AP Statistics
2001 Solutions and Scoring Guidelines

These materials were produced by Educational Testing Service (ETS), which develops and administers the examinations of the Advanced Placement Program for the College Board. The College Board and Educational Testing Service (ETS) are dedicated to the principle of equal opportunity, and their programs, services, and employment policies are guided by that principle.

Question 1 - Solution

Part (a):

An outlier is any value that is more than 1.5 *IQR below the lower quartile or 1.5 *IQR above the upper quartile. If MIN < Q_1 – 1.5* IQR, there is at least one outlier on the low side and if MAX > Q_3 + 1.5*IQR, there is at least one outlier on the high side.

OR

An outlier is any observation that is more than 2 (or 3) standard deviations above or below the mean.

Part (b):

IQR = 19.250 – 9.680 = 9.57
1.5*IQR = 14.355

Q_1 – 1.5*IQR = 9.680 – 14.355 = –4.675
Q_3 + 1.5*IQR = 19.250 + 14.355 = 33.605

There is at least one outlier on the high side because the maximum value is greater than 1.5*IQR above the upper quartile, Q_3.

OR

mean – 2*std.dev = 1.447 OR mean – 3*std.dev = –5.300
mean + 2*std.dev = 28.435 mean + 3*std.dev = 35.182
Since 38.180 > 28.435 (or 38.180 > 35.182), there is at least one outlier on the high side.

Part (c):

Since Q_1 = 9.68 inches, more than 25% of the years had less than 10 inches of rain. Hence, 10 inches of rain is not an unusual value.

OR

Since 10 inches is within one standard deviation of the mean (Z = –0.732), 10 inches of rain is not an unusual value.

Scoring

The solution of this problem has four components:

1. **Outlier decision rule:**
   The student must state how to make a decision, using correct boundary values (one of the two standard procedures), that identifies outliers on both sides.

Question 1 (cont'd.)

2. **Two boundary values:**
   The student must describe how to get both upper and lower plausible boundary values using his or her decision rule - symbolically, numerically, or graphically.

Note: Incorrect but plausible boundaries:

- Give boundary values where the lower value is less than Q_1 and the upper value is greater than Q_3.

AND

- Must be of the form:
  Outlier < (location measure) – (multiplier) * (spread measure)
  Outlier > (location measure) + (multiplier) * (spread measure)

3. **Execution of outlier decision rule:**
   The student must give a correct conclusion using his or her outlier rule. Only the upper side must be checked. Boundary values and outlier decision rule must be plausible to receive credit for this part. (Remember, the execution must be consistent with the rule given in part (a).)

4. **Comment on "only 10 inches of rainfall":**
   The student must state that 10 inches of rain is not (or is) an unusual value and must explain why.
   Examples of incorrect comments include stating that 10 inches of rain is not an outlier or basing his or her conclusion on an assumption of normality.

Parts (a) and (b) need to be read together. Credit for components 1 and 2 may be given if found in part (b) instead of part (a). If components 1 or 2 are incorrect in part (a), credit for components 1 and 2 cannot be given based on work in part (b).

4 **Complete Response**
All four components are essentially correct.

3 **Substantial Response**
Three components are essentially correct.

2 **Developing Response**
Two components are essentially correct.

1 **Minimal Response**
Only one component is essentially correct.

Remember: Before assigning a score, assess the entire paper holistically.

Question 2 - Solution

For A:

Total 3-year cost: $10,000 + 36($50) = $10,000 + $1,800 = $11,800
This cost is fixed.

For B:

Expected number of repairs in 3 years = 3[0(.5) + 1(.25) + 2(.15) + 3(.1)] = 3(.85) = 2.55
Expected cost of repairs in 3 years = 3($200)(0.85) = $510
Expected 3-year cost = $10,500 + $510 = $11,010

Choice:

Choose B because it has a lower expected (or average) cost. (A has a fixed cost that is $790 ($11,800 - $11,010) higher than the expected cost of B.)

