Free-Response Questions

In the free-response section of the AP Statistics Examination, students are asked to answer five questions and complete an investigative task. Each question is designed to be answered in approximately 12 minutes. The longer investigative task is designed to be answered in approximately 30 minutes.

Statistics is a discipline in which clear and complete communication is an essential skill. The free-response questions on the AP Statistics Examination require students to use their analytical, organizational, and communication skills to formulate cogent answers and provide students with an opportunity to:

- Relate two or more different content areas (i.e., elementary data analysis, experimental design and sampling, probability, and statistical inference) as they formulate a complete response or solution to a statistics or probability problem, and
- Demonstrate their mastery of statistics in a response format that permits the students to determine *how* they will organize and present each response.

The purpose of the investigative task is not only to evaluate the student's understanding in several content areas, but also to assess his or her ability to integrate statistical ideas and apply them in a new context or in a nonroutine way.

Scoring of Free-Response Questions

The evaluation of student responses on the free-response section of the AP Statistics Examination reflects the dual importance of statistical knowledge and good communication. The free-response questions and the investigative task are scored "holistically"; that is, each question's response is evaluated as 'a complete package'. With holistic scoring, after reading through the details of a student's response, a judgment is made about the *overall quality* of the response, as opposed to "analytic" scoring, wherein the individual components to be evaluated in a student's response are specified in advance, and each component is given a value counting toward the overall score.

Holistic scoring is well suited for questions wherein the student is required to synthesize information and respond at least partially in written paragraphs, and for questions that could potentially generate multiple, and diverse, but equally correct, responses. For example, an open-ended question may present data from a real life study and ask the student not only to analyze the data, but also to comment on how the study's protocol might be improved. Comments on improving the protocol might focus on improving the sampling method, controlling confounding variables, or seeking more power by increasing the sample size. In this context, holistic scoring represents a recognition not only of the existence of multiple reasonable approaches to a statistical analysis, but a realization of the existence of a statistical synergy — i.e., that a quality student response is more than just the sum of its parts.

The AP Statistics scoring rubric for each free response question has five categories, numerically scored on a 0 to 4 scale. Each of these categories represents a level of quality in the student response. These levels of quality are defined on two dimensions: statistical knowledge and communication. The specific rubrics for each question are tied to a general template, which represents the descriptions of the quality levels as envisioned by the Development Committee. This general template is given in the following table, "A Guide to Scoring Free Response Statistics Questions."

A GUIDE TO SCORING FREE-RESPONSE STATISTICS QUESTIONS: THE CATEGORY DESCRIPTORS

Score Descriptors	Statistical Knowledge	Communication
	Identification of the important components of the problem Demonstration of the statistical concepts and techniques that result in a correct solution of the problem	Explanation of what was done and why, along with a statement of conclusions drawn
4 Complete	• shows complete understanding of the problem's statistical components • synthesizes a correct relationship among these components, perhaps with novelty and creativity • uses appropriate and correctly executed statistical techniques • May have minor arithmetic errors, but answers are still reasonable	• provides a clear, organized, and complete explanation, using correct terminology, of what was done and why • states appropriate assumptions and caveats • uses diagrams or plots when appropriate to aid in describing the solution • states an appropriate and complete conclusion

3 Substantial	• shows substantial understanding of the problem's statistical components • synthesizes a relationship among these components, perhaps with minor gaps • uses appropriate statistical techniques • may have arithmetic errors, but answers are still reasonable	• provides a clear but not perfectly organized explanation, using correct terminology, of what was done and why, but explanation may be slightly incomplete • may miss necessary assumptions or caveats • uses diagrams or plots when appropriate to aid in describing the solution • states a conclusion that follows from the analysis but may be somewhat incomplete
2 Developing	• shows some understanding of the problem's statistical components • shows little in the way of a relationship among these components • uses some appropriate statistical techniques, but misses or misuses others • may have arithmetic errors that result in unreasonable answers	• provides some explanation of what was done, but explanation may be vague and difficult to interpret and terminology may be somewhat inappropriate • uses diagrams in an incomplete or ineffective way, or diagrams may be missing • states a conclusion that is incomplete

1 Minimal	• shows limited understanding of the problem's statistical components by failing to identify important components • shows little ability to organize a solution and may use irrelevant information • misuses or fails to use appropriate statistical techniques • has arithmetic errors that result in unreasonable answers	• provides minimal or unclear explanation of what was done or why it was done, and explanation may not match the presented solution • fails to use diagrams or plots, or uses them incorrectly • states an incorrect conclusion or fails to state a conclusion
0	• shows little to no understanding of statistical components	• provides no explanation of a legitimate strategy

Some important points that students should remember when answering free-response questions on the AP Statistics Examination are given below.

