Question 2

Intent of Question

The three primary goals of this question are to assess a student's ability to: (1) calculate a probability from a display of population frequencies; (2) calculate a binomial probability; and (3) describe a sampling distribution of a sample mean for a moderately large sample.

Solution

Part (a):

$$P(X > 3) = 0.07 + 0.04 + 0.04 + 0.02 = 0.17.$$

Part (b):

Y = number of households in violation.

Y has a binomial distribution with $n = 10$ and $p = 0.17$.

$$P(Y = 2) = \binom{10}{2}(0.17)^2(0.83)^8 = 0.2929.$$

Part (c):

The distribution of $\bar{X}$ will:

1. be approximately normal;
2. have mean $\mu_{\bar{X}} = \mu = 1.65$;
3. have standard deviation $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{1.851}{\sqrt{150}} = 0.1511$.

Scoring

This question is scored in four sections. Each section is scored as either essentially correct (E), partially correct (P), or incorrect (I).

Section 1 is part (a), section 2 is part (b), and sections 3 and 4 consist of elements of part (c). This scoring gives part (c) double weight relative to either part (a) or part (b).

Section 1 is essentially correct (E) if $P(X > 3)$ is correctly computed and work is shown in part (a).

Section 1 is partially correct (P) if:

$P(X \geq 3) = 0.26$ is computed;

OR

a correct numerical answer is given but no work is shown.

Question 2 (continued)

Section 2 is essentially correct (E) if in part (b):

1. the probability from part (a) is correctly used to calculate the probability that exactly 2 households are in violation, either using the binomial pdf or using general probability rules, *AND*
2. work is shown.

Section 2 is partially correct (P) if in part (b):

the student computes $P(Y \geq 2) = 0.5270$ or $P(Y \leq 2) = 0.7659$ instead of $P(Y = 2)$;

OR

the correct probability is given but no work is shown;

OR

the binomial coefficient is omitted $\left[(0.83)^8 (0.17)^2 = 0.0065\right]$.

Section 3 is essentially correct (E) if the response to part (c):

recognizes that the distribution of $\bar{X}$ will be approximately normal;

OR

the response says that the distribution of $\bar{X}$ is more symmetric than the population distribution *AND* mentions that the population distribution is highly skewed.

Section 3 is partially correct (P) if the response to part (c) reports a normal distribution for $\bar{X}$ without indicating that the normal distribution is an approximation.

Section 4 is essentially correct (E) if the response to part (c) provides the appropriate mean $\mu_{\bar{X}} = \mu = 1.65$ and standard deviation $\sigma_{\bar{X}} = \dfrac{\sigma}{\sqrt{n}} = \dfrac{1.851}{\sqrt{150}} = 0.1511$ for $\bar{X}$.

Section 4 is partially correct (P) if the response to part (c):

provides either the correct mean or the correct standard deviation for $\bar{X}$, but not both;

OR

provides correct numerical values for both the mean and standard deviation but sample notation ($\bar{X}$ and s) is used instead of population notation (μ and $\sigma_{\bar{X}} = \sigma / \sqrt{n}$);

OR

says only that the $\bar{X}$ distribution is centered in the same place as the population and has a smaller standard deviation than the population (and does not give the values of 1.65 and 0.1511).

Question 2 (continued)

4 Complete Response

All four sections essentially correct

3 Substantial Response

Three sections essentially correct and no section partially correct

OR

Two sections essentially correct and two sections partially correct

2 Developing Response

Two sections essentially correct and no sections partially correct

OR

One section essentially correct and two sections partially correct

OR

No sections essentially correct and four sections partially correct

1 Minimal Response

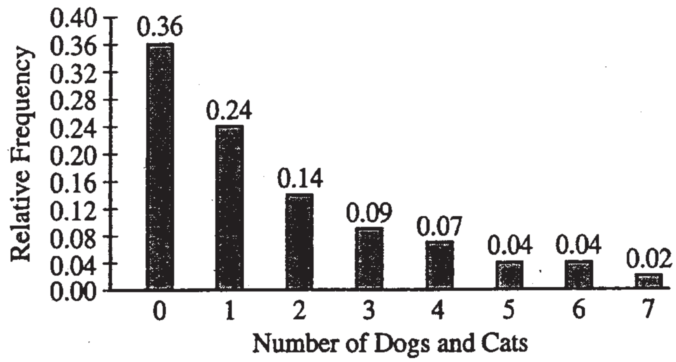
One section essentially correct and no sections partially correct

OR

No sections essentially correct and two sections partially correct

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. If the word "approximately" is missing in part (c), round down.