Scoring

The solution should include the following four elements:

1. Correct calculation of 3-year cost for A.
2. Correct calculation of a relevant expected **value** for B (expected number of repairs per year or per 3 years or expected cost of repairs per year or per 3 years). Calculation of expected value must be shown.
3. Correct calculation of expected total **cost** for B.
4. Choice of B with a *complete and coherent* explanation that is based on student's prior calculations for A & B.
   "Complete and coherent " means that:
   - costs for A & B are compared

   AND
   - B's cost has been indicated as "expected" or "average" or "mean" or "estimated" or "approximate" or "predicted," etc.

4 **Complete Response**
Solution includes all four of the required elements.

3 **Substantial Response**
Solution includes three of the required elements.

2 **Developing Response**
Solution includes two of the required elements.

1 **Minimal Response**
Solution includes one of the required elements.

Question 2 (cont'd.)

Notes:

1. If calculations are based on 1-year costs rather than 3-year costs, and then the student chooses A with explanation, the student can earn a score of up to 3.

   Total 1-year cost for A: $10,000 + 12($50) = $10,000 + $600 = $10,600
   And for B:
   Expected number of repairs in 1 year = 0(.5) + 1(.25) + 2(.15) + 3(.1) = .85
   Expected cost of repairs in 1 year = $0(.5) + $200(.25) + $400(.15) + $600(.1) = $170
   Expected 1-year cost for B: $10,500 + $170 = $10,670

2. If initial purchase prices are omitted from the calculations, the student can earn a score of up to 3.

3. Rounded calculations of the expected number of repairs for B: If a student rounds the expected number of repairs per year (.85) to 1, or rounds the expected number of repairs in 3 years (2.55) to 3, the maximum score is a 3. If the student identifies the rounded value as an upper bound on the expected cost, the paper may earn a maximum score of 4.

4. If choice of A or B is not based on expected cost for B, the student can still present a complete response. To earn 4 points with this solution, relevant and complete statistical reasoning must be demonstrated. This solution must include:

   a. Decision based on break-even analysis:

      1 point -- Correct calculation of 3-year fixed cost for A ($11,800)

      1 point -- Correct calculation that 7 <u>or more</u> repairs in the 3-year period would be necessary for B's cost to exceed A's 3-year cost

      1 point -- Says or calculates that the probability of 7 or more repairs for B is small and therefore chooses B,
      OR
      Chooses A because the probability of 7 or more repairs for B is not 0 and they want to guard against the possibility of paying more for B than A's fixed cost

      1 point -- Correctly calculates that the probability of 7 or more repairs for B is 0.01975 (or about 0.02, or about 2 percent of the time B's cost will exceed A's cost)
      AND
      States that this analysis depends on the assumption that repairs from year to year are independent

Question 2 (cont'd.)

b. Decision based on minimax analysis may earn a maximum of 3 points:

1 point -- Correct calculation of 3-year fixed cost for A ($11,800)

1 point -- Correct calculation of range of possible 3-year costs for B
$\$10,500 \leq \text{cost of B} \leq \$12,300$
AND
Probability calculation showing that chance of observing maximum cost is small (e.g., 0.001 or 0.1 percent that B costs $12,500, assuming independence)

1 point -- Relevant statistical justification for choice of A or B:

Gives convincing reasoning for minimizing maximum cost (minimax) and therefore chooses A. Student might argue, for example, that a company may prefer a known fixed cost to a variable cost that could be smaller but also has the chance of being larger.

Question 3 - Solution

Part (a):

1. **Scheme:** Obtain a two-digit random number from the random number table. If it is between 01 and 50, use it to represent the selected ticket. Ignore numbers 00 and 51 – 99.
2. **Stopping Rule:** Determine the amount of the prize associated with the chosen ticket, and add this amount to the total amount awarded so far. If the total amount awarded so far is less than $300, repeat this process.
3. **Count:** Note the total number of winners.
4. **Non-Replacement:** Ignore any ticket number that has already been awarded a prize in this trial**.**

Repeats steps 1 – 4 above a large number of times.