1. Read the questions carefully and answer them in context; for example, the results of a hypothesis test should always be followed by a conclusion in context and a confidence interval should always be followed by an interpretation of the interval in context. Explanations and conclusions in context are always required for a complete answer.
2. Know the vocabulary of statistics, and use that vocabulary correctly in all written responses.
3. Remember to define all symbols. Specifically, remember to distinguish between population parameters and sample statistics.
4. Remember to state and check all necessary assumptions when performing hypothesis tests and constructing interval estimates.
5. Be able to interpret data displayed in a variety of ways, including graphical and in computer outputs. Be able to represent data in a variety of forms and base sound statistical arguments on these representations.

AP Central contains free-response questions, rubrics, and selected student responses from past AP Statistics exams. This is an excellent place to become more familiar with the content of past free-response questions and how they were scored.

The following questions are examples of free-response questions. These questions were administered as part of a previous year's exam.

1. The summary statistics for the number of inches of rainfall in Los Angeles for 117 years, beginning in 1877, are shown below.

N	MEAN	MEDIAN	TRMEAN	STDEV	SE MEAN
117	14.941	13.070	14.416	6.747	0.624

MIN	MAX	Q1	Q3
4.850	38.180	9.680	19.250

   (a) Describe a procedure that uses these summary statistics to determine whether there are outliers.
   (b) Are there outliers in these data?_____
   Justify your answer based on the procedure that you described in part (a).
   (c) The news media reported that in a particular year, there were <u>only</u> 10 inches of rainfall. Use the information provided to comment on this reported statement.

2. A department supervisor is considering purchasing one of two comparable photocopy machines, *A* or *B*. Machine *A* costs \$10,000 and machine *B* costs \$10,500. This department replaces photocopy machines every three years. The repair contract for machine *A* costs \$50 per month and covers an unlimited number of repairs. The repair contract for machine *B* costs \$200 per repair. Based on past performance, the distribution of the number of repairs needed over any one-year period for machine *B* is shown below.

Number of Repairs	0	1	2	3
Probability	0.50	0.25	0.15	0.10

   You are asked to give a recommendation based on overall cost as to which machine, *A* or *B*, along with its repair contract, should be purchased. What would your recommendation be? Give a statistical justification to support your recommendation.

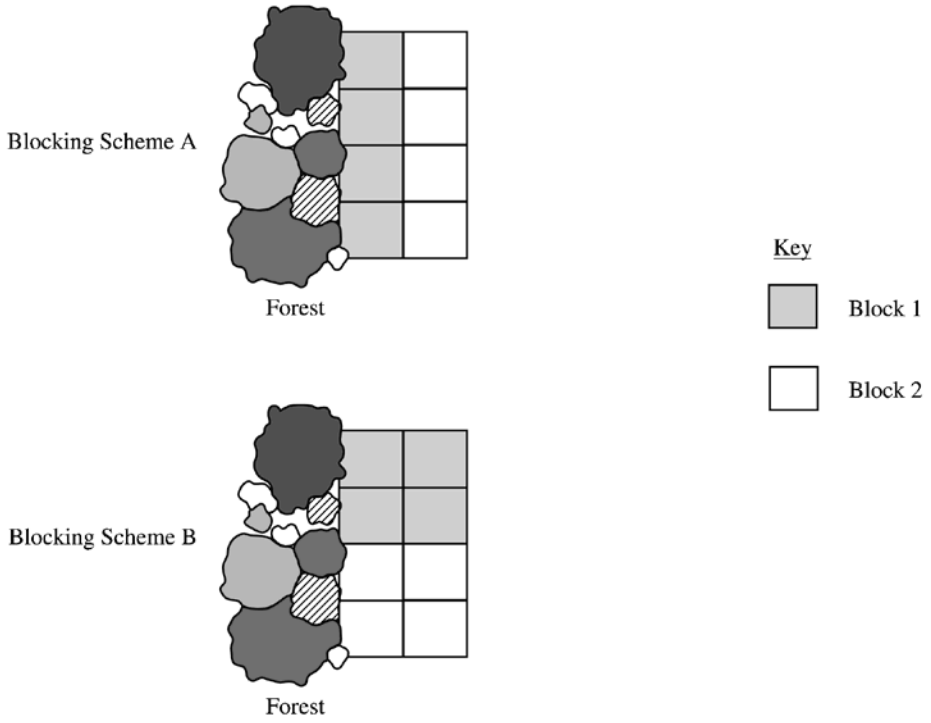
3. Every Monday a local radio station gives coupons away to 50 people who correctly answer a question about a news fact from the previous day's newspaper. The coupons given away are numbered from 1 to 50, with the first person receiving coupon 1, the second person receiving coupon 2, and so on, until all 50 coupons are given away. On the following Saturday, the radio station randomly draws numbers from 1 to 50 and awards cash prizes to the holders of the coupons with these numbers. Numbers continue to be drawn without replacement until the total amount awarded first equals or exceeds $300. If selected, coupons 1 through 5 each have a cash value of $200, coupons 6 through 20 each have a cash value of $100, and coupons 21 through 50 each have a cash value of $50.

