AP Statistics - HTM: Hypothesis Tests for Means

AP Statistics Hypothesis Tests for Means (HTM) - 33 questions

33 Questions | 66 min

---

1. An efficiency expert is interested in comparing the mean time taken for breaks by employees in areas with access to the Internet and those in areas that do not have this access. She interviews a simple random sample (SRS) of 10 employees with access to the Internet and an SRS of 10 without access. The efficiency expert then proceeds to run a t-test to compare the mean time taken for breaks in each group. Which of the following is a necessary assumption?

(A) The population standard deviations from each group are known.

(B) The population standard deviations from each group are unknown.

(C) The population standard deviations from each group are equal.

(D) The population of break times from each group is normally distributed.

(E) The samples must be independent samples and for each sample np and $n(1-p)$ must both be at least 10.

2. A high school coach claims that the average pulse rate of those trying out for sports is 62.4 beats per minute (bpm). The AP Statistics instructor suspects this is a made-up number and runs a hypothesis test on a simple random sample (SRS) of 32 students trying out for sports, calculating a mean of 65.0 bpm with a standard deviation of 10.3 bpm. What is the P-value?

(A) $P\left(z > \frac{65.0-62.4}{\frac{10.3}{\sqrt{32}}}\right)$

(B) $2P\left(z > \frac{65.0-62.4}{\frac{10.3}{\sqrt{32}}}\right)$

(C) $P\left(t > \frac{65.0-62.4}{\frac{10.3}{\sqrt{32}}}\right)$ with df = 31

(D) $2P\left(t > \frac{65.0-62.4}{\frac{10.3}{\sqrt{32}}}\right)$ with df = 31

(E) $P\left(t > \frac{65.0-62.4}{10.3}\right)$ with df = 32

3. In a one-sided hypothesis test for the mean for a random sample of size 20, the t-score of the sample mean is 2.615 in the direction of the alternative. Is this significant at the 5% level? At the 1% level?

(A) Significant at the 1% level but not at the 5% level

(B) Significant at the 5% level but not at the 1% level

(C) Significant at both the 1% and 5% levels

(D) Significant at neither the 1% nor 5% levels

(E) Cannot be determined from the given information

4. A confidence interval estimate is determined from the summer earnings of a simple random sample (SRS) of n students. All other things being equal, which of the following results in a larger margin of error?

(A) A lesser confidence level

(B) A smaller sample standard deviation

(C) A smaller sample size

(D) Introducing bias into the sampling

(E) Introducing blinding into an experiment

5. A spokesperson for the National Council of Teachers of Mathematics (NCTM) states that middle school math teachers spend an average of $1,250 per year of their own funds on classroom materials. A member of a Board of Education believes that the real figure is lower. So he interviews 12 randomly chosen middle school math teachers and comes up with a mean of $1,092 and a standard deviation of $308 spent out of pocket by middle school math teachers during the past academic year for classroom materials. Where is the P-value?

(A) Below 0.01

(B) Between 0.01 and 0.025

(C) Between 0.025 and 0.05

(D) Between 0.05 and 0.10

(E) Over 0.10

6. A critic of public education claims that during their four years of high school, students read an average of only 38 books. An English teacher plans to sample 100 high school graduates randomly as to the number of books read during their high school years and will reject the critic's claim if the sample mean exceeds 40 books. If the critic's claim is wrong and the true mean is 43 books, what is the probability that the random sample will lead to a mistaken failure to reject the critic's claim? Assume that the standard deviation in number of books read from the sample is 12 books.

(A) $P\left(t < \dfrac{43-38}{\frac{12}{\sqrt{100}}}\right)$

(B) $P\left(t > \dfrac{40-38}{\frac{12}{\sqrt{100}}}\right)$

(C) $P\left(t < \dfrac{40-43}{\frac{12}{\sqrt{100}}}\right)$

(D) $P\left(t > \dfrac{40-43}{\frac{12}{\sqrt{100}}}\right)$

(E) $P\left(t > \dfrac{43-38}{\frac{12}{\sqrt{100}}}\right)$

7. A dietician claims that a new weight loss program will result in an average loss of 9 pounds in the first month. The program developer believes that the average weight loss in the first month will be greater than this. The program developer runs a test on a random sample of 64 overweight volunteers. What conclusion is reached if the sample mean loss is 9.55 pounds with a standard deviation of 3.00 pounds?