Note: It is OK to also devise a scheme that uses 2 two-digit numbers to represent each ticket (for example, 01 and 51 both representing ticket 1; 02 and 52 both representing ticket 2; etc.) that also addresses the issue of assigning 2 two-digit numbers to each coupon correctly.

Part (b):

Solution will depend on answer to part (a).

For example, using scheme above:

Trial 1			Trial 2			Trial 3		
					Total so far			Total so far
72	ignore	0	02	200	200	06	100	100
74	ignore	0	61	ignore	200	70	ignore	100
91	ignore	0	28	50	250	29	50	150
33	50	50	48	50	300	04	200	350
47	50	100						
65	ignore	100	Total number of winners: 3			Total number of winners: 3		
03	200	300						
Total number of winners: 3								

Students should perform 3 trials. You will have to look at each student response carefully. Some will continue on 1st row, some will use 2nd row for second trial, etc.

Scoring

There are five components to the solution of this problem:

1. **Scheme:** must include a clear, correct statement of the assignment of two-digit random numbers to the coupon numbers (and/or values) AND clear directions as to how the table is to be used in the simulation.
2. **Stopping Rule:** must state that the trial ends when a total value of $300 is attained or exceeded.
3. **Count:** must state or demonstrate that the number of winners is the outcome of the simulation.

Question 3 (cont'd.)

4. **Non-Replacement:** must state that coupon numbers chosen cannot be used more than once in the same trial.
5. **Execution of #1 and #2:** must demonstrate a correct execution of a scheme with a stopping rule.

- **Credit for components #1 and #2, is given for statements in part (a).**
- **Credit for component #3 and/or #4 may be given for statements or demonstrations in parts (a) or (b).**
- **Credit for component #5 is given for a clear demonstration in part (b).**

Scoring Guide:

4 Complete Response
Essentially correct on all five components.

3 Substantial Response
Essentially correct on four of the five of components.

2 Developing Response
Essentially correct on three of five components.

1 Minimal Response
Essentially correct on component #5 only
or
Essentially correct on any two of the other components.

One **INCORRECT** solution is to use the random digit 1 to represent a $200 prize, random digits 2, 3, and 4 to represent a $100 prize, and random digits 6, 7, 8, 9, and 0 to represent a $50 prize.

Then a trial might look like

number	amount	total so far
7	50	50
2	100	150
7	50	200
4	100	300

A student using this scheme can merit **at most a score of 2** for the entire problem.

Question 4 - Solution

Part (a):

Blocking scheme A is preferable because it creates homogeneous blocks with respect to forest exposure. That is, plots in the same block have similar exposure to the forest.

Part (b):

Randomization of varieties of trees to the plots within each block should reduce any possible bias due to confounding variables, such as fertility or moisture, on the productivity of the two types of dwarf trees.

OR

Randomization of varieties of trees to the plots within each block should even out (or equalize) the effect of other characteristics of the plots that might be related to the productivity of the trees.

Scoring

Part (a) is
Essentially correct if:

A statement that blocking scheme A is preferable

AND

A good explanation that gets at the notion of wanting homogeneous experimental units (plots, not trees) within blocks

Partially correct if

Blocking scheme A is chosen with weak or no explanation

OR

Blocking scheme B is chosen and the student clearly demonstrates an understanding that trees of both varieties should appear in plots bordering the forest, and similarly, trees of both varieties should appear in plots away from the forest.

Note: If a student chooses blocking scheme B and indicates that s/he will create blocks within the blocks to deal with forest exposure as well as north/south exposure, part (a) should be scored as essentially correct.

Note: If a student attempts to describe analysis techniques, these should be considered extraneous and should be ignored.

Question 4 (cont'd.)

Part (b) is

Essentially correct if

The student clearly **explains** why randomization is important in the **context** of the problem.

Partially correct if

The student understands that randomization reduces bias and explains it in context, but does not make it clear that a non-random assignment may favor one **variety** of tree.

OR

The student has a correct explanation but contextual interpretation is poor or inappropriate.

