MAT 108: Exam 3 **Name:** *Caleb McWhorter — Solutions*
Spring – 2022
05/11/2022
85 Minutes

Write your name on the appropriate line on the exam cover sheet. This exam contains 10 pages (including this cover page) and 9 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	16	
2	8	
3	10	
4	10	
5	10	
6	10	
7	10	
8	13	
9	13	
Total:	100	

1. A student is given a set of data points on an exam to for which they are to create a linear regression. Suppose that the points are $(-2,1)$, $(0,5)$, $(2,3)$, and $(4,5)$. Their computation of the least square regression line is shown below. Unfortunately, they spilled coffee on their work, obscuring some of the numbers.

   (a) (7 points) Recompute (i)–(vii) for the student and place your answers in the appropriate space below.

$$\overline{x} = \text{(i)}$$

$$\overline{y} = \frac{1+5+3+5}{4} = \frac{14}{4} = 3.5$$

$$s_x^2 = \frac{1}{3} \cdot 20 \approx 6.6667 \Longrightarrow s_x \approx 2.582$$

$$s_y^2 = \text{(ii)} \Longrightarrow s_y = \text{(iii)}$$

x	y	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$y_i - \overline{y}$	$(y_i - \overline{y})^2$	$\frac{x_i - \overline{x}}{s_x}$	$\frac{y_i - \overline{y}}{s_y}$	$\frac{x_i - \overline{x}}{s_x} \cdot \frac{y_i - \overline{y}}{s_y}$
-2	1	(iv)	9	-2.5	6.25	-1.1619	-1.3056	1.517
0	5	-1	1	1.5	2.25	-0.3873	0.7833	-0.3034
2	3	1	1	-0.5	(v)	0.3873	-0.2611	-0.1011
4	5	3	9	1.5	2.25	1.1619	0.7833	(vi)
		Sum:	(vii)	Sum:	11		Sum:	2.0226

$$\text{(i)} = \frac{-2+0+2+4}{4} = 1 \qquad \text{(v)} = (-0.5)^2 = 0.25$$

$$\text{(ii)} = \frac{1}{3} \cdot 11 \approx 3.6667 \qquad \text{(vi)} = 1.1619 \cdot 0.7833 \approx 0.9101$$

$$\text{(iii)} = \sqrt{3.6667} \approx 1.9149 \qquad \text{(vii)} = 9+1+1+9 = 20$$

$$\text{(iv)} = -2 - 1 = -3$$

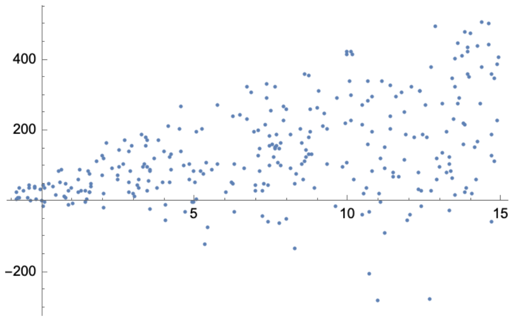
   (b) (5 points) Compute r^2 for the student. Based on this value, explain to the student whether the least square regression line is a 'good' model.

   We know $r = \frac{1}{3} \cdot 2.0226 = 0.6742$. *But then* $r^2 \approx 0.4545$. *Because this value is 'small', i.e. less than* 0.85, *this linear model is not a 'good' model.*

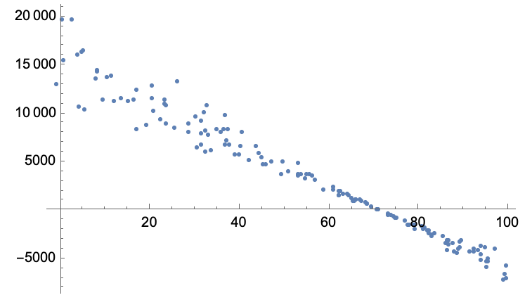
   (c) (4 points) Given that the resulting linear regression is $\widehat{y} = 0.5x + 3$, find the predicted value for $x = 2$ and its corresponding residual for the student.

   The predicted value for $x = 2$ *is* $\widehat{y}(2) = 0.5(2) + 3 = 4$ *with residual* $y_i - \widehat{y}_i = 3 - 4 = -1$.

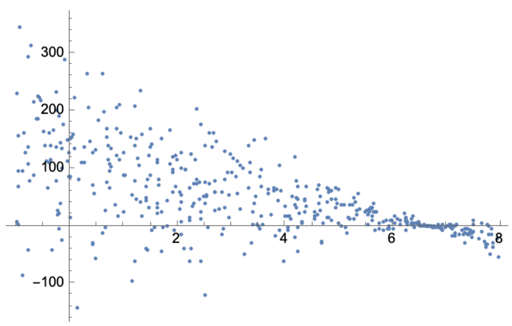
2. (8 points) The same student from Problem 1 also had to compute the regression coefficient, R, for several different data sets. Unfortunately, they dropped their papers—separating the data from the computed R value. Match the dataset to the most likely regression coefficient for the student.