(A) The P-value is less than 0.001, indicating very strong evidence against the 9-pound claim.

(B) The P-value is 0.01, indicating strong evidence against the 9-pound claim.

(C) The P-value is 0.07, indicating some evidence against the 9-pound claim.

(D) The P-value is 0.18, indicating very little evidence against the 9-pound claim.

(E) The P-value is 0.43, indicating no evidence against the 9-pound claim.

8. According to Major League Baseball (MLB) official rules, the average weight of manufactured baseballs should be 5.25 ounces. An MLB inspector weighs a simple random sample (SRS) of five baseballs, obtaining weights of 5.3, 5.15, 5.15, 5.2, and 5.25 ounces. The inspector then runs an appropriate hypothesis test. Which of the following gives the P-value of this test?

(A) $P\left(t < \frac{5.21 - 5.25}{\frac{0.0652}{\sqrt{5}}}\right)$ with df = 4

(B) $2P\left(t < \frac{5.21 - 5.25}{\frac{0.0652}{\sqrt{5}}}\right)$ with df = 4

(C) $P\left(t < \frac{5.21 - 5.25}{\frac{0.0652}{\sqrt{5}}}\right)$ with df = 5

(D) $2P\left(t < \frac{5.21 - 5.25}{\frac{0.0652}{\sqrt{5}}}\right)$ with df = 5

(E) $0.5P\left(t < \frac{5.21 - 5.25}{\frac{0.0652}{\sqrt{5}}}\right)$ with df = 5

9. An editor for a college magazine claims that the mean cost for 4-year tuition at private colleges is $258,000. A high school guidance counselor checks a random sample of 10 students who attended private colleges for four years and found the mean cost for tuition alone was $245,000 with a standard deviation of $38,000. What is the test statistic in performing a hypothesis test of $H_0 : \mu = 258000$ and $H_a : \mu \neq 258000$?

(A) $t = \dfrac{245000 - 258000}{\frac{38000}{\sqrt{10-1}}}$

(B) $t = \dfrac{245000 - 258000}{\frac{38000^2}{\sqrt{10}}}$

(C) $t = \dfrac{245000 - 258000}{\frac{(38000)^2}{\sqrt{10-1}}}$

(D) $t = \dfrac{245000 - 258000}{\frac{(38000)}{\sqrt{10}}}$

(E) $t = \dfrac{245000 - 258000}{\frac{(38000)}{\sqrt{10}}}$

10. In testing the null hypothesis $H_0 : \mu = 51$ against the alternative hypothesis $H_a : \mu > 51$, a sample from a normal population has a mean of 52.1 with a corresponding t-score of 2.2 and a P-value of 0.0377. Which of the following is not a reasonable conclusion?

(A) If the null hypothesis is assumed to be true, the probability of obtaining a sample mean as extreme as or more extreme than 52.1 is only 0.0377.

(B) In all samples using the same sample size and the same sampling technique, the null hypothesis will be wrong 3.77 percent of the time.

(C) There is sufficient evidence to reject the null hypothesis at the 5% significance level.

(D) If the population standard deviation σ had been known, the z-score of 2.2 would have resulted in a smaller P-value.

(E) If the alternative hypothesis had been two-sided, the same t-score would have resulted in a P-value of 0.0754.

11. A cell phone salesperson claims that the average teen processes 3,700 texts per month. A teacher believes this claim is high and randomly samples 30 teens. What conclusion is reached if the sample mean is 3,490 with a standard deviation of 1,120?

(A) There is sufficient evidence to prove the salesperson's claim is true.

(B) There is sufficient evidence to prove the salesperson's claim is false.

(C) The teacher has sufficient evidence to reject the salesperson's claim.

(D) The teacher does not have sufficient evidence to reject the salesperson's claim.

(E) There are not sufficient data to reach any conclusion.

12. The claim is made that the students at Lake Wobegon High School are smarter than other high school students. To test this, the principal gathers a simple random sample (SRS) of 30 students and finds their mean IQ is 104 with a standard deviation of 7.5. What is the test statistic in performing a hypothesis test of $H_0 : \mu = 100$ and $H_a : \mu > 100$ (average IQ is 100)?