Incorrect if

The student uses the word bias, confounding, or other general statistical terms, but does not explain the term(s) in context of the problem.

Note: If the student thinks of blocks as treatment groups (receiving partial credit in part (a)), then part (b) must be **logically consistent**. For example, if a student thinks of the shaded region as one treatment group, it is not sufficient to randomize within the shaded region. The student must address the randomization between the blocks (e.g., flip a coin to assign one variety of tree to one of the blocks and the other variety of tree to the other block).

4 **Complete Response**
Both parts essentially correct

3 **Substantial Response**
One part essentially correct and one part partially correct

2 **Developing Response**
One part essentially correct and one part incorrect

OR

Both parts partially correct

1 **Minimal Response**
One part partially correct and one part incorrect

Question 5 - Solution

Part 1: States a correct pair of hypotheses

μ_G = mean amount of active ingredient for generic
μ_B = mean amount of active ingredient for name brand
μ_D = mean difference in amount of active ingredient

$H_o: \mu_D = 0$ OR $H_o: \mu_G - \mu_B = 0$ OR $H_o: \mu_B - \mu_G = 0$
$H_a: \mu_D \neq 0$ $H_a: \mu_G - \mu_B \neq 0$ $H_a: \mu_B - \mu_G \neq 0$

Part 2: Identifies correct test by name or by formula and checks appropriate assumptions.

Paired t test

$$t = \frac{\bar{x}_D - 0}{\frac{s_D}{\sqrt{n_D}}}$$

Differences: 1, -4, -5, -13, 0, -7, -5, -8, -9, -16

Assumptions: (1) random samples (given) and (2) the population of differences is approximately normal.

Stem and Leaf of differences

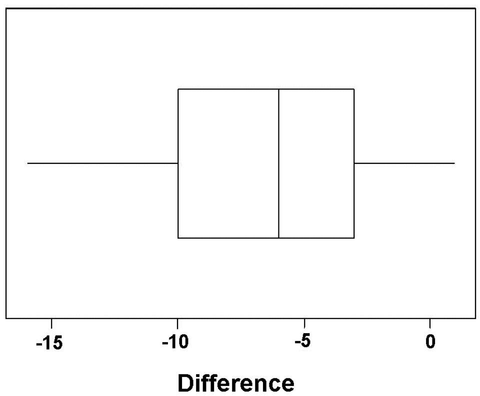
-1 | 6
-1 | 3
-0 | 9 8 7 5 5
-0 | 4
0 | 0 1

Question 5 (cont'd.)

It is reasonable to assume that the population of differences is approximately normal since the stem-and-leaf plot is roughly symmetric with no apparent outliers.

OR

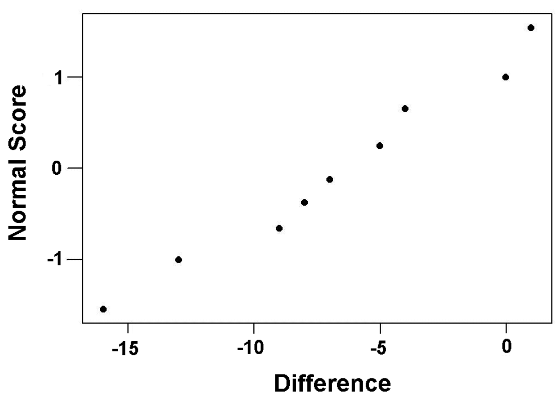
Box Plot of differences



The boxplot is roughly symmetric and shows no outliers. So, it is not unreasonable to assume that the distribution of differences is approximately normal.

