising company described in Exercise 37 is thinking about signing a WNBA star to an endorsement deal. In its poll, 27% of the respondents could identify her.
   a) Fans who never took Statistics can’t understand why the company did not offer this WNBA player an endorsement contract even though the 27% recognition rate in the poll is above the 25% threshold. Explain it to them.
   b) Suppose that further polling reveals that this WNBA star really is known to about 30% of the target audience. Did the company initially commit a Type I or Type II error in not signing her?
   c) Would the power of the company’s test have been higher or lower if the player were more famous? Explain.

39. **NIMBY.** In March 2007, the Gallup Poll split a sample of 1003 randomly selected U.S. adults into two groups at random. Half \( (n = 502) \) of the respondents were asked,

   “Overall, do you strongly favor, somewhat favor, somewhat oppose, or strongly oppose the use of nuclear energy as one of the ways to provide electricity for the U.S.?”

   They found that 53% were either “somewhat” or “strongly” in favor. The other half \( (n = 501) \) were asked,

   “Overall, would you strongly favor, somewhat favor, somewhat oppose, or strongly oppose the construction of a nuclear energy plant in your area as one of the ways to provide electricity for the U.S.?”

   Only 40% were somewhat or strongly in favor. This difference is an example of the NIMBY (Not In My Back-Yard) phenomenon and is a serious concern to policy makers and planners. How large is the difference between the proportion of American adults who think nuclear energy is a good idea and the proportion who would be willing to have a nuclear plant in their area? Construct and interpret an appropriate confidence interval.

40. **Dropouts.** One study comparing various treatments for the eating disorder anorexia nervosa initially enlisted 198 subjects, but found overall that 105 failed to complete their assigned treatment programs. Construct and interpret an appropriate confidence interval. Discuss any reservations you have about this inference.