(A) $t = \dfrac{104 - 100}{7.5}$

(B) $t = \dfrac{104 - 100}{\frac{7.5}{\sqrt{30}}}$

(C) $t = \dfrac{104 - 100}{\frac{7.5}{\sqrt{29}}}$

(D) $t = \dfrac{104 - 100}{\frac{7.5}{\sqrt{30}}}$

(E) $t = \dfrac{104 - 100}{\frac{7.5}{\sqrt{29}}}$

13. A t-test is being conducted at the 5% significance level with $H_0 : \mu = 20$ and $H_a : \mu \neq 20$. If a 95 percent t-interval constructed from the same data set contains the value 20, which of the following can be concluded about the t-test?

(A) The P-value is less than 0.05, and there is sufficient evidence to reject H_0.

(B) The P-value is greater than 0.05, and there is sufficient evidence to reject H_0.

(C) The P-value is less than 0.05, and there is not sufficient evidence to reject H_0.

(D) The P-value is greater than 0.05, and there is not sufficient evidence to reject H_0.

(E) Without knowing the sample standard deviation, none of the above can be concluded.

14. A test of the hypotheses $H_0 : \mu = 10$ versus $H_a : \mu \neq 10$ was conducted using a sample of size $n = 6$. The test statistic was $t = 1.641$. What was the P-value of the test?

(A) 0.0760 (B) 0.0809 (C) 0.1008 (D) 0.1519 (E) 0.1617

15. Suppose in a small town with a single Walmart, it is known that customers spend an average of $75 per visit. A new Target is opening, and its sales manager believes that the new store is taking in a higher mean amount per customer visit than Walmart. To test the Target manager's belief, which hypotheses should be used?

(A) H_0: the mean amount spent at Target is equal to $75 H_a: the mean amount spent at Target is less than $75

(B) H_0: the mean amount spent at Target is equal to $75 H_a: the mean amount spent at Target is greater than $75

(C) H_0: the mean amount spent at Target is less than $75 H_a: the mean amount spent at Target is greater than $75

(D) H_0: the mean amount spent at Walmart is equal to $75 H_a: the mean amount spent at Walmart is less than $75

(E) H_0: the mean amount spent at Walmart is equal to $75 H_a: the mean amount spent at Walmart is greater than $75

16. A spokesperson for the airline industry states that the mean number of pieces of lost luggage per 1,000 passengers is 3.09. A consumer agency believes the true figure is higher and runs an appropriate hypothesis test that results in a P-value of 0.075. What is an appropriate conclusion?

(A) Because 0.075 > 0.05, there is sufficient evidence at the 5% significance level to conclude that the mean number of pieces of lost luggage per 1,000 passengers is greater than 3.09.

(B) Because 0.075 > 0.05, there is sufficient evidence at the 5% significance level to conclude that the mean number of pieces of lost luggage per 1,000 passengers is less than 3.09.

(C) Because 0.075 < 0.10, there is sufficient evidence at the 10% significance level to conclude that the mean number of pieces of lost luggage per 1,000 passengers is less than 3.09.

(D) Because 0.075 < 0.10, there is not sufficient evidence at the 10% significance level to conclude that the mean number of pieces of lost luggage per 1,000 passengers is greater than 3.09.

(E) Because 0.05 < 0.075 < 0.10, there is sufficient evidence at the 10% significance level but not at the 5% significance level to conclude that the mean number of pieces of lost luggage per 1,000 passengers is greater than 3.09.

17. A researcher believes that a new diet should improve weight gain in laboratory mice. The average weight gain for 18 mice on the new diet is 4.2 ounces with a standard deviation of 0.4 ounces. The average weight gain for 16 control mice on the old diet is 3.8 ounces with a standard deviation of 0.3 ounces. Where is the P-value?

(A) Below 0.01

(B) Between 0.01 and 0.025

(C) Between 0.025 and 0.05

(D) Between 0.05 and 0.10

(E) Over 0.10

18. A teacher is interested in comparing the mean writing speeds of teenagers using cursive versus those printing. A simple random sample (SRS) of 50 students is chosen. The students are timed copying the same paragraph, once in cursive and once in printing. For each student, a coin flip decides which method he/she must use first. Which of the following is a proper test?

(A) Test of difference in two population means

(B) Test of difference in two population proportions

(C) One-sample test on differences of paired data

(D) Chi-square goodness-of-fit test

(E) Chi-square test for homogeneity

19. An entomologist believes that a certain species of large beetles decreased in size between two time periods. He plans to sample the remains of 45 fossilized beetles randomly from each period and reject any equality claim if the mean size in the second time period sample is at least 1 centimeter less than the mean size from the first period sample. If the standard deviation in size in all time periods for this species of beetle is known to be 1.8 centimeters, what is the probability the entomologist will commit a Type I error and mistakenly reject a correct null hypothesis of equality?