OR

Normal probability plot of differences

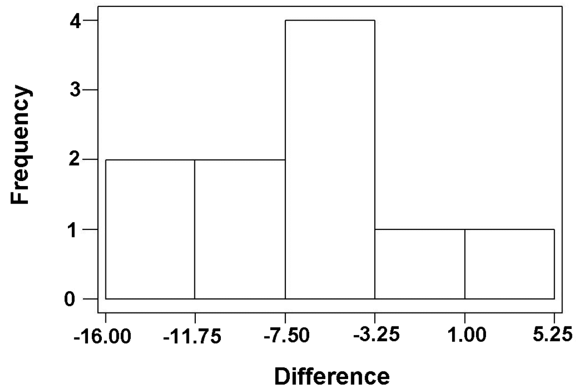


Normal probability plot is reasonably straight, so it is reasonable to assume that the population of differences is approximately normal.

OR

Histogram of differences (using the same intervals as the calculator)

Question 5 (cont'd.)



It is not unreasonable to assume that the population of differences is approximately normal since the histogram is roughly symmetric (especially considering the small sample size).

Note:

- It is acceptable for a student to comment, based on an appropriate graph, that there are no apparent outliers or extreme skewness, without having to mention normality.
- The student should recognize that the assumptions are about the difference distribution.
- If they look at plots of the two samples individually, they must comment on the fact that if it is reasonable to assume that the two population distributions are each approximately normal, then the distribution of the differences will also be approximately normal.

Part 3: Correct mechanics, including value of the test statistic, df, and P-value (or rejection region).

$\overline{x_D}$ = -6.6 $\qquad$ s_D = 5.27

$$t = \frac{-6.6 - 0}{\frac{5.27}{\sqrt{10}}} = \frac{-6.6}{1.67} = -3.96 \qquad \text{df} = n_D - 1 = 9 \qquad \text{P-value} = .00332$$

(Calculator: t = -3.956835797, df = 9, P-value = .0033201462)
(OR rejection region: α = .05, t critical value = ± 2.262
α = .01, t critical value = ± 3.250) (OR 95% C.I. : (-10.37 , -2.827))

Note: If the differences are formed using Name Brand – Generic, the differences, the mean of the differences, and the value of the test statistic will be opposite in sign, but the conclusion will be the same.

Question 5 (cont'd.)

Part 4: Stating a correct conclusion in the context of the problem <u>and</u> making a clear statement of linkage to the results of the statistical test.

Because the P-value is so small (or because P-value < α, or because t is in the rejection region, or because 0 is not contained in the confidence interval), reject H_o. There is evidence that the mean amount of active ingredient is not the same for the name brand and generic drugs.

The consumer group should report that the mean amount of active ingredient is not the same for the name brand and generic drugs.

Note: It is OK if the student says that the consumer group should report that the mean amount of active ingredient is lower for the generic drug **as long as it follows a two-sided conclusion**.

Notes for an incorrect test for means

2 sample t test (unequal variances): t = -4.438 ; df = 17.46; p =.0003391;
95% C.I. (-9.7309,-3.4691)
2 sample t test (equal variances): t = -4.439 ; df= 18; p = .0003172 ;
95% C.I. (-9.724,-3.476)
2 sample t test (conservative approach): t =-4.438 ; df= 9; p =.001627;
95% C.I. (-9.963,-3.236)
2 sample z test: z= -4.44 ; p=.000009 ; 95% C.I. (-9.512,-3.687)

1 sample z test on differences: z = -4.17 ; p = .00003; 95% C.I. (-9.866, -3.334)

Note: A student cannot receive credit for using a Chi-Square test or a test for proportions.

Question 5 (cont'd.)

Scoring for matched pairs t test

Each part of the hypothesis test is either completely correct or incorrect.

4 **Complete Response**
All four parts correct.

3 **Substantial Response**
Three parts correct.

2 **Developing Response**
Two parts correct.

1 **Minimal Response**
One part correct.

Scoring for an incorrect test for means

Each part of the hypothesis test is either completely correct or incorrect.

