STATISTICS
SECTION II
Part A
Questions 1-5

Spend about 65 minutes on this part of the exam.

Percent of Section II score—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be scored on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

1. An environmental group conducted a study to determine whether crows in a certain region were ingesting food containing unhealthy levels of lead. A biologist classified lead levels greater than 6.0 parts per million (ppm) as unhealthy. The lead levels of a random sample of 23 crows in the region were measured and recorded. The data are shown in the stemplot below.

Lead Levels

2	8
3	0
3	5 8 8
4	1 1 2
4	6 8 8
5	0 1 2 2 3 4
5	9 9
6	3 4
6	6 8

Key: 2|8 = 2.8 ppm

(a) What proportion of crows in the sample had lead levels that are classified by the biologist as unhealthy?

(b) The mean lead level of the 23 crows in the sample was 4.90 ppm and the standard deviation was 1.12 ppm. Construct and interpret a 95 percent confidence interval for the mean lead level of crows in the region.

2. An administrator at a large university wants to conduct a survey to estimate the proportion of students who are satisfied with the appearance of the university buildings and grounds. The administrator is considering three methods of obtaining a sample of 500 students from the 70,000 students at the university.

   (a) Because of financial constraints, the first method the administrator is considering consists of taking a convenience sample to keep the expenses low. A very large number of students will attend the first football game of the season, and the first 500 students who enter the football stadium could be used as a sample. Why might such a sampling method be biased in producing an estimate of the proportion of students who are satisfied with the appearance of the buildings and grounds?

   (b) Because of the large number of students at the university, the second method the administrator is considering consists of using a computer with a random number generator to select a simple random sample of 500 students from a list of 70,000 student names. Describe how to implement such a method.

   (c) Because stratification can often provide a more precise estimate than a simple random sample, the third method the administrator is considering consists of selecting a stratified random sample of 500 students. The university has two campuses with male and female students at each campus. Under what circumstance(s) would stratification by campus provide a more precise estimate of the proportion of students who are satisfied with the appearance of the university buildings and grounds than stratification by gender?

---

3. Each full carton of Grade A eggs consists of 1 randomly selected empty cardboard container and 12 randomly selected eggs. The weights of such full cartons are approximately normally distributed with a mean of 840 grams and a standard deviation of 7.9 grams.

   (a) What is the probability that a randomly selected full carton of Grade A eggs will weigh more than 850 grams?

   (b) The weights of the empty cardboard containers have a mean of 20 grams and a standard deviation of 1.7 grams. It is reasonable to assume independence between the weights of the empty cardboard containers and the weights of the eggs. It is also reasonable to assume independence among the weights of the 12 eggs that are randomly selected for a full carton.

   Let the random variable X be the weight of a single randomly selected Grade A egg.

   i) What is the mean of X ?

   ii) What is the standard deviation of X ?

AP® STATISTICS
2013 SCORING GUIDELINES

Question 1

Intent of Question

The primary goals of this question were to assess a student's ability to (1) use a stem-and-leaf plot to answer a question about a distribution of data; (2) identify and compute an appropriate confidence interval after checking the necessary conditions; and (3) interpret the interval in the context of the data.

Solution

Part (a):

Four of the 23 crows in the sample had a lead level greater than 6.0 ppm. Therefore, the proportion of crows in the sample that were classified as unhealthy is $\frac{4}{23} \approx 0.174$.

Part (b):

Step 1: Identifies the appropriate confidence interval (by name or by formula) and checks appropriate conditions.

The appropriate procedure is a one-sample *t*-interval for a population mean.

Conditions: 1. The sample is randomly selected from the population.
2. The population has a normal distribution, or the sample size is large.

The first condition is met because we were told that the crows were randomly selected. The sample size of 23 is not considered large, so we need to examine the sample data to assess whether it is reasonable to assume that the population distribution of lead levels for all crows in this region is normal. The stem-and-leaf plot shows no strong skewness or outliers, so we will consider the second condition to be met.

Step 2: Correct mechanics

A 95% confidence interval for the population mean μ is given by: $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$. The critical value for 95% confidence, based on $23 - 1 = 22$ degrees of freedom, is $t^* = 2.074$. The 95% confidence interval for μ is therefore

$$4.90 \pm 2.074 \times \frac{1.12}{\sqrt{23}} \approx 4.90 \pm 0.484,$$

which is the interval $(4.416, 5.384)$ ppm.

Using the raw data rather than the given summary statistics, the 95% confidence interval for μ is (4.411, 5.3803).

Step 3: Interpretation

We can be 95% confident that the population mean lead level among all crows in this region is between 4.416 and 5.384 parts per million.

AP® STATISTICS
2013 SCORING GUIDELINES

Question 1 (continued)

Scoring

This question is scored in four sections. Section 1 consists of part (a); section 2 consists of part (b), step 1; section 3 consists of part (b), step 2; and section 4 consists of part (b), step 3. Each section is scored as essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is scored as follows:

Essentially correct (E) if the response provides the correct numerical answer, as a decimal or as a fraction, with work shown.

Partially correct (P) if the response provides the correct numerical answer as a decimal but does not show the fraction that produced the answer, *OR* shows a fraction with the correct numerator but an incorrect denominator, *OR* shows the correct fraction but computes an incorrect answer.

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

Section 2 is scored as follows:

Essentially correct (E) if the response identifies a one-sample *t*-interval for a population mean (either by name or formula) *AND* also checks *both* the random sampling and the normality/large sample condition correctly.

Note: Any reasonable comment about the normality displayed in the stem-and-leaf plot (or another appropriately sketched plot) is acceptable.

Partially correct (P) if the response identifies the correct procedure *AND* checks only one of the two conditions correctly *OR* does not identify the correct procedure but does check both conditions correctly.

Incorrect (I) if the response identifies the correct procedure but does not check conditions correctly *OR* does not identify the correct procedure and checks at most one condition correctly.

Section 3 is scored as follows:

Essentially correct (E) if the response gives the correct confidence interval. Supporting work is not required, but if included, it must be correct.

Partially correct (P) if the response gives an incorrect but reasonable confidence interval with appropriate supporting work shown *OR* gives a correct confidence interval with incorrect (but appropriate) supporting work shown.

Note: If the response identifies a one-sample *z*-interval as the correct procedure in Section 2, then the response earns a P in Section 3 if this interval is calculated correctly.

Incorrect (I) if the response makes use of an inappropriate procedure for a confidence interval about a population mean.

Question 1 (continued)

Section 4 is scored as follows:

Essentially correct (E) if the response gives a reasonable interpretation of the interval that includes four elements:

1. Estimating a mean
2. Inference about a population
3. 95% confidence
4. Context (lead level/ppm and crows).

Partially correct (P) if the response gives a reasonable interpretation of the interval that includes both of the first two elements and one of the last two elements;

OR

if the response gives a correct interpretation of the confidence level in context (lead level/ppm and crows) but does not attempt to interpret the confidence interval.

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

Each essentially correct (E) response counts as 1 point, and a partially correct (P) response 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 strength of the response and communication.