Questions on Topic Fourteen: Tests of Significance—Proportions and Means

Multiple-Choice Questions

Directions: The questions or incomplete statements that follow are each followed by five suggested answers or completions. Choose the response that best answers the question or completes the statement.

1. Which of the following is a true statement?
   (A) A well-planned hypothesis test should result in a statement either that the null hypothesis is true or that it is false.
   (B) The alternative hypothesis is stated in terms of a sample statistic.
   (C) If a sample is large enough, the necessity for it to be a simple random sample is diminished.
   (D) When the null hypothesis is rejected, it is because it is not true.
   (E) Hypothesis tests are designed to measure the strength of evidence against the null hypothesis.

2. Which of the following is a true statement?
   (A) The *P*-value of a test is the probability of obtaining a result as extreme (or more extreme) as the one obtained assuming the null hypothesis is false.
   (B) If the *P*-value for a test is .015, the probability that the null hypothesis is true is .015.
   (C) The larger the *P*-value, the more evidence there is against the null hypothesis.
   (D) If possible, always examine your data before deciding whether to use a one-sided or two-sided hypothesis test.
   (E) The alternative hypothesis is one-sided if there is interest in deviations from the null hypothesis in only one direction.

Questions 3–4 refer to the following:

Video arcades provide a vehicle for competition and status within teenage peer groups and are recognized as places where teens can hang out and meet their friends. The national PTA organization, concerned about the time and money being spent in video arcades by middle school students, commissions a statistical study to investigate whether or not middle school students are spending an average of over two hours per week in video arcades. Twenty communities are randomly chosen, five middle schools are randomly picked in each of the communities, and ten students are randomly interviewed at each school.

3. What is the parameter of interest?

(A) Video arcades
(B) A particular concern of the national PTA organization
(C) 1000 randomly chosen middle school students
(D) The mean amount of money spent in video arcades by middle school students
(E) The mean number of hours per week spent in video arcades by middle school students

4. What are the null and alternative hypotheses which the PTA is testing?

(A) H_0: $\mu = 2$ H_a: $\mu < 2$
(B) H_0: $\mu = 2$ H_a: $\mu \leq 2$
(C) H_0: $\mu = 2$ H_a: $\mu > 2$
(D) H_0: $\mu = 2$ H_a: $\mu \geq 2$
(E) H_0: $\mu = 2$ H_a: $\mu \neq 2$

5. A hypothesis test comparing two population proportions results in a P-value of .032. Which of the following is a proper conclusion?

(A) The probability that the null hypothesis is true is .032.
(B) The probability that the alternative hypothesis is true is .032.
(C) The difference in sample proportions is .032.
(D) The difference in population proportions is .032.
(E) None of the above are proper conclusions.

6. A company manufactures a synthetic rubber (jumping) bungee cord with a braided covering of natural rubber and a minimum breaking strength of 450 kg. If the mean breaking strength of a sample drops below a specified level, the production process is halted and the machinery inspected. Which of the following would result from a Type I error?

   (A) Halting the production process when too many cords break.
   (B) Halting the production process when the breaking strength is below the specified level.
   (C) Halting the production process when the breaking strength is within specifications.
   (D) Allowing the production process to continue when the breaking strength is below specifications.
   (E) Allowing the production process to continue when the breaking strength is within specifications.

7. One ESP test asks the subject to view the backs of cards and identify whether a circle, square, star, or cross is on the front of each card. If p is the proportion of correct answers, this may be viewed as a hypothesis test with H_0: $p = .25$ and H_a: $p > .25$. The subject is recognized to have ESP when the null hypothesis is rejected. What would a Type II error result in?

   (A) Correctly recognizing someone has ESP
   (B) Mistakenly thinking someone has ESP
   (C) Not recognizing that someone really has ESP
   (D) Correctly realizing that someone doesn't have ESP
   (E) Failing to understand the nature of ESP

8. A coffee-dispensing machine is supposed to deliver 12 ounces of liquid into a large paper cup, but a consumer believes that the actual amount is less. As a test he plans to obtain a sample of 5 cups of the dispensed liquid and if the mean content is less than 11.5 ounces, to reject the 12-ounce claim. If the machine operates with a known standard deviation of 0.9 ounces, what is the probability that the consumer will mistakenly reject the 12-ounce claim even though the claim is true?

