MAT 107: Exam 3
Winter – 2022
01/21/2023
Time Limit: '∞'

Name: *Caleb McWhorter — Solutions*

Write your name on the appropriate line on the exam cover sheet. This exam contains 11 pages (including this cover page) and 10 questions. Check that you have every page of the exam. Answer the questions in the spaces provided on the question sheets. Be sure to answer every part of each question and show all your work. If you run out of room for an answer, continue on the back of the page — being sure to indicate the problem number.

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	10	
9	10	
10	10	
Total:	100	

1. (10 points) Consider a finite probability space with...

$$\begin{array}{ll} P(A) = 0.35 & P(B \text{ and } D) = 0.15 \\ P(B) = 0.40 & P(A \text{ and } C) = 0.10 \\ P(C) = 0.65 & P(C \text{ and } D) = 0 \\ P(D) = 0.25 & \end{array}$$

(a) Assuming A and B are independent, find $P(A \text{ and } B)$.

(b) Find $P(B \mid D)$.

(c) Find $P(A \text{ or } C)$.

(d) Are C and D independent events? Explain.

Solution.

(a) Because A and B are independent, we know $P(A \text{ and } B) = P(A) \cdot P(B)$. But then...

$$P(A \text{ and } B) = P(A) \cdot P(B) = 0.35 \cdot 0.40 = 0.14$$

(b) We know that $P(B \text{ and } D) = P(D)P(B \mid D)$. But then...

$$P(B \mid D) = \frac{P(B \text{ and } D)}{P(D)} = \frac{0.15}{0.25} = 0.60$$

(c)

$$\begin{aligned} P(A \text{ or } C) &= P(A) + P(C) - P(A \text{ and } C) \\ &= 0.35 + 0.65 - 0.10 \\ &= 0.90 \end{aligned}$$

(d) Because the probability space is finite and $P(C \text{ and } D) = 0$, we know that C and D are disjoint events. But disjoint events can never be independent. Therefore, C and D are not independent.

2. (10 points) Below is a summary of a survey of people about whether or not they preferred to shop in-person or online.

Age/Shopping	In-Person	Online	Total
18 – 30	17	34	51
30 – 50	54	61	115
50+	40	22	62
Total	111	117	228

(a) Find the percentage of people surveyed that preferred to shop in-person or were 18–30.

(b) Find the percentage of people surveyed that were 50+ and preferred to shop in-person.

(c) Find the percentage of people surveyed that preferred to shop online, if they were 30–50.

Solution. First, we find the totals of each row and column, which we insert in the table above.

(a)

$$P(\text{In-Person or 18–30}) = \frac{111 + 51 - 17}{228} = \frac{145}{228} \approx 0.6360$$

Therefore, 63.6% of people surveyed that preferred to shop in-person or were 18–30.

(b)

$$P(\text{50+ and In-Person}) = \frac{40}{228} = \frac{10}{57} \approx 0.1754$$

Therefore, 17.54% of people surveyed that were 50+ and preferred to shop in-person.

(c)

$$P(\text{Online} \mid \text{30–50}) = \frac{P(\text{Online and 30–50})}{P(\text{30–50})} = \frac{61}{115} \approx 0.5304$$

Therefore, 53.04% of people surveyed that preferred to shop online, if they were 30–50.

3. (10 points) Fifty people were surveyed about whether they had read any news online or in print in the last month. Of these people, 27 said they had read news online, 11 said they read news in print, and 5 said they read both.

   (a) Find the probability that a randomly selected person surveyed read news only online in the last month.

   (b) Find the probability that a randomly selected person surveyed had read no news in the last month.

   (c) Find the probability that a randomly selected person surveyed that read news online in the last month had also read news in print in the last month.

Solution.

(a)

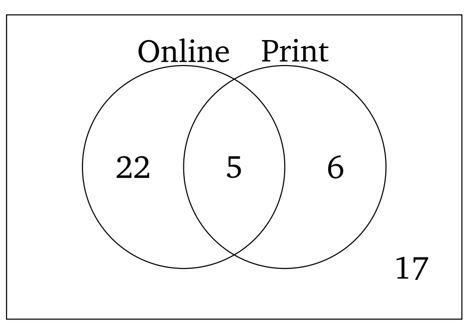
$$P(\text{Only Online}) = \frac{22}{50} = \frac{11}{25} \approx 0.44$$

(b)

$$P(\text{No News}) = \frac{17}{50} \approx 0.34$$

(c)

$$P(\text{In Print} \mid \text{Online}) = \frac{P(\text{In Print and Online})}{P(\text{Online})} = \frac{5}{22+5} = \frac{5}{27} \approx 0.1852$$



4. (10 points) Suppose in a certain town, 51% of people are women and 49% are men. Of men, 56% tend to lean conservative while only 52% of women lean conservative.

