5. A recent report stated that less than 35 percent of the adult residents in a certain city will be able to pass a physical fitness test. Consequently, the city's Recreation Department is trying to convince the City Council to fund more physical fitness programs. The council is facing budget constraints and is skeptical of the report. The council will fund more physical fitness programs only if the Recreation Department can provide convincing evidence that the report is true.

   The Recreation Department plans to collect data from a sample of 185 adult residents in the city. A test of significance will be conducted at a significance level of $\alpha = 0.05$ for the following hypotheses.

$$H_0 : p = 0.35$$
$$H_a : p < 0.35,$$

   where p is the proportion of adult residents in the city who are able to pass the physical fitness test.

   (a) Describe what a Type II error would be in the context of the study, and also describe a consequence of making this type of error.

   (b) The Recreation Department recruits 185 adult residents who volunteer to take the physical fitness test. The test is passed by 77 of the 185 volunteers, resulting in a p-value of 0.97 for the hypotheses stated above. If it was reasonable to conduct a test of significance for the hypotheses stated above using the data collected from the 185 volunteers, what would the p-value of 0.97 lead you to conclude?

   (c) Describe the primary flaw in the study described in part (b), and explain why it is a concern.

STATISTICS
SECTION II
Part B
Question 6

Spend about 25 minutes on this part of the exam.
Percent of Section II score—25

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.

6. Two students at a large high school, Peter and Rania, wanted to estimate μ, the mean number of soft drinks that a student at their school consumes in a week. A complete roster of the names and genders for the 2,000 students at their school was available. Peter selected a simple random sample of 100 students. Rania, knowing that 60 percent of the students at the school are female, selected a simple random sample of 60 females and an independent simple random sample of 40 males. Both asked all of the students in their samples how many soft drinks they typically consume in a week.

   (a) Describe a method Peter could have used to select a simple random sample of 100 students from the school.

   Peter and Rania conducted their studies as described. Peter used the sample mean $\overline{X}$ as a point estimator for μ. Rania used $\overline{X}_{\text{overall}} = (0.6)\overline{X}_{\text{female}} + (0.4)\overline{X}_{\text{male}}$ as a point estimator for μ, where $\overline{X}_{\text{female}}$ is the mean of the sample of 60 females and $\overline{X}_{\text{male}}$ is the mean of the sample of 40 males.

   Summary statistics for Peter's data are shown in the table below.

Variable	N	Mean	Standard Deviation
Number of soft drinks	100	5.32	4.13

   (b) Based on the summary statistics, calculate the estimated standard deviation of the sampling distribution (sometimes called the standard error) of Peter's point estimator $\overline{X}$.

AP® STATISTICS
2012 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this question were to assess students' ability to (1) describe a Type II error and its consequence in a particular study; (2) draw an appropriate conclusion from a *p*-value; (3) describe a flaw in a study and its effect on inference from a sample to a population.

Solution

Part (a):

In the context of the study, a Type II error means failing to reject the null hypothesis that 35 percent of adult residents in the city are able to pass the test when, in reality, less than 35 percent are able to pass the test. The consequence of this error is that the council would not fund the program, and the city would continue to have a smaller proportion of physically fit residents than the council would like.

Part (b):

Because the *p*-value of 0.97 is larger than $\alpha = 0.05$, we fail to reject the null hypothesis. There is not convincing evidence that the proportion of adult residents in the city who are able to pass the physical fitness test is less than 0.35. After all, the sample proportion of $\hat{p} = 0.416$ is actually higher than 0.35, which is in the opposite direction of the alternative hypothesis.

Part (c):

This is not a randomly selected sample because the sample was selected by recruiting volunteers. It seems reasonable to think that volunteers would be more physically fit than the population of city adults as a whole. Therefore, the sample proportion will likely overestimate the population proportion of adult residents in the city who are able to pass the physical fitness test.

Scoring

Parts (a), (b), and (c) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as follows:

Essentially correct (E) if the response correctly completes the following two components:

1. Describes the error in context by referring to the proportion of adult residents in the city who are able to pass the physical fitness test.
2. Describes the consequence as not funding the program and/or continuing poor physical fitness of the adult residents in the city.

