# AP Statistics - SIG: Logic of Significance Testing

AP Statistics Logic of Significance Testing (SIG) - 16 questions

16 Questions | 32 min

---

**1.** Which of the following is a true statement?

(A) A well-planned test of significance should result in a statement that either the null hypothesis is true or that it is false.

(B) The null hypothesis is one-sided and expressed using either < or > if there is interest in deviations in only one direction.

(C) When a true parameter value is further from the hypothesized value, it becomes easier to reject the alternative hypothesis.

(D) Tests of significance (hypothesis tests) are designed to measure the strength of evidence against the null hypothesis.

(E) Increasing the sample size makes it more difficult to conclude that an observed distance between observed and hypothesized values is significant.

**2.** Which of the following is a true statement?

(A) If the P-value is 0.16, the probability that the null hypothesis is correct is 0.16.

(B) The larger the P-value is, the more evidence there is against the null hypothesis.

(C) If the P-value is small enough, we can conclude that the alternative hypothesis is true.

(D) It is always wrong to use sample statistics in stating hypotheses.

(E) It is helpful to examine your data before deciding whether to use a one-sided or a two-sided hypothesis test.

**3.** Which of the following is a true statement?

(A) The P-value of a test is the probability of obtaining a result as extreme as or more extreme than the one obtained when assuming the null hypothesis is true.

(B) If the P-value for a test is 0.014, the probability that the null hypothesis is true is 0.014.

(C) When the null hypothesis is rejected, it is because the null hypothesis is not true.

(D) A very large P-value provides convincing evidence that the null hypothesis is true.

(E) The larger the P-value is, the greater the evidence of something statistically significant.

**4.** Which of the following is a true statement?

(A) If a population parameter is known, there is no reason to run a hypothesis test on that population parameter.

(B) The P-value can be negative or positive depending on whether the sample statistic is less than or greater than the claimed value of the population parameter in the null hypothesis.

(C) The P-value is based on a specific test statistic, so the P-value must be chosen before an experiment is conducted.

(D) If a P-value is larger than a specified value $\alpha$, the data are statistically significant at that level.

(E) The P-value is a probability calculated when assuming that the alternative hypothesis is true.

**5.** The mean miles per gallon (mpg) of a certain model car is 32.8. Concerned that a production process change might have lowered that efficiency, an inspector tests a random sample of six cars, calculating a mean of 31.9 mpg with a t-score of −2.16 and a P-value of 0.0416. Which of the following is the most reasonable conclusion?

(A) 95.84% of the cars produced under the new process will have an mpg under 31.9.

(B) 95.84% of the cars produced under the new process will have an mpg under 32.8.

(C) 4.16% of the cars produced under the new process will have an mpg over 32.8.

(D) There is evidence at the 5% significance level to conclude that the new process is producing cars with a mean mpg under 31.9.

(E) There is evidence at the 5% significance level to conclude that the new process is producing cars with a mean mpg under 32.8.

**6.** Which of the following is a true statement?

(A) A P-value is a conditional probability.

(B) The P-value is the probability that the null hypothesis is true.

(C) A P-value is the probability the null hypothesis is true given a particular observed statistic.

(D) The P-value is the same as the power of a hypothesis test.

(E) A large P-value is evidence against the null hypothesis because it says that observed results are unlikely to occur when the null hypothesis is true.

**7.** A small retail store employs five men and five women. When comparing the mean years of service of these men and women, which of the following is most appropriate?

(A) A two-sample z-test of population means

(B) A two-sample t-test of population means

(C) A one-sample z-test on a set of differences

(D) A one-sample t-test on a set of differences

(E) None of the above are appropriate.

**8.** In a simple random sample (SRS) of 625 families who do not live near any chemical plant, 10 had children with leukemia. In an SRS of 412 families living near chemical plants, 15 had children with leukemia. A 90% confidence interval of the difference is reported to be −0.020 ± 0.017. Which of the following is a proper conclusion?

(A) The interval is invalid because it does not contain 0.

(B) The interval is invalid because probabilities cannot be negative.

(C) Families who live near chemical plants are approximately 2.0 percent more likely to have children with leukemia than are families who do not live near chemical plants.

(D) 90 percent of families living near chemical plants are approximately 2.0 percent more likely to have children with leukemia than are families who do not live near chemical plants.

(E) None of the above are proper conclusions.

**9.** For which of the following is it appropriate to use a t-distribution with 12 degrees of freedom? I. Constructing a confidence interval from an SRS with $n = 11$. II. Doing a hypothesis test $H_0 : \mu_1 - \mu_2 = 0$ and $H_a : \mu_1 - \mu_2 > 0$, where $n_1 = 12$ and $n_2 = 12$. III. Doing a $\chi^2$-test of independence where the contingency table has 4 rows and 3 columns.

