

Name: Caleb McWhorter — Solutions

MATH 307

Spring 2023

HW 8: Due 03/01 (02)

“Statistically, the probability of any one of us being here is so small that you’d think the mere fact of existing would keep us all in a contented dazzlement of surprise.”

–Lewis Thomas

Problem 1. (10pt) The number of hours Elizabeth has spent on TokTik over the past few days is given below:

0 1 6 1 3 6 2 6 9 7 8

- (a) Find the median of this dataset.
- (b) Find the IQR of this dataset.
- (c) A 5-number summary for a dataset consists of the min, Q_1 , median, Q_3 , and max for the data. Find the 5-number summary for this dataset.

Solution.

- (a) Putting the data in order, we have...

0 1 1 2 3 6 6 6 7 8 9

There are 11 numbers. Because $\frac{11}{2} = 5.5$, the median is the 6th number. But then the median is 6.

- (b) The numbers less than the sixth number 6 that is the median are 0, 1, 1, 2, 3. The median of these numbers is 1. Therefore, we know that $Q_1 = 1$. The numbers greater than the sixth number 6 that is the median are 6, 6, 7, 8, 9. The median of these numbers is 7. Therefore, we know that $Q_3 = 7$. But then we have...

$$\text{IQR} = Q_3 - Q_1 = 7 - 1 = 6$$

- (c) We know for this dataset that $\text{min} = 0$ and $\text{max} = 8$. We found the median in (a), which was 5.5. Finally, we computed Q_1 and Q_3 in (b) as 1 and 7, respectively. Therefore, the 5-number summary is...

0 1 5.5 7 9

Problem 2. (10pt) Will took the SAT and received a 1650 while Chris took the ACT and received a 23. Suppose that both the SAT and the ACT had scores which were normally distributed. Furthermore, suppose that the SAT had a mean score of 1500 and standard deviation 300 while the ACT had a mean score of 21 and standard deviation of 5. Relative to their own exams, who did better? Be sure to justify your answer.

Solution. Because the distribution of exam scores for both the SAT and the ACT were normally distributed, we can compare both Will's and Chris' exam scores relative to the other test takers on their exams using z -scores. But this will give a comparison between their scores, despite not having taken the same exam. We have...

$$z_{\text{Will}} = \frac{x - \mu}{\sigma} = \frac{1650 - 1500}{300} = \frac{150}{300} \approx 0.50$$

$$z_{\text{Chris}} = \frac{x - \mu}{\sigma} = \frac{23 - 21}{5} = \frac{2}{5} \approx 0.40$$

Because $0.50 = |z_{\text{Will}}| > |z_{\text{Chris}}| = 0.40$, Will did better on his own exam (relative to other test takers) than Chris did on his (relative to other test takers).

Problem 3. (10pt) STACKS is a local college. At the school, GPAs are approximately normally distributed with mean 3.205 and standard deviation 0.27.

- (a) Find the percentage of students that have a GPA lower than 2.800.
- (b) Find the percentage of students that have a GPA greater than 3.500.
- (c) Find the percentage of students that have a GPA between 2.800 and 3.500.

Solution.

- (a) We have...

$$z_{2.800} = \frac{2.800 - 3.205}{0.27} = \frac{-0.405}{0.27} = -1.50 \rightsquigarrow 0.0668$$

Then we know that $P(X < 2.800) = 0.0668$. Therefore, 6.68% of students have a GPA lower than 2.800.

- (b) We have...

$$z_{3.500} = \frac{3.500 - 3.205}{0.27} = \frac{0.295}{0.27} = 1.09 \rightsquigarrow 0.8621$$

Then we know that $P(X < 3.500) = 0.8621$. But then we know that $P(X > 3.500) = 1 - P(X < 3.500) = 1 - 0.8621 = 0.1379$. Therefore, 13.79% of students have a GPA greater than 3.500.

- (c) We have...

$$P(2.800 < X < 3.500) = P(X < 3.500) - P(X < 2.800) = 0.8621 - 0.0668 = 0.7953$$

Therefore, 79.53% of students have a GPA between 2.800 and 3.500.

Problem 4. (10pt) Margarita claims that you can generate a random sample of the numbers 2 through 12 by continuously rolling two die and taking the sum of the numbers that appear.¹ Explain why Margarita is incorrect. Be sure to include the concept of bias in your response. How would you help Margarita understand why she is wrong?

Solution. We can create a table of each of the possible sums of the numbers on the two die, as shown below.

Roll	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

We can see the sums of the dice repeating along the diagonals. There are 36 possible combinations of what appears on the dice. However, there are fewer possibilities for their sum. From this table, we can see that...

$$\begin{array}{llll}
 P(2) = \frac{1}{36} & P(5) = \frac{4}{36} & P(8) = \frac{5}{36} & P(11) = \frac{2}{36} \\
 P(3) = \frac{2}{36} & P(6) = \frac{5}{36} & P(9) = \frac{4}{36} & P(12) = \frac{1}{36} \\
 P(4) = \frac{3}{36} & P(7) = \frac{6}{36} & P(10) = \frac{3}{36} &
 \end{array}$$

Therefore, not all of these combinations are equally likely. This shows that collecting a sample of numbers from 2 to 12 in this manner would result in bias because certain outcomes would be more likely (and hence likely appear more often) than others.

¹By a random sample of a finite set of numbers, we mean that the probability of all elements of the sample space are equal.