4 **Complete Response**
Not possible

3 **Substantial Response**
All four parts correct

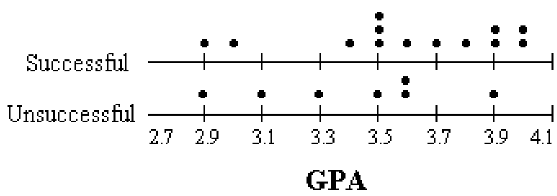
2 **Developing Response**
Three parts correct

1 **Minimal Response**
Two parts correct

Question 6 - Solution

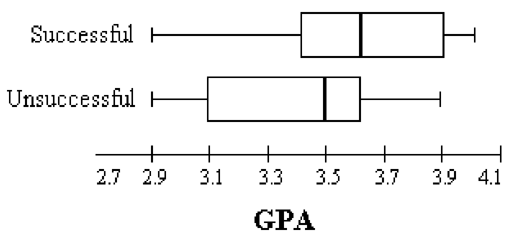
Part (a)

Dotplots, parallel boxplots or back-to-back stem-and-leaf plots:



Although the minimum GPA is the same for both groups, there are more GPAs at the high end for the successful group than those for the unsuccessful group.

OR



Although the minimum GPA is the same for both groups, GPAs in the successful group tend to be higher than the GPAs in the unsuccessful group. The lower quartile, median, and upper quartile are all higher for the successful group than those for the unsuccessful group.

OR

Success		No Success
0	2.9	0
0	3.0	
0	3.1	
	3.2	
	3.3	0
0	3.4	
0 0 0	3.5	0
0	3.6	0 0
0	3.7	
0	3.8	
0 0	3.9	0
0 0	4.0	

OR

Success		No Success
	2	
9	2	9
4 0	3	1 3
9 9 8 7 6 5 5 5 5	3	5 6 6 9
0 0	4	

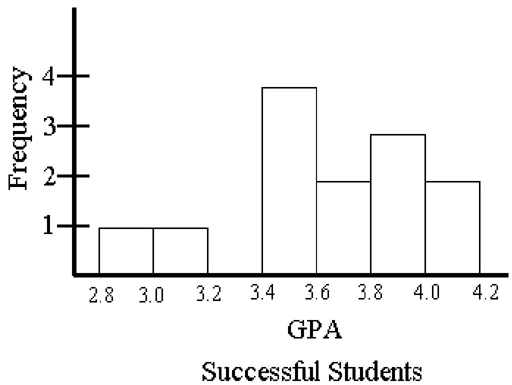
Although the minimum GPA is the same for both groups, there are more GPAs at the high end (bottom) for the successful group than for the unsuccessful group.

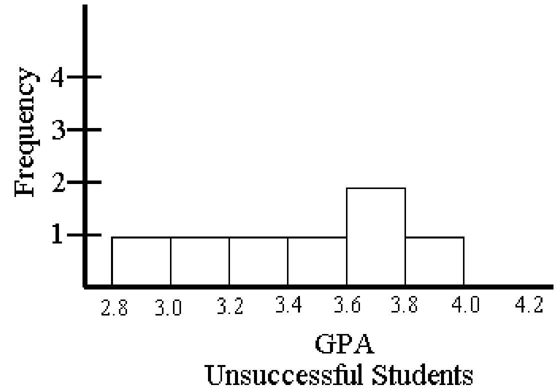
Question 6 (cont'd.)

|2.9| 0 = 2.90 |2| 9 = 2.90

OR

Histograms using the same scale may also be used.