(A) $P\left(z > \dfrac{1-0}{\sqrt{\frac{1.8^2}{45}+\frac{1.8^2}{45}}}\right)$

(B) $P\left(t > \dfrac{1-0}{\sqrt{\frac{1.8^2}{45}+\frac{1.8^2}{45}}}\right)$ with df = 44

(C) $P\left(t > \dfrac{1-0}{\sqrt{\frac{1.8^2}{45}+\frac{1.8^2}{45}}}\right)$ with df = 88

(D) $P\left(z > \dfrac{1-0}{\frac{1.8}{\sqrt{45}}}\right)$

(E) $P\left(t > \dfrac{1-0}{\frac{1.8}{\sqrt{45}}}\right)$ with df = 44

20. College students majoring in computer science and interested in pursuing careers in computer security were randomly put into either Python or PowerShell classes. In a standard security certification exam, 230 students who took Python scored an average of 76.5 with a standard deviation of 7.4, while 185 students who took PowerShell averaged 78.8 with a standard deviation of 5.6. Is there sufficient evidence of a difference in the test results? What is the test statistic for $H_0 : \mu_1 - \mu_2 = 0$ and $H_a : \mu_1 - \mu_2 \neq 0$?

(A) $t = \dfrac{76.5 - 78.8}{\sqrt{\frac{(7.4)^2}{230} + \frac{(5.6)^2}{185}}}$

(B) $t = \dfrac{76.5 - 78.8}{2\sqrt{\frac{(7.4)^2}{230} + \frac{(5.6)^2}{185}}}$

(C) $t = \dfrac{76.5 - 78.8}{\frac{7.4}{\sqrt{230}} + \frac{5.6}{\sqrt{185}}}$

(D) $t = \dfrac{76.5 - 78.8}{2\left(\frac{7.4}{\sqrt{230}} + \frac{5.6}{\sqrt{185}}\right)}$

(E) $t = \dfrac{76.5 - 78.8}{\frac{6.5}{\sqrt{207.5}}}$

21. To test whether husbands or wives have greater manual agility, a simple random sample (SRS) of 75 married couples is chosen. All 150 people are given the Illinois Agility Run Test, a commonly used fitness test of agility, and their scores are recorded. What is the proper conclusion at a 5% significance level if a two-sample hypothesis test, $H_0 : \mu_1 - \mu_2 = 0$, $H_a : \mu_1 - \mu_2 \neq 0$, results in a P-value of 0.038?

(A) The observed difference between husbands and wives is significant.

(B) The observed difference is not significant.

(C) A conclusion is not possible without knowing the mean scores of the husbands and of the wives.

(D) A conclusion is not possible without knowing both the mean and standard deviation of the scores of the husbands and of the wives.

(E) A two-sample hypothesis test should not be used in this example.

22. A random sample of college students tries out two different strategies of card counting at blackjack. The data give the following table.

Strategy	Sample Size	Mean Winnings	Standard Deviation
A	20	131	12
B	15	140	15

Assuming all conditions of inference are met, and performing a two-sample t-test with $H_0 : \mu_1 - \mu_2 = 0$, is there evidence of a significant difference in outcomes at the 5% significance level?

(A) P < 0.05, so reject H_0

(B) P < 0.05, so fail to reject H_0

(C) P > 0.05, so reject H_o

(D) P > 0.05, so fail to reject H_o

(E) The 5% significance level is inappropriate for gambling experiments.

23. To compare online prices between Amazon and Walmart, a consumer-oriented research organization picks 50 basic items and checks the prices of these items online at Amazon and Walmart. Which test should be used to determine if the prices are different at the two sites?

(A) χ^2-test for goodness-of-fit

(B) χ^2-test for independence

(C) Two-sample z-test

(D) Two-sample t-test

(E) Matched pairs t-test

24. The prescription drugs Coumadin and Plavix are both blood thinners known to increase clotting time. In one double-blind study, Coumadin outperformed Plavix. The 95% confidence interval estimate of the difference in mean clotting time (prothrombin time) increase was (2.1, 3.7) seconds. Which of the following is a reasonable conclusion?

(A) Plavix raises clotting time an average of 2.1 seconds, while Coumadin raises clotting time an average of 3.7 seconds.

