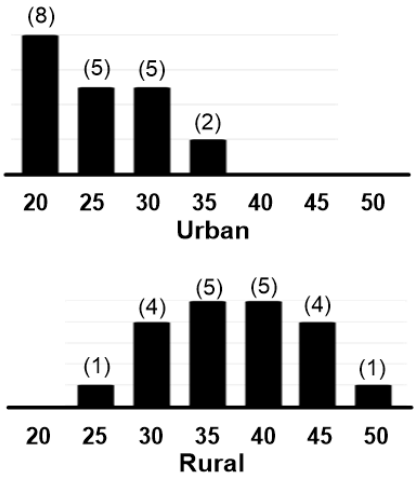
**AP Statistics Semester 1 Free-Response Final REVIEW**

#1. The goal of a nutritional study was to compare the caloric intake of adolescents living in rural areas of the United States with the caloric intake of adolescents living in urban areas of the United States. A random sample of ninth-grade students from one high school in a rural area was selected. Another random sample of ninth graders from one high school in an urban area was also selected. Each student in each sample kept records of all the food he or she consumed in one day.



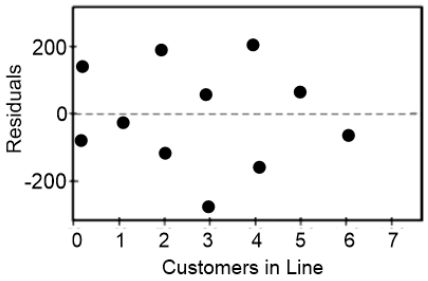

The bar charts for each region are shown to the right displaying the number of calories of food consumed per kilogram of body weight for each student on that day (the number of students in each column is shown in parentheses above the column)

(a) Write a few sentences comparing the distribution of the daily caloric intake of ninth-grade students in the rural high school with the distribution of the daily caloric intake of ninth-grade students in the urban high school.

(b) Is it reasonable to generalize the findings of this study to all rural and urban ninth-grade students in the United States? Explain.

(c) What percentage of all the students studied consumed 35 calories per kilogram of body weight?

#2. The manager of a grocery store selected a random sample of 11 customers to investigate the relationship between the number of customers in a checkout line and the time to finish checkout. As soon as the selected customer entered the end of a checkout line, data were collected on the number of customers in line who were in front of the selected customer and the time, in seconds, until the selected customer was finished with the checkout. A linear regression was performed on the data and the regression output and residual plot are shown below:




| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 72.95 | 110.36 | 0.66 | 0.525 |
| Customers in line | 174.40 | 35.06 | 4.97 | 0.001 |
| S = 200.01 | | R-Sq = 73.33% | R-Sq (adj) = 70.37% | |

(a) Is a linear model appropriate for modeling these data? Clearly explain your reasoning.

(b) Write the LSRL for the association between customers in line and time to finish checkout (be sure to define your variables)

(c) What is the predicted time to check out if there are 4 customers in line?

(d) Interpret the value of the slope of the LSRL in the context of this problem.

(e) Interpret the value of the y-intercept of the LSRL in the context of this problem.

(f) Interpret the value of $r^2$ in the context of this problem.

#3. A coed youth sports league includes multiple teams each with roughly the same number of players. For advertising purposes, the organization which manages the league wants to produce a brochure which contains a picture taken by a professional photographer of a 'typical' team – which represents the youth league's participants well. The photograph is tasked with selecting which players should appear in the photograph. She is considering two sampling strategies: a cluster sample, and a stratified random sample.

(a) Explain the difference between cluster sampling and stratified random sampling in this context.

(b) If cluster sampling is used, and the sample turns out to be biased, explain why this would <u>not</u> be called voluntary response bias.

(c) If cluster sampling is used, and the sample turns out to be biased, what would the correct term be for the type of bias which may occur?

#4. Airlines routinely overbook flights because they expect a certain number of no-shows. An airline runs a 5 P.M. commuter flight from Washington, D.C., to New York City on a plane that holds 38 passengers. Past experience has shown that if 41 tickets are sold for the flight, then the probability distribution for the number who actually show up for the flight is as shown in the table below.

| Number who actually show up | 36 | 37 | 38 | 39 | 40 | 41 |
|---|---|---|---|---|---|---|
| Probability | 0.46 | 0.30 | 0.16 | 0.05 | 0.02 | 0.01 |

Assume that 41 tickets are sold for each flight.

(a) There are 38 passenger seats on the flight. What is the probability that all passengers who show up for this flight will get a seat?

(b) What is the expected number of no-shows for this flight?

(c) Given that not all passenger seats are filled on a flight, what is the probability that only 36 passengers showed up for the flight?