   (A) .054
   (B) .107
   (C) .214
   (D) .289
   (E) .393

9. A pharmaceutical company claims that a medicine will produce a desired effect for a mean time of 58.4 minutes. A government researcher runs a hypothesis test of 40 patients and calculates a mean of $\overline{x} = 59.5$ with a standard deviation of $s = 8.3$. What is the P-value?

(A) $P\left(t > \dfrac{59.5 - 58.4}{8.3/\sqrt{40}}\right)$ with $df = 39$

(B) $P\left(t > \dfrac{59.5 - 58.4}{8.3/\sqrt{40}}\right)$ with $df = 40$

(C) $2P\left(t > \dfrac{59.5 - 58.4}{8.3/\sqrt{40}}\right)$ with $df = 39$

(D) $2P\left(t > \dfrac{59.5 - 58.4}{8.3/\sqrt{40}}\right)$ with $df = 40$

(E) $2P\left(z > \dfrac{59.5 - 58.4}{8.3/\sqrt{40}}\right)$

10. You plan to perform a hypothesis test with a level of significance of $\alpha = .05$. What is the effect on the probability of committing a Type I error if the sample size is increased?

(A) The probability of committing a Type I error decreases.
(B) The probability of committing a Type I error is unchanged.
(C) The probability of committing a Type I error increases.
(D) The effect cannot be determined without knowing the relevant standard deviation.
(E) The effect cannot be determined without knowing if a Type II error is committed.

11. A fast food chain advertises that their large bag of french fries has a weight of 150 grams. Some high school students, who enjoy french fries at every lunch, suspect that they are getting less than the advertised amount. With a scale borrowed from their physics teacher, they weigh a random sample of 15 bags. What is the conclusion if the sample mean is 145.8 g and standard deviation is 12.81 g?

(A) There is sufficient evidence to prove the fast food chain advertisement is true.
(B) There is sufficient evidence to prove the fast food chain advertisement is false.
(C) The students have sufficient evidence to reject the fast food chain's claim.
(D) The students do not have sufficient evidence to reject the fast food chain's claim.
(E) There is not sufficient data to reach any conclusion.

16. What is the probability of mistakenly failing to reject a false null hypothesis when a hypothesis test is being conducted at the 5% significance level ($\alpha = .05$)?

(A) .025
(B) .05
(C) .10
(D) .95
(E) There is insufficient information to answer this question.

17. A research dermatologist believes that cancers of the head and neck will occur most often of the left side, the side next to a window when a person is driving. In a review of 565 cases of head/neck cancers, 305 occurred on the left side. What is the resulting *P*-value?

(A) $P\left(z > \dfrac{.54 - .50}{\sqrt{(.5)(.5)/565}}\right)$

(B) $2P\left(z > \dfrac{.54 - .50}{\sqrt{(.5)(.5)/565}}\right)$

(C) $P\left(z > \dfrac{.54 - .50}{\sqrt{(.54)(.46)/565}}\right)$

(D) $P\left(z \geq \dfrac{.54 - .50}{\sqrt{(.54)(.46)/565}}\right)$

(E) $2P\left(z > \dfrac{.54 - .50}{\sqrt{(.54)(.46)/565}}\right)$

18. Suppose you do five independent tests of the form H_0: $\mu = 38$ versus H_a: $\mu > 38$, each at the $\alpha = .01$ significance level. What is the probability of committing a Type I error and incorrectly rejecting a true null hypothesis with at least one of the five tests?

(A) .01
(B) .049
(C) .05
(D) .226
(E) .951

19. Given an experiment with H_0: $\mu = 35$, H_a: $\mu < 35$, and a possible correct value of 32, you obtain a sample statistic of $\bar{x} = 33$. After doing analysis, you realize that the sample size n is actually larger than you first thought. Which of the following results from reworking with the increase in sample size?

