9. **Teen smoking, part I.** A Vermont study published in December 2001 by the American Academy of Pediatrics examined parental influence on teenagers' decisions to smoke. A group of students who had never smoked were questioned about their parents' attitudes toward smoking. These students were questioned again two years later to see if they had started smoking. The researchers found that among the 284 students who indicated that their parents disapproved of kids smoking, 54 had become established smokers. Among the 41 students who initially said their parents were lenient about smoking, 11 became smokers. Do these data provide strong evidence that parental attitude influences teenagers' decisions about smoking?

a) What kind of design did the researchers use?
b) Write appropriate hypotheses.
c) Are the assumptions and conditions necessary for inference satisfied?
d) Test the hypothesis and state your conclusion.
e) Explain in this context what your P-value means.
f) If that conclusion is actually wrong, which type of error did you commit?

10. **Depression.** A study published in the Archives of General Psychiatry in March 2001 examined the impact of depression on a patient's ability to survive cardiac disease. Researchers identified 450 people with cardiac disease, evaluated them for depression, and followed the group for 4 years. Of the 361 patients with no depression, 67 died. Of the 89 patients with minor or major depression, 26 died. Among people who suffer from cardiac disease, are depressed patients more likely to die than nondepressed ones?

a) What kind of design was used to collect these data?
b) Write appropriate hypotheses.
c) Are the assumptions and conditions necessary for inference satisfied?
d) Test the hypothesis and state your conclusion.
e) Explain in this context what your P-value means.
f) If your conclusion is actually incorrect, which type of error did you commit?

13. **Pregnancy.** In 1998, a San Diego reproductive clinic reported 42 live births to 157 women under the age of 38, but only 7 live births for 89 clients aged 38 and older. Is this strong evidence of a difference in the effectiveness of the clinic's methods for older women?

a) Test an appropriate hypothesis and state your conclusion.
b) If you concluded there was a difference, estimate that difference with a confidence interval and interpret your interval in context.

14. **Suicide.** The August 2001 issue of Pediatrics reported on a study of adolescent suicide attempts. Questionnaires were given to 6577 middle and high school students, 214 of who were adopted. Of the 6577,213 youngsters said they had attempted suicide within the last year—16 of those who were adopted and 197 of those who were not. Does this indicate a significantly different rate of suicide among adopted teens?

a) Test an appropriate hypothesis and state your conclusion.
b) If you concluded there was a difference, estimate that difference with a confidence interval and interpret your interval in context.

15. **Politics and sex.** One month before the election, a poll of 630 randomly selected voters showed 54% planning to vote for a certain candidate. A week later it became known that he had had an extramarital affair, and a new poll showed only 51% of 1010 voters supporting him. Do these results indicate a decrease in voter support for his candidacy?

a) Test an appropriate hypothesis and state your conclusion.
b) If you concluded there was a difference, estimate that difference with a confidence interval and interpret your interval in context.

16. **Retirement.** The Employee Benefit Research Institute reports that 27% of males anticipate having enough money to live comfortably in retirement, but only 18% of females express that confidence. If these results were based on samples of 250 people of each gender, would you consider this strong evidence that men and women have different outlooks?

a) Test an appropriate hypothesis and state your conclusion.
b) If you concluded there was a difference, estimate that difference with a confidence interval and interpret your interval in context.
17. **Twins.** In 2001, one county reported that among 3,132 white women who had babies, 94 were multiple births. There were also 20 multiple births to 606 black women. Does this indicate any racial difference in the likelihood of multiple births?
   a) Test an appropriate hypothesis and state your conclusion.
   b) If your conclusion is incorrect, which type of error did you commit?

18. **Shopping.** A survey of 430 randomly chosen adults found that 21% of the 222 men and 18% of the 208 women had purchased books online.
   a) Is there evidence that men are more likely than women to make online purchases of books? Test an appropriate hypothesis and state your conclusion in context.
   b) If your conclusion in fact proves to be wrong, did you make a Type I or Type II error?

19. **Mammograms.** It's widely believed that regular mammogram screening may detect breast cancer early, resulting in fewer deaths from the disease. One study that investigated this issue over a period of 18 years was published during the 1970s. Among 30,565 women who had never had mammograms, 196 died of breast cancer, while only 153 of 30,131 who had undergone screening died of breast cancer.
   a) Do these results suggest that mammograms may be an effective screening tool to reduce breast cancer deaths?
   b) If your conclusion is incorrect, what type of error have you committed?

