4. High cholesterol levels in people can be reduced by exercise, diet, and medication. Twenty middle-aged males with cholesterol readings between 220 and 240 milligrams per deciliter (mg/dL) of blood were randomly selected from the population of such male patients at a large local hospital. Ten of the 20 males were randomly assigned to group A, advised on appropriate exercise and diet, and also received a placebo. The other 10 males were assigned to group B, received the same advice on appropriate exercise and diet, but received a drug intended to reduce cholesterol instead of a placebo. After three months, posttreatment cholesterol readings were taken for all 20 males and compared to pretreatment cholesterol readings. The tables below give the reduction in cholesterol level (pretreatment reading minus posttreatment reading) for each male in the study.

Group A (placebo)

Reduction (in mg/dL)	2	19	8	4	12	8	17	7	24	1

Mean Reduction: 10.20 Standard Deviation of Reductions: 7.66

Group B (cholesterol drug)

Reduction (in mg/dL)	30	19	18	17	20	−4	23	10	9	22

Mean Reduction: 16.40 Standard Deviation of Reductions: 9.40

Do the data provide convincing evidence, at the $\alpha = 0.01$ level, that the cholesterol drug is effective in producing a reduction in mean cholesterol level beyond that produced by exercise and diet?

5. Windmills generate electricity by transferring energy from wind to a turbine. A study was conducted to examine the relationship between wind velocity in miles per hour (mph) and electricity production in amperes for one particular windmill. For the windmill, measurements were taken on twenty-five randomly selected days, and the computer output for the regression analysis for predicting electricity production based on wind velocity is given below. The regression model assumptions were checked and determined to be reasonable over the interval of wind speeds represented in the data, which were from 10 miles per hour to 40 miles per hour.

Predictor	Coef	SE Coef	T	P
Constant	0.137	0.126	1.09	0.289
Wind velocity	0.240	0.019	12.63	0.000

S = 0.237 R-Sq = 0.873 R-Sq (adj) = 0.868

(a) Use the computer output above to determine the equation of the least squares regression line. Identify all variables used in the equation.

(b) How much more electricity would the windmill be expected to produce on a day when the wind velocity is 25 mph than on a day when the wind velocity is 15 mph? Show how you arrived at your answer.

(c) What proportion of the variation in electricity production is explained by its linear relationship with wind velocity?

(d) Is there statistically convincing evidence that electricity production by the windmill is related to wind velocity? Explain.

AP® STATISTICS
2011 SCORING GUIDELINES

Question 4

Intent of Question

The primary goal of this question was to assess students' ability to set up, perform and interpret the results of a hypothesis test. More specific goals were to assess students' ability to: (1) state hypotheses; (2) identify the name of an appropriate statistical test and check appropriate assumptions/conditions; (3) compute the test statistic and *p*-value; (4) draw a conclusion, with justification, in the context of the problem.

Solution

Step 1: States a correct pair of hypotheses.

Let μ_A represent the mean cholesterol reduction if all such male patients at this hospital are advised on appropriate exercise and diet and also receive a placebo.

Let μ_B represent the mean cholesterol reduction if all such male patients at this hospital are advised on appropriate exercise and diet but receive the drug instead of a placebo.

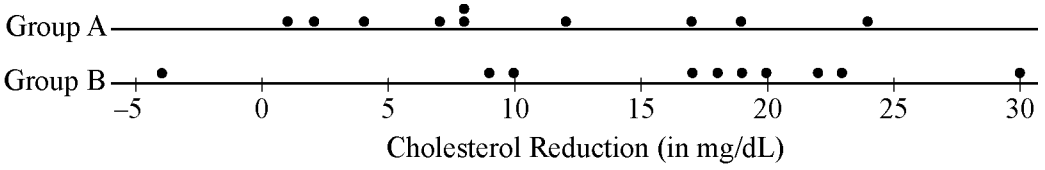
The hypotheses to be tested are $H_0: \mu_A = \mu_B$ versus $H_a: \mu_A < \mu_B$.

Step 2: Identifies a correct test procedure (by name or by formula) and checks appropriate conditions.

The appropriate procedure is a two-sample *t*-test.

