

2005 AP[®] STATISTICS FREE-RESPONSE QUESTIONS

2. Let the random variable X represent the number of telephone lines in use by the technical support center of a software manufacturer at noon each day. The probability distribution of X is shown in the table below.

x	0	1	2	3	4	5
$p(x)$	0.35	0.20	0.15	0.15	0.10	0.05

- (a) Calculate the expected value (the mean) of X .
- (b) Using past records, the staff at the technical support center randomly selected 20 days and found that an average of 1.25 telephone lines were in use at noon on those days. The staff proposes to select another random sample of 1,000 days and compute the average number of telephone lines that were in use at noon on those days. How do you expect the average from this new sample to compare to that of the first sample? Justify your response.
- (c) The median of a random variable is defined as any value x such that $P(X \leq x) \geq 0.5$ and $P(X \geq x) \geq 0.5$. For the probability distribution shown in the table above, determine the median of X .
- (d) In a sentence or two, comment on the relationship between the mean and the median relative to the shape of this distribution.

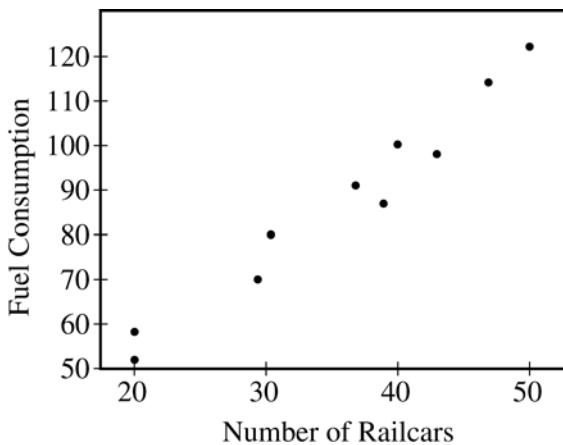
2005 AP[®] STATISTICS FREE-RESPONSE QUESTIONS

3. The Great Plains Railroad is interested in studying how fuel consumption is related to the number of railcars for its trains on a certain route between Oklahoma City and Omaha.

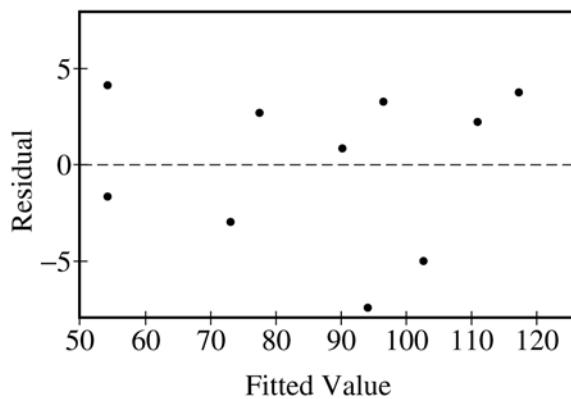
A random sample of 10 trains on this route has yielded the data in the table below.

Number of Railcars	Fuel Consumption (units/mile)
20	58
20	52
37	91
31	80
47	114
43	98
39	87
50	122
40	100
29	70

A scatterplot, a residual plot, and the output from the regression analysis for these data are shown below.



RESIDUALS VERSUS THE FITTED VALUES



**AP[®] STATISTICS
2005 SCORING GUIDELINES**

Question 2

Solution

Part (a):

The expected number of telephone lines in use by the technical support center at noon is:

$$\begin{aligned}E(X) &= 0 \times 0.35 + 1 \times 0.2 + 2 \times 0.15 + 3 \times 0.15 + 4 \times 0.1 + 5 \times 0.05 \\&= 1.6\end{aligned}$$

Part (b):

We would expect the average based on 1,000 days to be closer to 1.6 than the first average based on 20 days. Both averages have the same expected value (1.6), but the variability for sample averages based on 1,000 days is smaller than the variability for sample averages based on 20 days.

Part (c):

The median of X is 1.

x	$P(X \leq x)$	$P(X \geq x)$
0	0.35	1.0
1	0.55	0.65
2	0.70	0.45
3	0.85	0.30
4	0.95	0.15
5	1.0	0.05

OR

The median of X is 1 because $P(X \leq 1) = 0.55 \geq 0.50$ and $P(X \geq 1) = 0.65 \geq 0.50$.

Part (d):

The probability histogram is clearly skewed to the right (or toward the larger values) so the mean (1.6) is larger than the median (1), as is typical for a right-skewed distribution.

**AP[®] STATISTICS
2005 SCORING GUIDELINES**

Question 2 (continued)

Scoring

Parts (a) and (c) are combined as one computational part. Each part is scored as essentially correct (E), partially correct (P), or incorrect (I).

Collectively parts (a) and (c) are essentially correct (E) if both parts are calculated correctly, with the exception of minor arithmetic errors.

Collectively parts (a) and (c) are partially correct (P) if one of the two parts is calculated correctly, with the exception of minor arithmetic errors.

Collectively parts (a) and (c) are incorrect (I) if both parts are calculated incorrectly.

Note: Unsupported answers in parts (a) and (c) are scored as incorrect.

Part (b) is essentially correct (E) if the student:

1. States the new estimate based on 1,000 days should be closer to the expected value of 1.6; OR the new estimate will increase, or decrease if the answer in part (a) is less than 1.25.

AND

2. Provides justification by stating the variability for sample averages based on 1,000 days will be smaller than the variability for sample averages based on 20 days; OR as the sample size increases the sample average approaches the expected value of X .

Part (b) is partially correct (P) if the student provides one of the two items above.

Part (d) is essentially correct (E) if the student states that since the distribution is skewed to the right, the mean is greater than the median; OR since the mean is greater than the median, the distribution is skewed to the right.

Note: There must be evidence that the student looked at the given distribution.

Part (d) is partially correct (P) if the student:

- States that since the mean is greater than the median, the distribution is skewed to the right (with no evidence that the student looked at the given distribution); OR
- Compares the two measures of center by referring to the inappropriate or incomplete shape of the distribution (e.g., “skewed to the left” or “skewed”); OR
- Makes a correct statement about the measures of center and the shape without connecting the two.