Question 5 (continued)

Notes

- If a response provides more than one description of a Type II error, score the weakest attempt.
- Referring to the symbolic hypotheses is *not sufficient* for context.
- Referring to funding and/or the city council is *not sufficient* for context.
- If a response describes a Type II error incorrectly, the response can get the consequence component correct if it is consistent with the incorrectly described error.
- If a response provides more than one description of a Type II error, the response can get the consequence component correct if the consequence is clearly linked to one of the error descriptions and is consistent with the error to which it is linked.
- If a response gives an incomplete description of a Type II error (for example, "we fail to reject the null hypothesis that the proportion of adult residents who are able to pass is 0.35"), the response can get the consequence component correct if the consequence is consistent with the partial description of the error.
- If a response provides no description of a Type II error, the response cannot get the consequence component correct.

Partially correct (P) if the response correctly completes only one of the two components listed above.

Incorrect (I) if the response correctly completes neither of the two components listed above.

Note: Describing the Type II error only in terms of the consequence (for example, "They don't fund the program when they should") should get credit for the consequence but should not get credit for the error, because there is no reference to the proportion of adult residents in the city who are able to pass the test.

Part (b) is scored as follows:

Essentially correct (E) if the response correctly completes the following three components:

1. Links the p-value to the conclusion by stating that the p-value is greater than $\alpha = 0.05$,
   OR
   by stating that the p-value is large,
   OR
   by correctly interpreting the p-value.
2. Uses context by referring to the proportion of adult residents who are able to pass the test,
   OR
   by referring to the funding of the program.
3. Makes a correct conclusion that describes the lack of evidence for the alternative hypothesis ($H_a : p < 0.35$).

Notes

- If a response includes an incorrect interpretation of the p-value, then the response cannot earn credit for the linkage component, even if the response explicitly compares the p-value to α or describes the p-value as large.
- Referring to the symbolic hypotheses is not sufficient for context.
- Accepting the null hypothesis or some equivalent statement such as "the population proportion is (or is likely to be, or is about) 0.35" cannot receive credit for the conclusion component, even if the student makes additional correct statements about the alternative hypothesis.

Question 5 (continued)

- Stating that the null hypothesis should not be rejected is not sufficient for the conclusion, because it does not address the direction of H_a.
- Correctly addressing the consequence ("They don't fund the program") is sufficient if the response also indicates that the null hypothesis is not being rejected.
- Drawing a conclusion about the sample proportion (for example, "proportion who passed the test") is not sufficient for the conclusion, because it does not properly address the parameter in H_a.

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

Incorrect (I) if the response correctly completes one or none of the three components listed above.

Notes

- A response that says the *p*-value is very large, recognizes that the sample proportion $\left(\hat{p} = 0.416\right)$ is greater than $p = 0.35$, and consequently concludes there is no evidence to support H_a (in context) is scored as essentially correct (E).
- A response that *rejects* H_0 is scored as incorrect (I).

Part (c) is scored as follows:

Essentially correct (E) if the response correctly completes the following three components:

1. States that the sample is not random and/or says that volunteers were used.
2. Describes *how* the sample is "different" with regard to physical fitness or another variable related to the ability to pass the physical fitness test.
3. Addresses the idea of making an inference from the sample to the population by stating that the sample statistic will overestimate the population parameter or that the sample will not be representative of the population.

Notes

- If for the first component a student provides additional proposed flaws (for example, "the sample size is too small"), score the weakest attempt.
- Saying only that the sample is different or not representative does not address *how* the sample is different.
- Saying "physically fit people will be overrepresented" or "the results cannot be generalized" or "the results will be inaccurate" lack a specific reference to the population and is not sufficient for the third component.
- Referring to "bias" is not sufficient for the first component unless the concept of bias is clearly explained (for example, saying that the sample proportion will tend to overestimate the population proportion).
- Incorrect application of statistical concepts (for example, saying that the statistic is "skewed," discussing cause and effect) results in a loss of credit for the third component.

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

Incorrect (I) if the response correctly addresses one or none of the three components listed above.

AP® STATISTICS
2012 SCORING GUIDELINES

Question 5 (continued)

4 Complete Response

All three parts essentially correct

3 Substantial Response

Two parts essentially correct and one part partially correct

2 Developing Response

Two parts essentially correct and one part incorrect

OR

One part essentially correct and one or two parts partially correct

OR

Three parts partially correct

1 Minimal Response

One part essentially correct and two parts incorrect

OR

Two parts partially correct and one part incorrect