    (A) Decrease in probability of a Type I error; decrease in probability of a Type II error; decrease in power.
    (B) Increase in probability of a Type I error; increase in probability of a Type II error; decrease in power.
    (C) Decrease in probability of a Type I error; decrease in probability of a Type II error; increase in power.
    (D) Increase in probability of a Type I error; decrease in probability of a Type II error; decrease in power.
    (E) Decrease in probability of a Type I error; increase in probability of a Type II error; increase in power.

20. Thirty students volunteer to test which of two strategies for taking multiple-choice exams leads to higher average results. Each student flips a coin, and if heads, uses Strategy A on the first exam and then Strategy B on the second, while if tails, uses Strategy B first and then Strategy A. The average of all 30 Strategy A results is then compared to the average of all 30 Strategy B results. What is the conclusion at the 5% significance level if a two-sample hypothesis test, H_0: $\mu_1 = \mu_2$, H_a: $\mu_1 \neq \mu_2$, results in a P-value of .18?

    (A) The observed difference in average scores is significant.
    (B) The observed difference in average scores is not significant.
    (C) A conclusion is not possible without knowing the average scores resulting from using each strategy.
    (D) A conclusion is not possible without knowing the average scores and the standard deviations resulting from using each strategy.
    (E) A two-sample hypothesis test should not be used here.

21. Choosing a smaller level of significance, that is, a smaller α-risk, results in

    (A) a lower risk of Type II error and lower power.
    (B) a lower risk of Type II error and higher power.
    (C) a higher risk of Type II error and lower power.
    (D) a higher risk of Type II error and higher power.
    (E) no change in risk of Type II error or in power.

22. The greater the difference between the null hypothesis claim and the true value of the population parameter,

    (A) the smaller the risk of a Type II error and the smaller the power.
    (B) the smaller the risk of a Type II error and the greater the power.
    (C) the greater the risk of a Type II error and the smaller the power.
    (D) the greater the risk of a Type II error and the greater the power.
    (E) the greater the probability of no change in Type II error or in power.

23. A company selling home appliances claims that the accompanying instruction guides are written at a 6th grade reading level. An English teacher believes that the true figure is higher and with the help of an AP Statistics student runs a hypothesis test. The student randomly picks one page from each of 25 of the company's instruction guides, and the teacher subjects the pages to a standard readability test. The reading levels of the 25 pages are given in the following table:

Reading grade level	5	6	7	8	9	10
Number of pages	6	10	4	2	2	1

Is there statistical evidence to support the English teacher's belief?

(A) No, because the *P*-value is greater than .10.
(B) Yes, the *P*-value is between .05 and .10 indicating some evidence for the teacher's belief.
(C) Yes, the *P*-value is between .01 and .05 indicating evidence for the teacher's belief.
(D) Yes, the *P*-value is between .001 and .01 indicating strong evidence for the teacher's belief.
(E) Yes, the *P*-value is less than .001 indicating very strong evidence for the teacher's belief.

24. Suppose H_0: $p = .4$, and the power of the test for the alternative hypothesis $p = .35$ is .75. Which of the following is a valid conclusion?

(A) The probability of committing a Type I error is .05.
(B) The probability of committing a Type II error is .65.
(C) If the alternative $p = .35$ is true, the probability of failing to reject H_0 is .25.
(D) If the null hypothesis is true, the probability of rejecting it is .25.
(E) If the null hypothesis is false, the probability of failing to reject it is .65.

25. A factory is located close to a city high school. The manager claims that the plant's smokestacks spew forth an average of no more than 350 pounds of pollution per day. As an AP Statistics project, the class plans a one-sided hypothesis test with a critical value of 375 pounds. Suppose the standard deviation in daily pollution poundage is known to be 150 pounds and the true mean is 385 pounds. If the sample size is 100 days, what is the probability that the class will mistakenly fail to reject the factory manager's false claim?