2. The graph below displays the relative frequency distribution for X, the total number of dogs and cats owned per household, for the households in a large suburban area. For instance, 14 percent of the households own 2 of these pets.



(a) According to a local law, each household in this area is prohibited from owning more than 3 of these pets. If a household in this area is selected at random, what is the probability that the selected household will be in violation of this law? Show your work.

$P(X>3) = 0.07 + 0.04 + 0.04 + 0.02$

$P(X>3) = \boxed{0.17}$

(b) If 10 households in this area are selected at random, what is the probability that exactly 2 of them will be in violation of this law? Show your work.

$P(X>3)$ for 10 households $= \binom{10}{2}(0.17)^2(0.83)^8$

$\boxed{= 0.2929}$

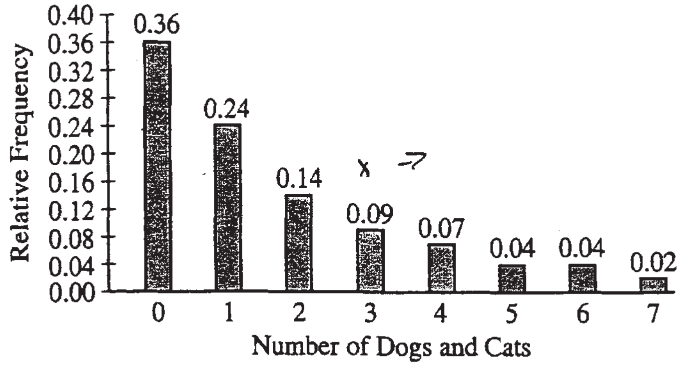
$P(X>3) = \binom{n}{3}(p)^3(1-p)^{n-3}$

(c) The mean and standard deviation of X are 1.65 and 1.851, respectively. Suppose 150 households in this area are to be selected at random and $\overline{X}$, the mean number of dogs and cats per household, is to be computed. Describe the sampling distribution of $\overline{X}$, including its shape, center, and spread.

The shape of the sampling distribution of $\overline{x}$ will be approximately normal (per the Central Limit Theorm), with the center located at 1.65 and the spread equaling $\frac{1.851}{\sqrt{150}}$, or about 0.1511. $\boxed{N(1.65, 0.1511)}$

GO ON TO THE NEXT PAGE.

2. The graph below displays the relative frequency distribution for X, the total number of dogs and cats owned per household, for the households in a large suburban area. For instance, 14 percent of the households own 2 of these pets.



(a) According to a local law, each household in this area is prohibited from owning more than 3 of these pets. If a household in this area is selected at random, what is the probability that the selected household will be in violation of this law? Show your work.

$.09 + .07 + .04 + .04 + .02$

$= .26$

26% of selected households will be in violation of the law.

(b) If 10 households in this area are selected at random, what is the probability that exactly 2 of them will be in violation of this law? Show your work.

violation = .26

$_{10}C_{8}\ (.26)^2(.74)^8$

$= .2735$

27.4%

(c) The mean and standard deviation of X are 1.65 and 1.851, respectively. Suppose 150 households in this area are to be selected at random and $\overline{X}$, the mean number of dogs and cats per household, is to be computed. Describe the sampling distribution of $\overline{X}$, including its shape, center, and spread.