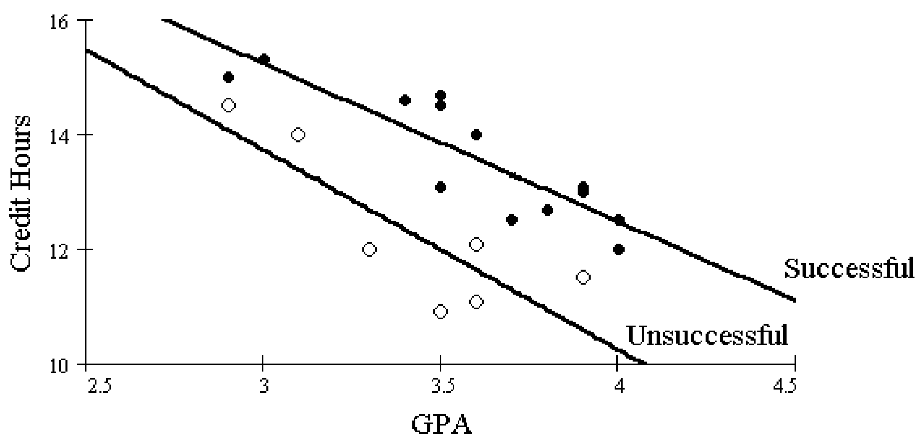


In general, there are more successful students with high GPAs than unsuccessful students with high GPAs. In general, the GPAs for successful students are higher than the GPAs for unsuccessful students.

OR

Scatter plot with linear fits may be used to compare GPA through credit hours. A clear comparison of GPAs must be made. The comparison is more difficult using the linear equations and must be carefully and clearly stated.

Note: If a student presents a scatterplot, the best possible score is partially correct.



For a given number of credit hours, in general, the GPAs of successful students tend to be higher than the GPAs of unsuccessful students.

Question 6 (cont'd.)

Note: (a) A graph with no comment is incorrect.
(b) A correct comment that clearly indicates what graphical display was used, but the display is not shown, is partially correct.

Part (b):

Part 1: States a correct pair of hypotheses in symbols or in words

$$H_0: \beta = 0 \qquad \qquad H_0: \rho = 0$$
$$\text{OR}$$
$$H_a: \beta \neq 0 \qquad \qquad H_a: \rho \neq 0$$

OR

H_0: There is no relationship between GPA and mean number of credit hours
H_a: There is a relationship between GPA and mean number of credit hours

Part 2: Identifies correct test by name or by formula and states that the appropriate assumptions were met.

t test for slope or linear regression t-test (LinReg t test) or gives t from the computer output.

$$t = \frac{\hat{\beta} - 0}{s_{\hat{\beta}}}$$

(Calculator output: t = -5.903, p = 0.0001026)

Problem states that assumptions necessary for inference were judged to be reasonable. This should be mentioned, but then there is no need to go further with checking of assumptions. (Correct assumptions are: random sample, the residuals from the line are normally distributed with mean zero and constant variance.)

Part 3: Correct mechanics, including value of test statistic and P-value (or rejection region)

The computer output provided: t = -5.90 P-value = .000

Part 4: States a correct conclusion in the context of the problem, linked to the P-value or rejection region.

Since the P-value is 0.000, reject the null hypothesis. For students who successfully completed the program, there is a significant relationship between GPA and mean number of credit hours per semester.

Note: If conclusion immediately follows P-value, the linkage is implied.

Note: If the student just appeals to the P-value and then goes directly to a correct conclusion, part (b) should be scored as partially correct.

Note: If the student appeals to high correlation as evidence of a significant relationship, part (b) should be scored as incorrect.

Question 6 (cont'd.)

Note: If the student indicates that the assumptions necessary for inference have been met or that a linear equation is a good fit, and appeals to the high correlation, part (b) should be scored as partially correct.

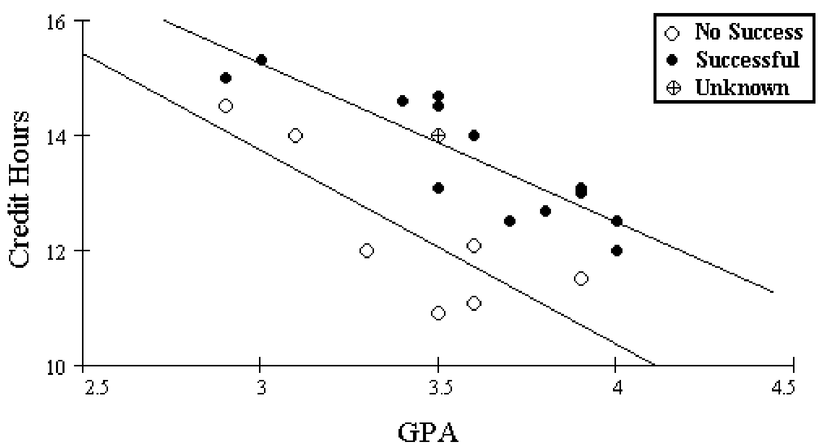
Part (c):

For a complete response, the student must look at the data as bivariate (combine information from both GPA and credit hours) and compare to both successful and unsuccessful students. It is not enough to consider just GPA or just number of credit hours or to look at GPA and credit hours for successful students only.