   (a) Explain how you would conduct a simulation using the random number table provided below to estimate the distribution of the number of prize winners each week.
   (b) Perform your simulation three times. (That is, run three trials of your simulation.) Start at the leftmost digit in the first row of the table and move across. Make your procedure clear so that someone can follow what you did. You must do this by marking directly on or above the table. Report the number of winners in each of your three trials.

72749	13347	65030	26128	49067	02904	49953	74674	94617	13317
81638	36566	42709	33717	59943	12027	46547	61303	46699	76423
38449	46438	91579	01907	72146	05764	22400	94490	49833	09258

4. Students are designing an experiment to compare the productivity of two varieties of dwarf fruit trees. The site for the experiment is a field that is bordered by a densely forested area on the west (left) side. The field has been divided into eight plots of approximately the same area. The students have decided that the test plots should be blocked. Four trees, two of each of the two varieties, will be assigned at random to the four plots within each block, with one tree planted in each plot.

   The two blocking schemes shown below are under consideration. For each scheme, one block is indicated by the white region and the other block is indicated by the gray region in the figures.



   (a) Which of the blocking schemes, A or B, is better for this experiment? Explain your answer.
   (b) Even though the students have decided to block, they must randomly assign the varieties of trees to the plots within each block. What is the purpose of this randomization in the context of this experiment?

5. A growing number of employers are trying to hold down the costs that they pay for medical insurance for their employees. As part of this effort, many medical insurance companies are now requiring clients to use generic brand medicines when filling prescriptions. An independent consumer advocacy group wanted to determine if there was a difference, in milligrams, in the amount of active ingredient between a certain "name" brand drug and its generic counterpart. Pharmacies may store drugs under different conditions. Therefore, the consumer group randomly selected ten different pharmacies in a large city and filled two prescriptions at each of these pharmacies, one for the "name" brand and the other for the generic brand of the drug. The consumer group's laboratory then tested a randomly selected pill from each prescription to determine the amount of active ingredient in the pill. The results are given in the following table.

ACTIVE INGREDIENT
(in milligrams)

Pharmacy	1	2	3	4	5	6	7	8	9	10
Name brand	245	244	240	250	243	246	246	246	247	250
Generic brand	246	240	235	237	243	239	241	238	238	234

Based on these results, what should the consumer group's laboratory report about the difference in the active ingredient in the two brands of pills? Give appropriate statistical evidence to support your response.

6. The statistics department at a large university is trying to determine if it is possible to predict whether an applicant will successfully complete the Ph.D. program or will leave before completing the program. The department is considering whether GPA (grade point average) in undergraduate statistics and mathematics courses (a measure of performance) and mean number of credit hours per semester (a measure of workload) would be helpful measures. To gather data, a random sample of 20 entering students from the past 5 years is taken. The data are given below.

Successfully Completed Ph.D. Program

Student	A	B	C	D	E	F	G	H	I	J	K	L	M
GPA	3.8	3.5	4.0	3.9	2.9	3.5	3.5	4.0	3.9	3.0	3.4	3.7	3.6
Credit hours	12.7	13.1	12.5	13.0	15.0	14.7	14.5	12.0	13.1	15.3	14.6	12.5	14.0

Did Not Complete Ph.D. Program

Student	N	O	P	Q	R	S	T
GPA	3.6	2.9	3.1	3.5	3.9	3.6	3.3
Credit hours	11.1	14.5	14.0	10.9	11.5	12.1	12.0

The regression output below resulted from fitting a line to the data in each group. The residual plots (not shown) indicated no unusual patterns, and the assumptions necessary for inference were judged to be reasonable.

Successfully Completed Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	23.514	1.684	13.95	0.000
GPA	−2.7555	0.4668	−5.90	0.000
S = 0.5658	R-Sq = 76.0%			

Did Not Complete Ph.D. Program

Predictor	Coef	StDev	T	P
Constant	24.200	3.474	6.97	0.001
GPA	−3.485	1.013	−3.44	0.018
S = 0.8408	R-Sq = 70.3%			

(a) Use an appropriate graphical display to compare the GPA's for the two groups. Write a few sentences commenting on your display.

(b) For the students who successfully completed the Ph.D. program, is there a significant relationship between GPA and mean number of credit hours per semester? Give a statistical justification to support your response.
(c) If a new applicant has a GPA of 3.5 and a mean number of credit hours per semester of 14.0, do you think this applicant will successfully complete the Ph.D. program? Give a statistical justification to support your response.