

1. Suppose that the population of the scores of all high school seniors who took the SAT-V (SAT verbal) test this year follows a normal distribution with mean $\mu = 480$ and standard deviation $\sigma = 90$. You read a report that claims that 10,000 students who took part in a national program for improving one's SAT-V score had significantly better scores (at the 0.05 level of significance) than the population as a whole. In order to determine if the improvement is of practical significance, one should

- A. find out the actual mean score of the 10,000 students.
 - B. find out the actual P-value.
 - C. use a two-sided test rather than the one-sided test implied by the report.
-

2. Does taking garlic tablets twice a day provide significant health benefits? To investigate this issue, a researcher conducted a study with fifty adult subjects who took garlic tablets twice a day for a period of six months. At the end of the study, 100 variables related to the health of the subjects were measured on each subject and the means compared to known means for these variables in the population of all adults. Four of these variables were significantly better (in the sense of statistical significance) at the 5% level for the group taking the garlic tablets as compared to the population as a whole, and one variable was significantly better at the 1% level for the group taking the garlic tablets as compared to the population as a whole. It would be correct to conclude

- A. that there is good statistical evidence that taking garlic tablets twice a day provides some health benefits.
 - B. that there is good statistical evidence that taking garlic tablets twice a day provides benefits for the variable that was significant at the 1% level. We should be somewhat cautious about making claims for the variables that were significant at the 5% level.
 - C. None of the above.
-

3. A radio station wants to know if residents in their area are in favor of a proposed tax increase. They invite listeners to call in to respond to the poll. 645 of the 800 who responded were against the tax. The station calculated a 95% confidence interval and declared "Between 77.9% and 83.4% of residents oppose the new tax."

- A. This is a valid interval, so the station must be right.
 - B. We can't say for certain that between 77.9% and 83.4% of residents oppose the new tax because it is a 95% confidence interval.
 - C. Because of the way the poll was conducted, the results are invalid.
-

4. An instructor wanted to construct a confidence interval for the mean GPA of the students in his class. He used the campus records system to obtain all their GPA's and computed the 95% interval as (2.32, 2.87). If he wants to use this interval to describe the students in his class

- A. he's 95% confident the interval contains the real average GPA.
- B. there's a 5% chance the interval is wrong.
- C. he didn't need an interval after all.

5. A statistics instructor wants to know if an expensive computer program will help his students learn. He tries the program on a section of 35 students one semester and finds a P-value of 0.078 when comparing the class' average grade to that of the semester before. The vendor, still confident about the program's potential success, lets him try it again. That time, the P-value for the test was 0.006 with a class of 150 students, indicating the students did better on average, with the program. What's the difference?

- A. The second class was smarter.
 - B. The second class had more students — a larger sample size.
 - C. There were probably some outliers that affected the analysis.
-

6. National opinion polls are normally conducted by telephone over a short time span, usually 2 or 3 days. A "Presidential Approval Poll" may state that the president's approval rating is $x\%$, "with a margin of error $\pm 3\%$." What they usually don't say is that this is really a 95% confidence interval estimate. Can we truly be 95% confident that the true percentage for all adult Americans is within 3% of the estimate?

- A. Yes
 - B. No
 - C. There is not enough information to be able to tell.
-

7. Several years ago, some researchers at my university asked me to recheck the statistics on a paper they were writing about the impacts of a group of new science courses. They had performed 63 statistical tests on their data at $\alpha = 5\%$, and had found 4 "significant" results. They were excited. Should they have been?

- A. Yes — this could be important.
 - B. Maybe — depends on the actual results.
 - C. No — too many tests were done.
-

8. Students at my university wanted to move spring break later in the semester, so that weather at the beach would be warmer. They stood outside a campus eatery and tried to flag down passersby to sign their petition. The student newspaper later announced that "an overwhelming proportion of those contacted were in favor of moving spring break." It wasn't moved. Why did the administration reject the plan?

- A. The sample wasn't random.
 - B. The administration just didn't care.
 - C. The sample wasn't large enough.
-

9. There is an old saying "Statistics means never having to say you're certain." Why might this be true?

- A. It isn't — if I get a P-value of 0.0000001 (or less) I'm sure the null hypothesis is wrong.
- B. It's just a joke on statisticians.
- C. To be totally confident, I would need a z^* of ∞ .

10. Criminal jury trials are analogous to statistical tests of hypotheses. If we consider the decision as a test of H_0 : not guilty versus H_a : guilty, and the jury must find "beyond a reasonable doubt" that the defendant is guilty, what might be considered an appropriate α level for the test?

- A. 10%
- B. 5%
- C. 1%