AP® Statistics
2001 Scoring Commentary

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.

The College Board is a national nonprofit membership association dedicated to preparing, inspiring, and connecting students to college and opportunity. Founded in 1900, the association is composed of more than 3,900 schools, colleges, universities, and other educational organizations. Each year, the College Board serves over three million students and their parents, 22,000 high schools, and 3,500 colleges, through major programs and services in college admission, guidance, assessment, financial aid, enrollment, and teaching and learning. Among its best-known programs are the SAT®, the PSAT/NMSQT™, the Advanced Placement Program® (AP®), and Pacesetter®. The College Board is committed to the principles of equity and excellence, and that commitment is embodied in all of its programs, services, activities, and concerns.

AP® STATISTICS
2001 SCORING COMMENTARY

Question 1

Response: 1 of 3
Score: 4
This response earned a score of 4 because it provided a good description of one of the standard procedures for finding outliers, justified why the maximum is an outlier on the upper end, and justified why 10 inches of rainfall was not unusual by specifying where 10 is relative to the entire data set.

Response: 2 of 3
Score: 4
In this response, the student clearly described one of the standard procedures for finding outliers, checked for outliers on both the upper and lower ends, and justified why 10 inches of rainfall is not rare using a *z*-score approach.

Response: 3 of 3
Score: 3
The student used the standard IQR rule for finding outliers, and indicated that there could be multiple outliers on the upper end. However, in the last part of the question, the use of the normal approximation to the distribution of rainfall is incorrect because the summary statistics indicate the distribution is not normal.

Question 2

Response: 1 of 3
Score: 4
The student computed the fixed and expected costs, compared the two values and made a correct conclusion. The student demonstrated careful and complete statistical reasoning by recognizing that the cost of B is a random variable and that one needs to assume its need for repairs in the future will be similar to its need for repairs in the past.

Response: 2 of 3
Score: 3
This response earned a score of 3 because it did not state that the assumption of independence was used to compute the probability of 7 or more repairs. Otherwise, a correct procedure was used.

Response: 3 of 3
Score: 3
This response presented a complete solution, but because the problem stated that the department replaces photocopy machines every 3 years, the solution should have addressed the 3-year costs rather than just the first-year costs. Therefore, it earned a score of 3.

Question 3

Response: 1 of 3
Score: 4
All five components of the solution in this response are clearly stated or demonstrated. This student used a different row of the random number table for each trial. Note that in part (b), the second "38" in row 3 of the table was disregarded, consistent with the non-repetition statement in part (a).

Response: 2 of 3
Score: 4
In this response, the student assigned the random digits 00-99 to the 50 coupons and stated a correct rule to address the fact that two random digits were assigned to each coupon number. The student displayed the reduction of the double assignment to the numbers 01-50 in part (b). The separate issue of non-replacement was also addressed successfully.

Response: 3 of 3
Score: 3
The student clearly discussed or demonstrated every component of the solution in the response except non-replacement. Note that there is no mention of repeated values in part (a) and the second "38" in row 3 of the table is not disregarded in part (b).

Question 4

Response: 1 of 3
Score: 4
While making a case for scheme A, the student described an undesirable possibility that could occur under scheme B, namely that all 4 trees of one variety could end up next to the forest. In part (b), the student understood that "the element of landscape" (meaning, proximity to the forest) is not a problem that needs to be solved with randomization since the blocking scheme chosen in part (a) spreads the effect of the forest equally among the two varieties. The student cited several possible confounding variables, and stated that randomization "evens out," rather than eliminates, these factors.

Response: 2 of 3
Score: 4
The student made the correct selection and justified the answer by appealing to the homogeneity of the blocks in scheme A and the lack of homogeneity of the blocks in scheme B. Later, the student acknowledged that there will still be differences in the plots that may affect the productivity of the trees, understanding that randomization will not make the differences disappear, but rather that randomization spreads these "inequities" between the two varieties, so that one variety should not be favored over the other.

Response: 3 of 3
Score: 3
In this response, even though the student improperly used the term "blocks" to mean "treatment groups," there was understanding that the purpose of blocking is to reduce the effects of the forest by giving each treatment group equal exposure to the forest. Consistent with the choice of scheme B in part (a), the student made random assignments of tree variety so that both varieties were given an equal chance to be planted in the upper or lower treatment group.

Question 5

Response: 1 of 3
Score: 4
The response correctly states the hypotheses with respect to mean difference. Two different graphical displays are used to check and comment on the reasonable use of the matched pairs *t*-test with respect to normality of the differences. Work is given for the correct *t*-test statistic and the correct *p*-value is stated. The conclusion is well formulated.

Response: 2 of 3
Score: 3
This response earned a score of 3 because the student shows a one-sided alternative hypothesis, when a two-sided was needed. However, the student correctly checked and commented on the normality assumption of the differences, used the correct matched pairs *t*-test statistic, and stated the correct *p*-value for a one-sided alternative hypothesis.

Response: 3 of 3
Score: 3
Although the mechanics and conclusion are correct, the student fails to make the connection that if two separate samples are reasonably normal, then the distribution of differences will be reasonably normal as well. There are errors in the student's notation.

Question 6

Response: 1 of 3
Score: 4
The solution includes comments on the center, shape, and spread of the GPAs using a back-to-back stemplot in part (a), inference about the relationship between the GPA and mean number of credit hours per semester using a *t*-test for slope in part (b), and a conclusion using a bivariate approach based on a comparison of residuals that includes both (successful and non-successful) groups in part (c).

Response: 2 of 3
Score: 3
The solution includes comments on the GPAs using parallel boxplots in part (a), inference about the relationship between the GPA and mean number of credit hours per semester using a *t*-test for slope in part (b), but the conclusion using a bivariate approach based on the fitted value includes only the successful group in part (c). The student fails to check the possibility of new applicants not completing the Ph.D. program.

Response: 3 of 3
Score: 3
The boxplots in part (a) are acceptable but communication in this part is weak. The statistical justification using an appropriate inferential procedure about the relationship between GPA and mean number of credit hours per semester is not presented in part (b).