This question is not asking for a general solution to a variety of classification problems. It is asking students to create from their limited knowledge of statistics, a reasonable, statistically supported solution for these data.

Approach #1 – Scatterplot or regression lines drawn, with two groups indicated

Plotting the "new applicant point" on the scatterplot, it appears that the data point is more similar to the successful students than the unsuccessful students.



Approach #2 – Using the regression lines and comparing predicted values to 14 or comparing residuals

Predicted number of credit hours for a GPA of 3.5

For the successful group predicted hours = 23.514 - 2.7555(3.5) = 13.86975
For the unsuccessful group predicted hours = 24.200 – 3.485(3.5) = 12.0025

The actual value of 14 is much closer to the prediction for the successful group than that for the unsuccessful group. Therefore, we believe this student will be successful.

Question 6 (cont'd.)

Approach #3 – Add the point (3.5, 14) to each set of data and compare the effect on the correlation coefficient.

Successful without (3.5, 14)	r = -0.8718.
Successful with (3.5, 14)	r = -0.8725.
Unsuccessful without (3.5, 14)	r = -0.8384.
Unsuccessful with (3.5, 14)	r = -0.718.

The point (3.5, 14) is more consistent with the data for successful students since the correlation is essentially unchanged by including this point in the data while the correlation for unsuccessful students is substantially weaker.

Note: Correct conclusions using bivariate analysis supported by a statistical argument should be scored at least partially correct.

Scoring

Part (a) is

Essentially correct if

1. Gives a correct graphical display with appropriate scales and labels.

   AND

2. Draws some conclusion about GPAs for the two groups from the graphical presentation,

   OR

Uses the graphical display to comment on at least two of center, shape and spread, or some equivalent comparison.

Partially correct if

Constructs correct graphs but gives weak or partial comparisons.

OR

Does not show graphs or shows graphs that do not meet (1) above, but gives a reasonable interpretation based on the given or implied graphs.

Incorrect if

Constructs correct graph but makes no comments.

Question 6 (cont'd.)

Part (b) is

Essentially correct if

All 4 parts of the appropriate test are correct.

Part 1:	Hypotheses
Part 2:	Identifies test
Part 3:	Mechanics
Part 4:	Conclusion

Partially correct if

2 or 3 parts of the test are correct

OR

Student does not carry out a formal test, but correctly reasons from P-value on output to a correct conclusion.

OR

Student indicates that the assumptions necessary for inference have been met or that a linear equation is a good fit, and appeals to the high correlation.

Part (c)

Part (c) must be scored holistically.

To be considered essentially correct, it must be treated as a bivariate problem—both variables (GPA and credit hours) must be used. Also, the possibility of inclusion must be checked for both groups. The correct prediction (successful) must follow from the student's analysis.

The solution should be considered partially correct if it is treated as a bivariate problem but the communication is weak or the statistical justification is weak or difficult to follow.

The solution should be considered partially correct if it is treated as a bivariate problem but only one group (successful or unsuccessful) is used and the point is never considered in relation to the other group.

If a decision is made based on looking at either of the two variables individually, part (c) is incorrect.

Question 6 (cont'd.)

For parts (a) and (b), essentially correct responses count as 1 part and partially correct responses count as ½ part. For part (c), an essentially correct response counts as 2 parts and a partially correct response counts as 1 part.

4 **Complete Response**
Four parts.

3 **Substantial Response**
Three parts.

2 **Developing Response**
Two parts.

1 **Minimal Response**
One part.

If a paper is between two scores (for example, 2 ½ parts) use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.