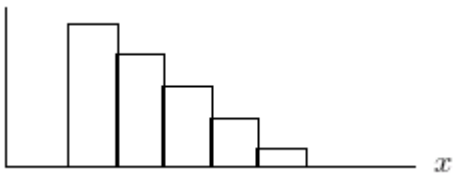
Math 263, Section 4, Solutions to Exam 1, Spring 2005

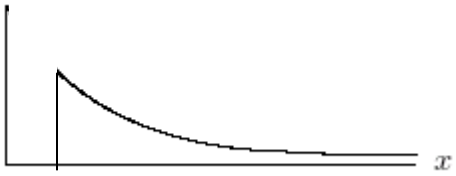
1.

(a) Frequency or relative frequency of incomes. Units are numbers (integers, percents, or fractions).

(b) Incomes. Units are dollars.

(c) Distribution has tail to the right, as in either of the following diagrams. Histogram does not start at 0.





2.

(a) Observational study, because experimenters did not intervene. Students were not randomly assigned to learn foreign languages.
(b) Whether or not have studied foreign language.
(c) Categorical.
(d) Score on verbal SAT.
(e) Quantitative.
(f) No, we cannot conclude from these studies that studying a foreign language causes students' SAT verbal scores to improve. We know there is an association between studying a foreign language and higher verbal SAT score, but there may be lurking variables. For example, students who are more academically inclined may be the ones who chose to study foreign languages and they also tend get higher verbal SAT scores. However, these experiments do not rule out the possibility that foreign language does cause improvement students' verbal skill and hence raise verbal SAT scores. Common sense suggests that studying a new language may help students understand their own.

3.

(a) $\bar{x} = \dfrac{\sum_{i=1}^{i=3} x_i}{3}$

(b) $s = \sqrt{\dfrac{1}{2}\left((x_1 - 17)^2 + (x_2 - 17)^2 + (x_3 - 17)^2\right)}$

4.

I. is 30
II. is 20
III. is 40
IV. is 50
V. is 10

5.

(a) The fraction of the variation explained by the regression line is r^2, so the fraction not explained by the line is $1 - r^2$.
(b) Any value near –1, say – 0.9.
(c) Millions of people.
(d) The number of people living on farms when years = 0, that is in AD 0.
(e) Millions of people per year.
(f) The average decrease in the number of people living on farms, per year.
(g) Using the line, we predict the population in 2004 to be $y = 1167 - 0.587*2004 = -9.35$. Since population cannot be negative, this prediction must be incorrect. The reason is that it is dangerous to extrapolate outside the range of the data given.

(h) Since $(\bar{x}, \bar{y})$ lies on the regression line, we need to find out if (1957.5, 21.5) lies on the line. If x = 1957.5, the equation gives y = 1167 – 0.587*1957.5 = 17.95.Thus (1957.5, 17.95) is on the line, but (1957.5, 21.5) is not. Thus (1957.5, 21.5) could not be $(\bar{x}, \bar{y})$.

6.
(a) The shortest two groups, sopranos and altos, are women.
(b) 70 inches.
(c) Tenors. IQR = Q3 – Q1 = 71 – 67 = 4 inches.
(d) Sopranos; Q2 = Q3 = 65 inches.
(e) Smaller.
(f) For tenors, 1.5*IQR = 1.5*4 = 6 inches. Thus, a height above 71 + 6 = 77 inches is an outlier.