

**AP<sup>®</sup> STATISTICS**  
**2010 SCORING GUIDELINES**

**Question 4**

**Intent of Question**

The primary goals of this question were to (1) assess students' ability to calculate an expected value and a standard deviation; (2) recognize the applicability of a binomial distribution and perform a relevant binomial probability calculation (or recognize the applicability of a normal approximation and use it to perform a relevant probability calculation); (3) suggest an appropriate sampling method to achieve a given goal.

**Solution**

**Part (a):**

Because the population size is so large compared with the sample size ( $\frac{297,354}{2,000} \approx 148.7$  times the sample size), far greater than the usual standard of 10 or 20 times larger, we can use the binomial probability distribution even though this is technically sampling without replacement. The parameters of this binomial distribution are the sample size,  $n$ , which has a value of  $n = 2,000$ , and the proportion of new car buyers who bought model E,  $p$ , which has a value of  $p = \frac{2,323}{297,354} \approx 0.0078$ . The expected value of the number of model E buyers in a simple random sample of 2,000 is therefore  $n \times p = 2,000 \times 0.0078 \approx 15.62$ . The variance is  $n \times p \times (1 - p) = 2,000 \times 0.0078 \times (1 - 0.0078) \approx 15.50$ , so the standard deviation is  $\sqrt{15.50} \approx 3.94$ .

**Part (b):**

For the reason given in part (a), the binomial distribution with  $n = 2,000$  and  $p \approx 0.0078$  can be used here. The probability that the sample would contain fewer than 12 owners of model E is calculated from the binomial distribution to be  $\sum_{x=0}^{11} \binom{2,000}{x} (0.0078)^x (0.9922)^{2,000-x} \approx 0.147$ . This probability is small enough that the result (fewer than 12 owners of model E in the sample) is not likely, but this probability is also not small enough to consider the result very unlikely.

This binomial probability can also be evaluated using a normal approximation. This is reasonable because  $n \times p = (2,000) \times (0.0078) = 15.6$  is larger than 10 and  $n(1 - p) = (2,000) \times (0.9922) = 1,984.4$  is much larger than 10. Using the mean and standard deviation from part (a) gives

$$P(X \leq 11) \approx P\left(Z < \frac{12.0 - 15.62}{3.94}\right) = P(Z < -0.92) = 0.179.$$

**Part (c):**

Stratified random sampling addresses the concern about the number of owners for models D and E. By stratifying on car model and then taking a simple random sample of at least 12 owners from the population of owners for each model, the company can ensure that at least 12 owners are included in the sample for each model while maintaining a total sample size of 2,000. For example, the company could select simple random samples of sizes 755, 647, 560, 22 and 16 for models A, B, C, D and E, respectively, to make the sample size approximately proportional to the size of the owner population for each model.

# AP<sup>®</sup> STATISTICS

## 2010 SCORING GUIDELINES

### Question 4 (continued)

#### **Scoring**

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

**Part (a)** is scored as follows:

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

- Calculation of the expected number of owners, showing a proper method for the calculation and providing the correct numerical value
- Calculation of the standard deviation for the number of owners, indicating recognition of the appropriate binomial distribution and providing the calculation and the correct numerical value

Partially correct (P) if the response contains only one of the two components listed above *OR* displays correct formulas for both the expected value and the standard deviation of a binomial distribution but fails to show both of the correct numerical values.

Incorrect (I) if the response provides only numerical values without showing how they were calculated.

**Part (b)** is scored as follows:

Essentially correct (E) if the student does any of the following:

- Recognizes the applicability of the binomial distribution, identifies the correct parameters, sets up the relevant probability calculation, and completes the calculation correctly
- Uses a normal probability approximation, identifying the relevant mean and standard deviation, and shows a correct calculation of the probability
- Provides an argument based on an appropriate z-score, or the number of standard deviations away from the mean, with a reasonable conclusion about likeliness

Partially correct (P) if the student does any of the following:

- Recognizes the applicability of the binomial distribution and identifies the correct parameters *BUT* sets up an incorrect cumulative binomial probability calculation
- Recognizes the applicability of the binomial distribution and shows the calculation correctly *BUT* does not identify the correct parameters in either part (a) or part (b)
- Recognizes the applicability of the normal approximation and identifies the correct parameters *BUT* incorrectly calculates the z-score or probability

Incorrect (I) otherwise.

#### *Notes*

- If the parameter values were properly identified in part (a), they do not have to be identified in part (b).
- If the response shows a correct calculation of the probability, no comment about likeliness is necessary. But such a comment is necessary if the response contains only a z-score without a probability or discusses standard deviations from the mean.
- With the normal calculation, it is acceptable for the response to show the probability that the normal value is below 11 or 11.5 or 12.

# AP<sup>®</sup> STATISTICS

## 2010 SCORING GUIDELINES

### Question 4 (continued)

**Part (c)** is scored as follows:

Essentially correct (E) if the response describes an appropriate sampling method (e.g., stratified random sampling) that ensures all of the following:

- Total sample size of 2,000
- At least 12 owners for each of the five car models
- Random selection of owners

Partially correct (P) if the response mentions stratified random sampling but gives a weak description, or no description, of how to implement the procedure *OR* describes another appropriate sampling method but includes only two of the three components listed above.

Incorrect (I) otherwise.

**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