(A) .0475
(B) .2525
(C) .7475
(D) .7514
(E) .9525

Free-Reponse Questions

Directions: You must show all work and indicate the methods you use. You will be graded on the correctness of your methods and on the accuracy of your final answers.

Ten Open-Ended Questions

1. Next to good brakes, proper tire pressure is the most crucial safety issue on your car. Incorrect tire pressure compromises cornering, braking, and stability. Both underinflation and overinflation can lead to problems. The number on the tire is the maximum allowable air pressure—not the recommended pressure for that tire. At a roadside vehicle safety checkpoint, officials randomly select 30 cars for which 35 psi is the recommended tire pressure and calculate the average of the actual tire pressure in the front right tires. What is the parameter of interest, and what are the null and alternative hypotheses that the officials are testing?

2. No vaccinations are 100% risk free, and the theoretical risk of rare complications always have to be balanced against the severity of the disease. Suppose the CDC decides that a risk of one in a million is the maximum acceptable risk of GBS (Guillain-Barre syndrome) complications for a new vaccine for a particularly serious strain of influenza. A large sample study of the new vaccine is conducted with the following hypotheses:

   H_0: The proportion of GBS complications is .000001 (one in a million)
   H_a: The proportion of GBS complications is greater than .000001 (one in a million)

   The *P*-value of the test is .138.

   (a) Interpret the *P*-value in context of this study.

   (b) What conclusion should be drawn at the $\alpha = .10$ significance level?

   (c) Given this conclusion, what possible error, Type I or Type II, might be committed, and give a possible consequence of committing this error.

3. A particular wastewater treatment system aims at reducing the most probable number per ml (MPN/ml) of *E. coli* to 1000 MPN/100 ml. A random study of 40 of these systems in current use is conducted with the data showing a mean of 1002.4 MPN/100 ml and a standard deviation of 7.12 MPN/100 ml. A test of significance is conducted with:

   H_0: The mean concentration of *E. coli* after treatment under this system is 1000 MPN/100 ml.

   H_a: The mean concentration of *E. coli* after treatment under this system is greater than 1000 MPN/100 ml.

   The resulting *P*-value is .0197 with $df = 39$ and $t = 2.132$.

   (a) Interpret the *P*-value in context of this study.

   (b) What conclusion should be drawn at a 5% significance level?

   (c) Given this conclusion, what possible error, Type I or Type II, might be committed, and give a possible consequence of committing this error.

4. A 20-year study of 5000 British adults noted four bad habits: smoking, drinking, inactivity, and poor diet. The study looked to show that there is a higher death rate (proportion who die in a 20-year period) among people with all four bad habits than among people with none of the four bad habits.

   (a) Was this an experiment or observational study? Explain.

   (b) What are the null and alternative hypotheses?

   (c) What would be the result of a Type I error?

   (d) What would be the result of a Type II error?

   Of the 314 people who had all four bad habits, 91 died during the study, while of the 387 people with none of the four bad habits, 32 died during the study.

   (e) Calculate and interpret the *P*-value in context of this study.

5. It is estimated that 17.4% of all U.S. households own a Roth IRA. The American Association of University Professors (AAUP) believes this figure is higher among their members and commissions a study. If 150 out of a random sample of 750 AAUP members own Roth IRAs, is this sufficient evidence to support the AAUP belief?

6. A long accepted measure of the discharge rate (in 1000 ft^3/sec) at the mouth of the Mississippi River is 593. To test if this has changed, ten measurements at random times are taken: 590, 596, 592, 588, 589, 594, 590, 586, 591, 589. Is there statistical evidence of a change?

7. A behavior study of high school students looked at whether a higher proportion of boys than girls met a recommended level of physical activity (increased heart rate for 60 minutes/day for at least 5 days during the 7 days before the survey). What is the proper conclusion if 370 out of a random sample of 850 boys and 218 out of an independent random sample of 580 girls met the recommended level of activity?