When comparing two experimental treatments using a two-sample *t*-test, the subjects must be randomly assigned to the treatments. This condition is stated in the question (10 men were randomly assigned to group A and the remaining 10 men to group B).

The second condition is that the two populations are approximately normally distributed or the sample sizes are sufficiently large. Because of the small sample sizes (10 in each treatment group), we need to check whether it is reasonable to assume that the samples came from populations that are normally distributed. The following dotplots reveal slight skewness and a possible outlier for group B, but it appears reasonable to proceed with the two-sample *t*-test.



Step 3: Demonstrates correct mechanics, including the value of the test statistic and *p*-value (or the rejection region).

The test statistic is: $t = \dfrac{\bar{x}_A - \bar{x}_B}{\sqrt{\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}}} = \dfrac{10.20 - 16.40}{\sqrt{\dfrac{7.66^2}{10} + \dfrac{9.40^2}{10}}} \approx -1.62$

With df = 17.3, *p*-value ≈ 0.062.

AP® STATISTICS
2011 SCORING GUIDELINES

Question 4 (continued)

Step 4: States a correct conclusion in the context of the problem, using the result of the statistical test.

Because the *p*-value is greater than the significance level of $\alpha = 0.01$, we fail to reject H_0. The data do not provide enough evidence at the 0.01 level of significance to conclude that the drug is effective in producing a mean cholesterol reduction beyond that provided by exercise and dietary advice.

Scoring

Steps 1, 2, 3 and 4 are each scored as essentially correct (E), partially correct (P) or incorrect (I).

Step 1 is scored as follows:

Essentially correct (E) if the response states hypotheses with correct comparisons between the means and defines the population means as the parameters.

Partially correct (P) if the response states hypotheses with correct comparisons between the means *OR* correctly defines the population means as the parameters, but not both.

Incorrect (I) if the response does not meet the criteria for E or P.

Note: Defining the parameter symbols in context or simply using μ_A and μ_B, with subscripts clearly relevant to the context is sufficient for defining parameters.

Step 2 is scored as follows:

Essentially correct (E) if the response correctly includes the following three components:
1. Identifies the correct test procedure (by name or by formula).
2. Checks for random assignment to treatments.
3. Checks for normality.

Partially correct (P) if the response correctly includes exactly two of the three components listed above.

Incorrect (I) if the response fails to meet the criteria for E or P.

Notes
- Graphs of both distributions must be produced and described to check the normality condition.
- If the response calls for a *pooled* two-sample *t*-test, step 2 can be scored as E as long as the condition of equal variances is mentioned and checked by comparing the variability in the graphs or the sample standard deviations.
- If the response calls for applying a paired *t*-test, then step 2 is scored as I, but steps 3 and 4 can be scored as E if the test mechanics are correct in step 3 and the conclusion is correct in step 4.

AP® STATISTICS
2011 SCORING GUIDELINES

Question 4 (continued)

Step 3 is scored as follows:

Essentially correct (E) if both the test statistic and *p*-value are correctly calculated.

Partially correct (P) if the test statistic is correctly calculated but not the *p*-value
OR
if the test statistic is calculated incorrectly, but the correct *p*-value for the computed test statistic is given.

Incorrect (I) if the response fails to meet the criteria for E or P.

Step 4 is scored as follows:

Essentially correct (E) if the response provides a correct conclusion in context, also providing justification based on the linkage between the size of the *p*-value and the conclusion.

Partially correct (P) if the response provides a correct conclusion, including justification based on the size of the *p*-value, but not in context
OR
if the response provides a correct conclusion, written in context, but without justification based on linkage to the *p*-value.

Incorrect (I) if the response does not meet the criteria for E or P.

Notes

- If the conclusion is consistent with the p-value from step 3, and also in context with justification based the size of the p-value, then step 4 is scored as E (even if the p-value in step 3 is incorrect).
- A conclusion in step 4 that is equivalent to "accept H_0" (such as "we conclude that the drug is not effective") is not acceptable for an E. Such a response should be scored as P, provided that the conclusion is in context with justification based on the size of the p-value. Such a response should be scored as I if it lacks either context or linkage to the p-value.

Each essentially correct (E) step counts as 1 point. Each partially correct (P) step counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down, depending on the overall strength of the response and communication.