(B) There is a 0.95 probability that Coumadin will outperform Plavix in raising clotting time for any given individual.

(C) There is a 0.95 probability that Coumadin will outperform Plavix by at least 2.1 seconds in raising the clotting time for any given individual.

(D) We should be 95% confident that Coumadin will outperform Plavix as the drug that raises a clotting time.

(E) None of the above.

25. Material thickness plays a vital part in the quality of soccer balls. Researchers plan to test two layers versus five layers of lining between the cover and the bladder of soccer balls. Two identical-looking soccer balls, one with two internal layers and the other with five, are each kicked and given scores ranging from 1–10 by 12 professional soccer players. A score of 10 is given for a top-quality feel to the ball. A score of 1 is given for a poor-quality feel to the ball. For each player, a coin flip decides which ball is to be kicked first. What type of study and what type of inference test are appropriate here?

(A) Observational study and t-test of difference of means

(B) Matched pairs design and one-sample t-test for a mean difference

(C) Matched pairs design and two-sample t-test for difference of population means

(D) Completely randomized design and one-sample t-test for a mean difference

(E) Completely randomized design and two-sample t-test for difference of population means

26. An experiment was performed to determine whether a waitress would receive larger tips if she gave her name when greeting customers. A sample of waitresses was randomly assigned to two groups. One group was instructed to give their names when greeting customers, while the other group was instructed not to give their names. After one week, the mean tip received for each group was calculated. The P-value of this one-sided test was 0.065. Consider the three factors: original sample size, two sample standard deviations, and magnitude of the difference of the two sample means. Which of the following would have resulted in a smaller P-value?

(A) The magnitude of the difference of the two sample means and the two sample standard deviations remains the same, but the original sample size is smaller.

(B) The original sample size remains the same, but the magnitude of the difference of the two sample means is larger and the two sample standard deviations are smaller.

(C) The original sample size and the magnitude of the difference of the two sample means remain the same, but the two sample standard deviations are larger.

(D) The original sample size and the two sample standard deviations remain the same, but the magnitude of the difference of the two sample means is smaller.

(E) The original sample size remains the same, but the magnitude of the difference of the two sample means is smaller and the two sample standard deviations are larger.

27. In which of the following is a matched pairs t-test not appropriate?

(A) Heights of twins for 100 randomly selected pairs of twins

(B) Heights of 100 randomly selected children when they are age 10 and again when they are age 11

(C) Heights of both spouses of 100 randomly selected married couples

(D) Heights of both people in pairs formed by 100 randomly selected people

(E) Heights of 100 randomly selected children from England and from France in pairs by matching family income.

28. A doctors' association gathers data on weights from a random sample of elementary school children who are allowed to watch TV all week and from an independent random sample of elementary school children who are allowed to watch TV only on weekends. The data are summarized in the table below.

	TV All Week	Limited TV
Number in Sample	175	160
Mean Weight	75.4	71.3
Standard Deviation	9.8	8.2

What is the t-statistic for an appropriate test of whether elementary school children who watch TV all week have a higher mean weight than elementary school children who watch TV only on weekends: $H_0 : \mu_1 - \mu_2 = 0$ and $H_a : \mu_1 - \mu_2 > 0$?

(A) $t = \dfrac{75.4 - 71.3}{\sqrt{\frac{9.8}{175} + \frac{8.2}{160}}}$

(B) $t = \dfrac{9.8 - 8.2}{\sqrt{\frac{75.4^2}{175} + \frac{71.3^2}{160}}}$

(C) $t = \dfrac{9.8 - 8.2}{\sqrt{\frac{9.8^2}{175} + \frac{8.2^2}{160}}}$

(D) $t = \dfrac{75.4 - 71.3}{\sqrt{\frac{9.8^2}{175} + \frac{8.2^2}{160}}}$

(E) $t = \dfrac{75.4 - 71.3}{\sqrt{\frac{9.8^2 + 8.2^2}{175 + 160}}}$

29. A sports equipment manufacturer is considering producing one of two new golf tees. The manufacturer chooses 60 sports stores. The manufacturer randomly assigns 30 stores to sell a prototype of one of the tees and the other 30 stores to sell a prototype of the other tee. After a period of time, the mean number of sales in each group of 30 stores is calculated. A statistician concludes that the difference in means is statistically significant. Which of the following is the appropriate conclusion?