   (a) Find the probability that a randomly selected person in the town leans conservative.

   (b) Find the probability that a randomly selected person in the town is a woman or leans conservative.

   (c) Find the probability that a randomly selected person in the town is a man, assuming they do not lean conservative.

Solution.

(a)

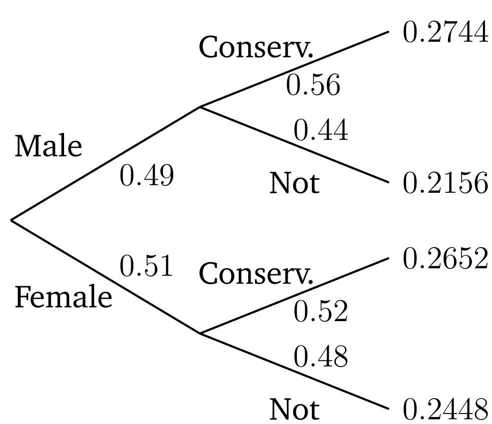
$$P(\text{Conservative}) = 0.2744 + 0.2652 = 0.5396$$

(b)

$$P(\text{Woman or Conservative}) = 0.2744 + 0.2652 + 0.2448 = 0.7844$$

(c)

$$\begin{aligned} P(\text{Man} \mid \text{Not Conservative}) &= \frac{P(\text{Man and Conservative})}{P(\text{Not Conservative})} \\ &= \frac{0.2744}{0.2156 + 0.2448} \\ &= \frac{0.2744}{0.4604} \approx 0.5960 \end{aligned}$$



5. (10 points) A certain professor's Wordle scores are given below:

# Guesses	1	2	3	4	5	6
Times Occurred	0	2	17	48	39	5

Find the average number of guesses it takes the professor to guess the word, i.e. the expected number of guesses.

Solution. The professor played a total of $0 + 2 + 17 + 48 + 39 + 5 = 111$ games of Wordle. We first find the probability that each guess occurred, which $P(\#\text{ guesses}) = \frac{\#\text{ times guessed}}{\text{Number Games Played}}$. That gives us the following table:

# Guesses	1	2	3	4	5	6
Times Occurred	0	2	17	48	39	5
Probability	0	0.0180	0.1532	0.4324	0.3514	0.0450

Therefore, the average number of guesses, i.e. the expected value, is...

$$\begin{aligned}
EX &= \sum xP(X = x) \\
&= 1(0) + 2(0.0180) + 3(0.1532) + 4(0.4324) + 5(0.3514) + 6(0.0450) \\
&= 0 + 0.036 + 0.4596 + 1.7296 + 1.757 + 0.27 \\
&= 4.2522
\end{aligned}$$

6. (10 points) Consider the following dataset:

$$\begin{array}{cccccccccc} -4 & 10 & 1 & 8 & 8 & 21 & 17 & 2 & 0 & \\ 6 & 14 & -3 & 9 & 11 & 20 & 12 & 13 & -6 & 16 \end{array}$$

(a) Find the 5-number summary.

(b) Find the IQR.

(c) Find P_{32}.

Solution. First, we put the data in order:

$$-6 \quad -4 \quad -3 \quad 0 \quad 1 \quad 2 \quad 6 \quad 8 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad 14 \quad 16 \quad 17 \quad 20 \quad 21$$

(a) The five-number summary contains the minimum, Q_1, median, Q_3, and the maximum. The minimum and maximum are clear. Observe there are 19 values. Therefore, the median is the $\frac{19+1}{2} = \frac{20}{2} = 10$th value in the dataset, which is 9. To find Q_1 and Q_3, we find the median of the values below and above the median, respectively.

There are 9 values above and below the median. Therefore, the median of these sets is the $\frac{9+1}{2} = \frac{10}{2} = 5$th value below/above the median, respectively. But then $Q_1 = 1$ and $Q_3 = 14$. Therefore, the five-number summary is...

Min	Q_1	Median	Q_3	Max
-6	1	9	14	21

(b)

$$\text{IQR} = Q_3 - Q_1 = 14 - 1 = 13$$

(c) We know P_{32} is the $19 \cdot \frac{32}{100} = 6.08 \rightsquigarrow 7$th value in the dataset. Therefore, $P_{32} = 6$.

7. (10 points) Consider the following dataset:

$$2 \qquad 4 \qquad 4 \qquad 9$$

(a) Find the mean.

(b) Find the standard deviation.

Solution.

