STAT230 - McClave *et al.*, - Answers - Chapters 2 & 3

Note that these are bare answers. What you present as homework must have problem statements or otherwise be self-contained and must have explanations.

There is a reward for finding misprints in these solutions.

2.54

a. The sample mean is 80/20 = 4. There are 20 samples, so the sample median is found by arranging them in ascending order, then taking the mean of the 10th and 11th samples. That's $\frac{3+4}{2} = 3.5$. The mode is the observation that appears the most. That's 1.

b. We do the same set of steps as in a., except we eliminate the largest number, 13. That results in $\bar{x} = 3.53$, median 3, mode still 1.

c. Here we drop the highest and lowest two observations, because 10% of 20 is 2, so there are 16, that is, 1 . . . 7. The sample 10% trimmed mean is 3.5.

2.68

a.

Worker A:

The average for subtask 1 is $\bar{x} = \frac{\sum x}{n} = \frac{211}{7} = 30.14$

For subtask 2 it's $\bar{x} = \frac{\sum x}{n} = \frac{21}{7} = 3$

The overall time is just the times for 1 and 2 added. That's 33.14.

Worker B:

The average for subtask 1 is $\bar{x} = \frac{\sum x}{n} = \frac{213}{7} = 30.43$

For subtask 2 it's $\bar{x} = \frac{\sum x}{n} = \frac{29}{7} = 4.14$

The overall time is just the times for 1 and 2 added. That's 34.57.

b.

Worker A:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{6455 - \frac{211^2}{7}}{7-1}} = \sqrt{15.8095} = 3.98$$

Note here the number of decimal places carried. For calculations, carry all that your calculaor will hold. The final answer, though, must make sense and must be in concordance with data. Here we give 0.01 second accuracy, about as much as can be measured for motion of human beings.

Worker B:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{6487 - \frac{213^2}{7}}{7-1}} = \sqrt{0.9524} = 0.98$$

d.

Worker A:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{67 - \frac{21^2}{7}}{7-1}} = \sqrt{0.6667} = 0.82$$

Worker B:

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{147 - \frac{29^2}{7}}{7-1}} = \sqrt{4.4762} = 2.12$$

e. Worker B for Subtask 1 because of consistency. Worker A for because of smaller average time and smaller variability.

2.96

a. $\mu = 2.7, \sigma = 0.5$

$$z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + z\sigma$$

Thus

For $z = 2.0, x = 2.7 + 2.0(0.5) = 3.7$;

For $z = -1.0, x = 2.7 - 1.0(0.5) = 2.2$;

For $z = 0.5, x = 2.7 + 0.5(0.5) = 2.95$;

For $z = -2.5, x = 2.7 - 2.5(0.5) = 1.45$.

b. For $z = -1.6, x = 2.7 - 1.6(0.5) = 1.9$.

c. Assume the GPA distribution is mound-shaped, and use the empirical rule.

We know that $\approx 0.25 = 2.5\%$ of the students will have GPA above 3.7 (with $z = 2$), thus *summa* is > 3.7.

Likewise $\approx 0.16 = 16\%$ of the students will have GPA above 3.2 (with $z = 1$). Thus *cum* is > 3.2.

2.98 a. The data are approximately mound-shaped, so we can use the Empirical Rule. For the blue exam, the mean is 53% and the SD is 15%. Thus approximately

68% of students are within $\bar{x} \pm s = 53 \pm 15 \Rightarrow (38, 68)$

95% of students are within $\bar{x} \pm s = 53 \pm 2(15) \Rightarrow (23, 83)$

99.7% of students are within $\bar{x} \pm s = 53 \pm 3(15) \Rightarrow (8, 98)$

b. By the same logic, the intervals are $(27, 51), (15, 63), (3, 75)$.

c. Red exam.

2.104

a.

$\bar{x} = 52.33$ and $s = 9.22$.

The highest salary is 75 (thousand). The z-score is $z = \frac{x-\bar{x}}{s} = \frac{75-53.22}{9.22} = 2.46$. Therefore, the highest salary is 2.46 standard deviations above the mean.