(A) I only (B) II only (C) III only (D) I, II, and III (E) None are appropriate.

**10.** Are generic batteries just as good as brand-name batteries? A consumer group runs a test on random samples of 20 brand-name alkaline AA batteries and 20 generic alkaline AA batteries. The group lets all the batteries run continuously and notes the number of hours when they die. Which of the following shows the proper set of hypotheses?

(A) $H_0 : \overline{x}_{\text{name}} = \overline{x}_{\text{generic}}$ and $H_a : \overline{x}_{\text{name}} \neq \overline{x}_{\text{generic}}$

(B) $H_0 : \overline{x}_{\text{name}} = \overline{x}_{\text{generic}}$ and $H_a : \overline{x}_{\text{name}} > \overline{x}_{\text{generic}}$

(C) $H_0 : \overline{x}_{\text{name}} = \overline{x}_{\text{generic}}$ and $H_a : \overline{x}_{\text{name}} < \overline{x}_{\text{generic}}$

(D) $H_0 : \overline{x}_{\text{difference}} = 0$ and $H_a : \overline{x}_{\text{difference}} \neq 0$

(E) None of the above show the proper set of hypotheses.

**11.** A driving instructor claims that the probability of passing the driving test on the first try after finishing his course is 0.85. A reporter believes this claim is high and plans to survey a random sample of people who took the instructor's class. Assuming the sample size will be large, what statistical test will be most appropriate?

(A) z-test of a proportion

(B) z-test for difference of two proportions

(C) chi-square test of homogeneity of proportions

(D) t-test of mean of a proportion

(E) t-test for a slope where a proportion is interpreted as a slope

**12.** The P-value for a two-sided t-test is 0.10. If the test had been one-sided, what would the P-value have been?

(A) 0.05

(B) 0.10

(C) 0.20

(D) 0.95

(E) It depends on the direction of the alternative hypothesis.

**13.** Medical researchers keep trying different molecular variations of a drug. They test each of these variations for effectiveness in relieving pain after minor surgeries. At the 5% significance level, 2 out of 75 of these variations show some statistical significance. Which of the following should be concluded?

(A) A sample size of 75 is not large enough for any serious medical conclusion.

(B) $\frac{2}{75} = 0.0267 < 0.05$, so there is sufficient evidence to conclude that the drug is effective in relieving pain after minor surgeries.

(C) Although no conclusion can be made about the drug in general, positive conclusions can be made about 2 of the 75 molecular variations in their effectiveness in relieving pain after minor surgeries.

(D) Pain relief is very susceptible to the placebo effect, so no conclusion is reasonable from this study without knowing whether or not random assignment and blinding with one group receiving a placebo was part of the procedure.

(E) There is insufficient evidence that any of the molecular variations are effective in relieving pain after minor surgeries because at the 5% significance level, statistically significant results in 2 out of 75 tests could simply be due to chance variability.

**14.** One of the conditions in performing a one-sample t-test is that the sample size be less than 10 percent of the population size. What is the reason for this condition?

(A) This allows for the application of the central limit theorem.

(B) This allows for using the sample standard deviation as an approximation to the population standard deviation.

(C) This guarantees an unbiased sample.

(D) If the sample is not too large when compared to the population, the dependence among observations is negligible even though we sample without replacement.

(E) This guarantees a large enough sample for the test procedure.

**15.** Suppose a two-sided hypothesis test results in a P-value of 0.32. If a one-sided test is conducted on the same data, which of the following would be true about the possible values of the resulting P-value?

(A) The only possible value is 0.16.

(B) The only possible value is 0.32.

(C) The only possible value is 0.64.

(D) The only possible values are 0.16 and 0.84.

(E) The only possible values are 0.32 and 0.68.

**16.** A politician is concerned about whether he has a problem with support from women voters. A large number of polls are taken by various polling organizations, each surveying independent random samples of 500 male voters and 500 female voters. The difference in the proportion of male support and proportion of female support is calculated for each poll. If there is actually no difference in support for the politician between male and female voters, which of the following results would be anticipated? I. If for each poll tests are run at the 5% significance level with the hypotheses $H_0 : p_M = p_W$ versus $H_a : p_M > p_W$, the P-value will be less than 0.05 for approximately 5 percent of the tests. II. There will be approximately equal numbers of polls showing higher support levels from males and higher support levels from females. III. If for each poll 95% confidence intervals for the difference in proportions, $p_M - p_W$, are calculated, approximately 95 percent of the confidence intervals will contain 0.

(A) I only (B) II only (C) III only (D) I and II only (E) I, II, and III