(a) The mean is...

$$\overline{x} = \frac{\sum x_i}{n} = \frac{2+4+4+9}{4} = \frac{19}{4} \approx 4.75$$

(b)

x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
2	-2.75	7.5625
4	-0.75	0.5625
4	-0.75	0.5625
9	4.25	18.0625
	Total:	26.75

Therefore, the variance is...

$$\sigma^2 = \frac{1}{n-1} \sum (x_i - \overline{x})^2 = \frac{1}{4-1} \cdot 26.75 = \frac{1}{3} \cdot 26.75 = 8.91667$$

But then the standard deviation is $\sigma = \sqrt{\sigma^2} = \sqrt{8.91667} \approx 2.98608$.

8. (10 points) Suppose that the amount of time teenagers spend on their phone per week is normally distributed with mean 752 and standard deviation 279.

   (a) Find the percentage of teens that spend less than 600 minutes on their phone.
   (b) Find the percentage of teens that spend more than 600 minutes on their phone.
   (c) How many minutes would a teen minimally need to spend on their phone to be in the greatest 20% of minutes teenagers spend on their phone per week?

Solution.

(a) We have...

$$z_{600} = \frac{600 - 752}{279} = -0.54 \rightsquigarrow 0.2946$$

Therefore, $P(X < 600) = 0.2946$; that is, 29.46% of teens spend less than 600 minutes per week on their phone.

(b) We have...

$$P(X > 600) = 1 - P(X < 600) = 1 - 0.2946 = 0.7054$$

Therefore, 70.54% of teens spend more than 600 minutes per week on their phone.

(c) If X is the minimum number of minutes needed to be spent on the phone to be in the top 20% of teenagers. But then X is greater than 80% of the values in the distribution. But then $z_X \rightsquigarrow 0.80$. Examining the z-score table, we see that $z_X \approx 0.84$. But then...

$$\begin{aligned} z_X &\approx 0.84 \\ \frac{X - 752}{279} &\approx 0.84 \\ X - 752 &\approx 234.36 \\ X &\approx 986.36 \text{ minutes} \end{aligned}$$

9. (10 points) Suppose that a certain model of car gets an average of 38.6 miles per gallon (mpg) with standard deviation 4.2 mpg. If you took a simple random sample of 45 cars, what is the probability that their average miles per gallon was less than 38 mpg?

Solution. We are given $\mu = 38.6$ miles per gallon and $\sigma = 4.2$ mpg but *not* that the distribution of gas milage is normally distributed. Observe that the sample is a simple random sample and the sample size $n = 45 \geq 30$ is 'sufficiently large.' Therefore, the Central Limit Theorem applies. The distribution of sample averages of sample size $n = 45$ is given by...

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) = N\left(38.6, \frac{4.2}{\sqrt{45}}\right) = N(38.6, 0.6261)$$

But then...

$$z_{38} = \frac{38 - 38.6}{0.6261} = \frac{-0.6}{0.6261} \approx -0.96 \rightsquigarrow 0.1685$$

But then $P(\overline{X} < 38) = 0.1685$. That is, there is a 16.85% chance that this sample of 45 cars will get an average gas milage of less than 38 mpg.

10. (10 points) Suppose that only 3% of people can identify Moldova on a map. If you randomly surveyed 490 people, what is the probability that more than 8 people surveyed could identify Moldova on a map?

Solution. Each individual can either identify Moldova on the map or not. There are a fixed number of people surveyed, namely $n = 490$. We assume the probability that an individual can identify Moldova is a fixed $p = 0.03$. Finally, we assume that the same was independent. Then the number of individuals surveyed that could identify Moldova on a map, X, is given the binomial distribution $B(n, p) = B(490, 0.03)$. We want $P(X > 8) = P(X = 9) + P(X = 10) + \cdots + P(X = 490)$. Of course, this is too many values to compute. Instead, we will use the normal approximation to the binomial distribution.

Observe that $np = 490(0.03) = 14.7 \geq 10$ and $n(1-p) = 490(1-0.03) = 490(0.97) = 475.3 \geq 10$. Therefore, the normal approximation can be used. We know that $B(n, p) \approx N\big(np, \sqrt{np(1-p)}\big)$. But then the normal approximation is...

$$\begin{aligned} N\big(np, \sqrt{np(1-p)}\big) &= N\big(\,490(0.03),\ \sqrt{490(0.03)(1-0.03)}\,\big) \\ &= N(14.7, \sqrt{14.259}\,) \\ &= N(14.7, 3.77611) \end{aligned}$$

Therefore, $B(490, 0.03) \approx N(14.7, 3.77611)$. Now observe...

$$z_8 = \frac{8 - 14.7}{3.77611} = \frac{-6.7}{3.77611} \approx -1.77 \rightsquigarrow 0.0384$$

But then $P(X < 8) \approx 0.0384$. Therefore, $P(X > 8) \approx 1 - P(X < 8) \approx 1 - 0.0384 = 0.9616$. Therefore, there is an approximately 96.16% chance that more than 8 individuals surveyed could identify Moldova on a map.