The lowest salary is 35 (thousand). The z-score is $z = \frac{x-\bar{x}}{s} = \frac{35-52.33}{9.22} = -1.88$. Therefore, the lowest salary is 1.88 standard deviations below the mean.

The z-score for the mean salary is zero, 0. (Obviously — verify that you understand that.)

The highest salary is not unusually high. For *any* distribution, at least 8/9 of the salaries should have z- scores between -3 and 3. Thus $z = 2.46$ isn't very unusual.

b. From the boxplot, which you should do, no salaries are outside the fences, so none are potentially faulty. (Doing a five-number summary is a good exercise, also.)

3.86

a. The two probability rules for a sample space are that the probability for any sample point is between 0 and 1 and that the sum of all probabilities is 1.

Here, the probabilities are between 0 and 1 and

$$\sum_{n=1}^{4} P(s_i) = 0.2 + 0.1 + 0.3 + 0.4 = 1.0.$$

b. $P(A) = P(S_1) + P(S_4) = 0.2 + 0.4 = 0.6.$

3.120 There are a total of $6 \times 6 = 36$ outcomes when rolling 2 dice. If we let the first number in the pair represent the outcome of die number 1 and the second number in the pair represent the outcome of die number 2, then the possible outcomes are:

1,1	2,1	3,1	4,1	5,1	6,1
1,2	2,2	3,2	4,2	5,2	6,2
1,3	2,3	3,3	4,3	5,3	6,3
1,4	2,4	3,4	4,4	5,4	6,4
1,5	2,5	3,5	4,5	5,5	6,5
1,6	2,6	3,6	4,6	5,6	6,6

If both dice are fair, then each of these outcomes are equally like and have a probability of 1/36.

a. To win on the first roll, a player must roll a 7 or 11. There are 6 ways to roll a 7 and 2 ways to roll an 11. Thus the probability of winning on the first roll is:

$$P(7 \text{ or } 11) = \frac{8}{36} = 0.2222$$

b. To lose on the first roll, a player must roll a 2 or 3. There is I way to roll a 2 and 2 ways to roll a 3. Thus the probability of losing on the first roll is:

$$P(2 \text{ or } 3) = \frac{3}{36} = .0833$$

c. If a player rolls a 4 on the first roll, the game will end on the next roll if the player rolls the original roll again (player wins) or if the player rolls a seven (player loses). Now, there are 3 ways of getting a 4 on the first roll: 1,3, 2,2, or 3,1.

If the first roll was 2,2, then the game would end on the next roll if the player threw a 2,2, 1,6, 2,5, 3,4, 4,3, 5,2, or 6,1 on the next roll. The probability of the game ending on the next roll would be:

$$P(2,2 \text{ or } 7 \text{ on second toss } |12,2 \text{ on first}) = \frac{7}{36} = 0.1944$$

Now, suppose the first roll ended with a 1 and a 3. Since the dice are not marked, this result could have happened two ways: 1,3 or 3,1. Regardless of how the original 1 and 3 were obtained, the player would have 2 ways of winning on the next roll: 1,3 or 3,1. For the game to end on the next roll, the player could throw 1,3, 3,1, 1,6, 2,5, 3,4, 4,3, 5,2, or 6,1. The probability of the game ending on the next roll would be:

$$P(1,3 \text{ or } 3,1 \text{ or } 7 \text{ on second toss } |1 \text{ and } 3 \text{ on firsr}) = \frac{8}{36} = 0.2222$$

Since there were 3 ways to get a 4 on the first roll, and each were equally likely, P(2,2) = 1/3 and P[I and 3 (any order)] = 2/3.

The probability that the game ends on the second roll is

$$P(2,2 \text{ or} 7 \;\text{ on second toss } P(2,2) \text{ on first})|P(2,2 \text{ on first})$$
$$+P(l,3 \text{ or} 3,1 \text{ or } 7 \text{ on second toss}|1 \text{ and } 3 \text{ on first})P(1 \text{ and } 3 \text{ on first} =$$
$$0.1944\left(\frac{1}{3}\right)0.222\left(\frac{2}{3}\right) = 0.0648 + 0.1481 = 0.2129$$

.