(A) It is reasonable to conclude that the difference in sales is caused by the difference in tees because the tees were randomly assigned to stores.

(B) It is reasonable to conclude that the difference in sales is caused by the difference in tees because the sample size was ≥30.

(C) It is not reasonable to conclude that the difference in sales is caused by the difference in tees because the 60 stores were not randomly selected.

(D) It is not reasonable to conclude that the difference in sales is caused by the difference in tees because there was no control group for comparison.

(E) It is not reasonable to conclude that the difference in sales is caused by the difference in tees because this was an observational study, not an experiment.

30. A car simulator was used to compare the effect on reaction time between DWI (driving while intoxicated) and DWT (driving while texting). A total of 20 volunteers were instructed to drive at 60 mph and then hit the brakes in response to the sudden image of a child darting into the road. One day, each driver was tested for stopping distance while driving while texting. Another day, each driver was tested after consuming a quantity of alcohol. For each driver, which test was done on the first day was decided by a coin toss. Which of the following is an appropriate test and hypotheses to determine if there is a difference in stopping distances between DWI and DWT?

31. A sports statistician is interested in whether the mean number of home runs hit in American League stadiums (with designated hitters batting instead of pitchers) is greater than the mean number hit in National League stadiums (where pitchers do bat). A random sample of games is analyzed as to the number of home runs. Computer output for a random sample of American and National League games is as follows. Do the confidence interval and t-test lead to the same conclusion, and was this expected?

(A) With 0.0386 < 0.05 and the entire confidence interval > 0, both lead to the conclusion that there is sufficient evidence that the mean number of home runs hit in American League games is greater than the mean number hit in National League games. It should be expected that both lead to the same conclusion.

(B) With 0.0386 < 0.05 and the entire confidence interval > 0, both lead to the conclusion that there is sufficient evidence that the mean number of home runs hit in American League games is greater than the mean number hit in National League games. It is possible that this confidence interval and t-test contradict each other.

(C) With 0.0086 < 0.0386 < 0.2374, both lead to the conclusion that there is sufficient evidence that the mean number of home runs hit in American League games is greater than the mean number hit in National League games. It should be expected that both lead to the same conclusion.

(D) With 0.0386 < 0.05 and 0.05 < 0.2374, they lead to opposite conclusions, with the t-test showing a significant difference and the confidence interval not. This is not unexpected because these are two different approaches.

(E) With 0.0386 < 0.05 and 0.05 < 0.2374, they lead to opposite conclusions with the t-test showing a significant difference and the confidence interval not. This is unexpected as only one conclusion at the given significance level is possible.

32. In an often-quoted study reported in the British Medical Journal, researchers examined the number of people admitted to the ER for auto accidents on Friday the 6th of several months versus on Friday the 13th of several months. Is there evidence that more people are admitted, on average, on Friday the 13th? Some data are summarized below. Number of ER Admissions from Auto Accidents

	Oct 1989	Jul 1990	Sep 1991	Dec 1991	Mar 1992	Nov 1992
Friday 6th	9	6	11	11	3	5
Friday 13th	13	12	14	10	4	12

Let μ_1 = the mean number of ER admissions from auto accidents on all Fridays that fall on the 6th Let μ_2 = the mean number of ER admissions from auto accidents on all Fridays that fall on the 13th Let μ_D = the mean of the differences (the 6th minus the 13th) in the number of ER admissions from auto accidents on all Fridays that fall on the 6th and 13th. What is the appropriate test and alternative hypothesis?

(A) Two-sample t-test with $H_a : \mu_1 \neq \mu_2$

(B) Two-sample t-test with $H_a : \mu_1 < \mu_2$

(C) Two-sample t-test with $H_a : \mu_1 > \mu_2$

(D) Match paired t-test with $H_a : \mu_D < 0$

(E) Match paired t-test with $H_a : \mu_D > 0$

33. A study compared the math scores on a standardized test for homeschooled and public-schooled high school students. The scores of the eight students in each group are shown in the following table.

Home-Schooled	78	65	70	83	96	66	70	81
Public-Schooled	75	74	78	89	91	70	75	89

Assume the groups were simple random samples and all conditions for inference were met. What is the appropriate test to determine if there is a significant difference in the average test scores of the two groups?

(A) Matched pair t-test for means

(B) Two-sample t-test for means

(C) Chi-square test of independence

(D) Chi-square goodness-of-fit test

(E) Linear regression t-test