Univariate Data Analysis Graded Assignment Solutions
Applied Finite Math

Chapter 13

68. Answers will vary.

70. Consideration of one exception made them less likely to consider a second exception. This is why it is critical that the order a series of survey questions such as this are asked is randomized.

72. a. An Area Code 900 telephone poll represents an extreme case of selection bias. People that respond to these polls usually represent the extreme view points (strongly for or strongly against), leaving out much of the middle of the road point of view. Economics also plays some role in the selection bias. (While 50 cents is not a lot of money anymore, poor people are much more likely to think twice before spending the money to express their opinion.)

b. This survey was based on fairly standard modern-day polling techniques (random sample telephone interviews, etc.) but it had one subtle flaw. How reliable can a survey about the conduct of the newsmedia be when the survey itself is conducted by a newsmedia organization?

c. Both surveys seem to have produced unreliable data – survey 1 overestimating the public's disapproval of the role played by the newsmedia and survey 2 overestimating the public's support for the press coverage of the war. In survey 1, only those with strong opinions were surveyed (i.e. called the 900 number). In survey 2, people without strong opinions were called. Further, they may have been swayed by being surveyed by a newsmedia organization.

d. Many reasonable explanations possible.

Chapter 14

80. a. The smallest possible standard deviation for such a data set occurs when as many values in the data set as possible lie close to or at the mean. In this particular case, it occurs when 8 of the numbers take on the value of the average (A=7), one value takes on the minimum (Min = 2), and one value takes on the maximum (Max = 12).

The average of the data set {2, 7, 7, 7, 7, 7, 7, 7, 7, 12} is A=7

x	$Freq. f$	$x-7$	$f \cdot (x-7)^2$
2	1	-5	25
7	8	0	0
12	1	5	25
			50

So, the smallest value of $\sigma = \sqrt{\frac{50}{10}} = \sqrt{5}$.

b. The largest possible standard deviation for such a data set occurs when as many values in the data set as possible lie as far away as possible from the mean. In this particular case, in order to have a mean of 7, this occurs when 5 of the numbers take on the value of the minimum (Min = 2) and 5 values take on the maximum (Max = 12).
The average of the data set {2, 2, 2, 2, 2, 12, 12, 12, 12, 12} is A = 7

x	$Freq. f$	$x-7$	$f \cdot (x-7)^2$
2	5	-5	125
12	5	5	125
			250

So, the largest value of $\sigma = \sqrt{\frac{250}{10}} = 5$.

82. a.This follows since

$$\sigma = \sqrt{\frac{(x_1 - A)^2 + (x_2 - A)^2 + .. + (n - A)^2}{N}}$$

$$\geq \sqrt{\frac{(x_i - A)^2}{N}}$$

$$= \frac{|x_i - A|}{\sqrt{N}}$$

For every data value

b. From (a), $-\sigma\sqrt{N} \leq x_i - A \leq \sigma\sqrt{N}$ or equivalently, $A - \sigma\sqrt{N} \leq x_i \leq A + \sigma\sqrt{N}$

84. (Bonus) $1 + 2 + \ldots + N = \frac{N(N=1)}{2}$ implies the average $A = \frac{1+2+\ldots+N}{N} = \frac{N+1}{2}$. If N is odd, the median M is the “middle” number $M = \frac{N+1}{2}$. If N was even the median M is the average of $\frac{N}{2}$ and $\frac{N}{2} + 1$ which is $M = \frac{N+1}{2}$.

86. $V = \frac{(x_1-A)^2+\cdots+(x_N-A)^2}{N} = \frac{(x_1{}^2-2Ax_1+A)^2+\cdots+(x_N{}^2-2Ax_N+A)^2}{N} = \frac{(x_1{}^2+\cdots+x_N{}^2)}{N} - 2A\frac{(x_1+\cdots+x_N)}{N} + \frac{NA^2}{N} =$ $\mathrm{B} - 2A^2 + A^2 = B - A^2$

90. a. $\frac{8}{9}$ or approximately 89%

b. 10

c. Since $k > 0, 1 - \frac{1}{k^2} < 1$ will always be the case.

Chapter 16

68. a. $55 + 12 \times 2.33 = 82.96 \approx 83\ points$
b. $55 - 12 \times 0.52 = 48.76 \approx 49\ points$

70. a. 99th percentile – $55 + 12 \cdot w = 83$ - $w \approx 2.33$

b. 20th Percentile – $55 - 12 \cdot w = 45$ - $w \approx 0.84$

c. 5000 – $55 - 12 \cdot w = 35$ – $w = 1.666$

So a score of 35 is at (approximately) the 5th Percentile. 5% of $N = 250$ gives $N = 5000$ students.

78. $n = 720$

The probability of rolling a total of 7 is $p = \frac{6}{36} = \frac{1}{6}$. Thus, Y has an approximately normal distribution with mean $\mu = \frac{n}{6}$ and standard deviation $\sigma = \sqrt{n \cdot 1/6 \cdot 5/6} = \frac{1}{6}\sqrt{5n}$. If there is a 95% chance that Y will be between $\frac{n}{6} - 20$ and $\frac{n}{6} + 20$ then $20 = 2\sigma$ and $\sigma = 10$. Solving $\frac{1}{6}\sqrt{5n} = 10$ gives $n = 720$.

80. The standard deviation for the number of votes for Mrs. Butterworth is $\sqrt{800 \times 0.53 \times 0.47} \approx 14.12$, the standard error is approximately $\frac{14.12}{800} = 0.0176$ or 1.76%