20. **Mammograms redux.** In 2001, the conclusion of the study outlined in Exercise 19 was questioned. A new 9-year study was conducted in Sweden, comparing 21,088 women who had mammograms with 21,195 who did not. Of the women who underwent screening, 63 died of breast cancer, compared with 66 deaths among the control group. (New York Times, Dec 9, 2001)
   a) Do these results support the effectiveness of regular mammograms in preventing deaths from breast cancer?
   b) If your conclusion is incorrect, what kind of error have you committed?
Hypothesis Tests for 2 sample Proportions

Answers

9. a) Observational study
b) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 \neq 0 \) where \( p_1 \) is the proportion of students whose parents disapproved of smoking who became smokers and \( p_2 \) is the proportion of students whose parents are lenient about smoking who became smokers.
c) Yes. We assume the students were randomly selected; they are less than 10% of the population; samples are independent; more than 10 successes and failures in each sample.
d) \( z = -1.17 \), P-value = 0.2422. These samples do not show evidence that parental attitudes influence teens' decisions to smoke.
e) If there is no difference in the proportions, there is about a 24% chance of seeing the observed difference or larger by natural sampling variation.
f) Type II

10. a) Observational study
b) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 < 0 \) where \( p_1 \) is the proportion of people without depression who died within the 4 years and \( p_2 \) is the proportion of depressed people who died within the same time period.
c) Yes. We assume the patients were representative; they are less than 10% of the population; samples are independent; more than 10 successes and failures in each sample.
d) \( z = -2.22 \), P-value = 0.0131. With a P-value this low, we reject \( H_0 \). This study indicates that the death rate for patients with heart disease who die within 4 years is less in nondepressed patients than in depressed patients.
e) If there is no difference in the proportions, we will see an observed difference this large or larger only about 1.3% of the time by natural sampling variation.
f) Type I

12. a) (0.004, 0.209)
b) We are 95% confident, based on these data, that the proportion of heart disease patients who die within 4 years is between 0.4% and 20.9% higher for depressed patients than for nondepressed patients.
c) 95% of all random samples will produce intervals that contain the true value.

13. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 \neq 0 \). \( z = 3.56 \), P-value = 0.0004. With a P-value this low, we reject \( H_0 \). There is a significant difference in the clinic's effectiveness. Younger mothers have a higher birth rate than older mothers. Note that the success/failure condition is met based on the pooled estimate of \( p \).
b) We are 95% confident, based on these data, that the proportion of successful live births at the clinic is between 10.0% and 27.8% higher for mothers under 28 than those 38 and older. However, the Success/Failure Condition is not met for the older women, since # of successes < 10. We should be cautious in interpreting the level of the confidence interval.

14. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 \neq 0 \). \( z = 3.56 \), P-value = 0.0004. With a P-value this low, we reject \( H_0 \). This study shows a significant difference in the rate of teen suicides. Adopted teens are more likely to attempt suicide.
b) We are 95% confident, based on this study, that the proportion of teens who will attempt suicide is between 0.8% and 7.9% higher in adopted teens than in those who were not adopted.

15. \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 > 0 \). \( z = 1.18 \), P-value = 0.118. With a P-value this high, we fail to reject \( H_0 \). These data do not show evidence of a decrease in the voter support for the candidate.

16. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 \neq 0 \). \( z = 2.41 \), P-value = 0.0160. With a P-value this low, we reject \( H_0 \). This is moderately strong evidence that men and women have different outlooks.
b) We are 95% confident that between 12% and 24% more men than women expect to have enough money in retirement.

17. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 \neq 0 \). \( z = -0.39 \), P-value = 0.6951. With a P-value this high, we fail to reject \( H_0 \). There is no evidence of racial differences in the likelihood of multiple births, based on these data.
b) Type II

18. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 > 0 \). \( z = 0.78 \), P-value = 0.2166. With a P-value this high, we fail to reject \( H_0 \). There is no evidence, based on this information, that men are more likely than women to make online purchases of books.
b) Type II

19. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 > 0 \). \( z = 2.17 \), P-value = 0.0148. With a P-value this low, we reject \( H_0 \). These data do suggest that mammograms may reduce breast cancer deaths.
b) Type I

20. a) \( H_0: p_1 - p_2 = 0 \); \( H_A: p_1 - p_2 > 0 \). \( z = 0.24 \), P-value = 0.4068. With a P-value this high, we fail to reject \( H_0 \). These data do not suggest that mammograms are effective in reducing breast cancer deaths.
b) Type II