N(1.65, 1.851)

n=150

$\sigma_{\bar{x}} = \dfrac{1.851}{\sqrt{150}}$

$\bar{x}$ = ? 1.65

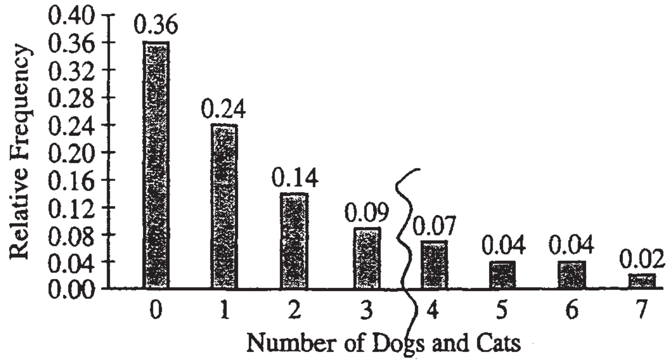
$\mu_{\bar{x}} = P$

$\sigma_{\bar{x}} = .151$

N(1.65, .151) ~ Normal spread Distribution with a center at approx. 1.65.

GO ON TO THE NEXT PAGE.

2. The graph below displays the relative frequency distribution for X, the total number of dogs and cats owned per household, for the households in a large suburban area. For instance, 14 percent of the households own 2 of these pets.



(a) According to a local law, each household in this area is prohibited from owning more than 3 of these pets. If a household in this area is selected at random, what is the probability that the selected household will be in violation of this law? Show your work.

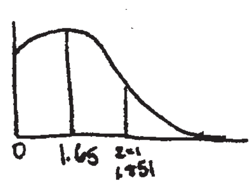
$P(>3 \text{ pets}) = .07 + .04 + .04 + .02 = .17$

(b) If 10 households in this area are selected at random, what is the probability that exactly 2 of them will be in violation of this law? Show your work.

$P(2 \text{ have} > 3) = (.17)^2 + (.83)^8 = .254$

(c) The mean and standard deviation of X are 1.65 and 1.851, respectively. Suppose 150 households in this area are to be selected at random and $\overline{X}$, the mean number of dogs and cats per household, is to be computed. Describe the sampling distribution of $\overline{X}$, including its shape, center, and spread.

The center would be about 1.65. It would be normaly distributed, but would stop abruptly at 0 on the left (can't have negative pets.)



GO ON TO THE NEXT PAGE.

Question 2

Sample: 2A
Score: 4

This is a complete response in which the probability that a randomly selected household will be in violation of the law is correctly calculated by summing the population frequencies of households owning 4, 5, 6, or 7 dogs and cats as $0.07 + 0.04 + 0.04 + 0.02 = 0.17$ in part (a). In part (b) the probability computed in part (a) is used to compute the appropriate binomial probability that exactly 2 of 10 randomly selected households will be in violation of the law. In part (c) the sampling distribution of the sample mean is correctly described as approximately normal with mean 1.65 and standard deviation 0.1511, although the less precise word “spread” is used instead of standard deviation. The response shows additional understanding by referring to the Central Limit Theorem to justify the normal approximation to the distribution of the sample mean, although this was not required for a complete response.

Sample: 2B
Score: 3

This is a substantial response in which the probability that a randomly selected house will be in violation of the law is not correctly calculated. This response incorrectly includes the population frequency of households owning exactly 3 cats or dogs to obtain an incorrect probability of 0.26 in part (a). The response to part (b) was scored as correct because it displays an appropriate binomial probability formula that is correctly evaluated with the probability computed in part (a). The correct mean and standard deviation of the large sample normal approximation to the distribution of the sample mean are reported in part (c), but this response fails to recognize that the normal distribution is an approximation.

Sample: 2C
Score: 2

This is a developing response because it demonstrates some understanding of probability concepts in part (a) and partial understanding of the sampling distribution of a sample mean in part (c). The probability that a randomly selected household will be in violation of the law on pet ownership is correctly calculated in part (a). Part (b) is incorrect. It reports the sum of two probabilities instead of the product needed to compute the appropriate binomial probability. The mean of the distribution of the population of possible sample means is correctly reported in part (c), but the standard deviation is not reported. By referring to the truncation issue, arising from the fact that the sample mean cannot be negative, this response shows some understanding that the normal distribution is only an approximation to the true distribution of the sample mean, but the display provided in this response reveals no understanding that the standard deviation of the sample mean will be much smaller than the standard deviation for a single observation.