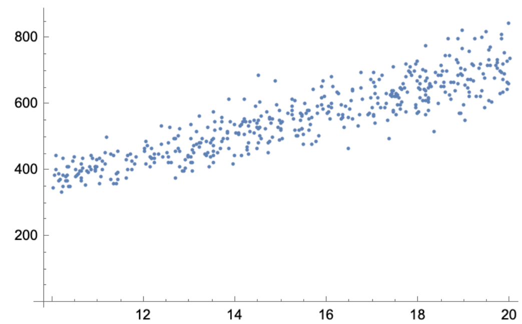
(a)



(b)



(c)



(d)

(i) ___(b)___: $R = -0.9831$

(ii) ___(c)___: $R = -0.6050$

(iii) ___(a)___: $R = 0.4635$

(iv) ___(d)___: $R = 0.9032$

3. Suppose A, B, and C are events in a finite probability space. Suppose $P(A) = 0.40$, $P(B) = 0.20$, $P(C) = 0.30$, $P(A \text{ and } B) = 0$, and $P(B \text{ and } C) = 0.05$.

   (a) (2 points) Assuming A and C are independent, compute $P(A \text{ and } C)$.

   Because A and C are independent,

   $$P(A \textit{ and } C) = P(A) \cdot P(C) = 0.40 \cdot 0.30 = 0.12$$

   (b) (2 points) Still assuming that A and C are independent, find $P(A \text{ or } C)$.

   $$P(A \text{ or } C) = P(A) + P(C) - P(A \text{ and } C) = 0.40 + 0.30 - 0.12 = 0.58$$

   (c) (2 points) Are A and B disjoint events? Explain.

   Because $P(A \textit{ and } B) = 0$, A and B are disjoint events.

   (d) (2 points) Are A and B independent events? Explain.

   By (c), we know that A and B are disjoint. Therefore, A and B are not independent because disjoint events can never be independent.

   (e) (2 points) Find $P(B \mid C)$.

   $$P(B \mid C) = \frac{P(B \text{ and } C)}{P(C)} = \frac{0.05}{0.30} \approx 0.1667$$

4. Students across various STEM majors at local colleges were surveyed to determine whether they preferred Netflix, Hulu, Disney+, or HBO Max. The results are shown below.

	Biology	Chemistry	Computer Science	Physics	Mathematics	Total
Netflix	3	7	14	4	8	36
Hulu	6	13	12	3	7	41
Disney+	15	4	1	4	2	26
HBO Max	10	2	5	2	1	20
Total	34	26	32	13	18	123

(a) (2 points) What is the probability that a randomly selected student was a Physics major that preferred Disney+?

$$P(\text{Physics and Disney+}) = \frac{4}{123} \approx 0.0325$$

(b) (2 points) What is the probability that a randomly selected student preferred Hulu?

$$P(\text{Hulu}) = \frac{41}{123} \approx 0.3333$$

(c) (2 points) What is the probability that a randomly selected student preferred Netflix or was a Chemistry major?

$$P(\text{Netflix or Chemistry}) = \frac{36 + 26 - 7}{123} = \frac{55}{123} \approx 0.4472$$

(d) (2 points) What is the probability that a Biology or Computer Science major preferred Netflix?

$$P(\text{Netflix} \mid \text{Bio or CS}) = \frac{3 + 14}{34 + 32} = \frac{17}{66} \approx 0.2576$$

(e) (2 points) What is the probability that a student that preferred HBO Max was a Mathematics major?

$$P(\text{Math} \mid \text{HBO Max}) = \frac{1}{20} = 0.05$$

5. Twenty-three students in a middle school class were asked whether they had any pets. Of the students, eight said that they had a dog, five said they had a cat, and two students said that they had both.

   (a) (3 points) What is the probability that a randomly selected student did not have a dog or cat?

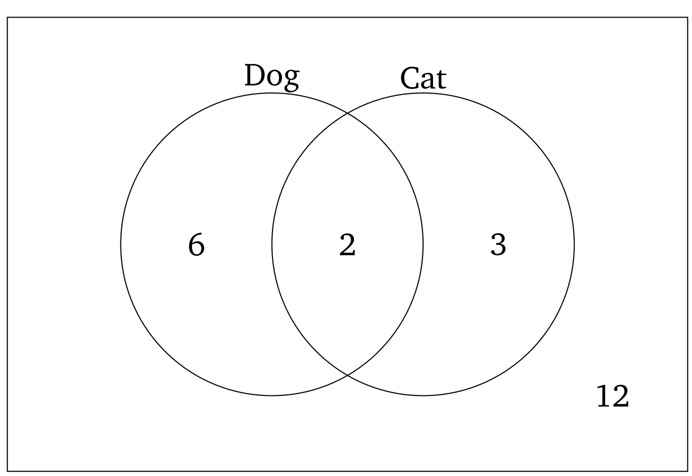
   $$P(\text{No Dog or Cat}) = \frac{12}{23} \approx 0.5217$$

   (b) (3 points) What is the probability that a randomly selected student had a dog or a cat?

   $$P(\text{Dog or Cat}) = \frac{6+2+3}{23} = \frac{11}{23} \approx 0.4783$$

   (c) (4 points) Assuming that a student with a cat was randomly chosen, what is the probability that they have a dog?

   $$P(\text{Dog} \mid \text{Cat}) = \frac{2}{2+3} = \frac{2}{5} = 0.40$$



6. When it comes to 'fine' Mexican dining, people agree the two best options are Chipotle or Moe's. One survey suggested 80% of people prefer Chipotle to Moe's. However of those surveyed, only 40% of people that went to Chipotle had a good experience whereas 60% of Moe's customers had a good experience.

   (a) (3 points) What is the probability that a randomly selected person had a good experience at Chipotle or Moe's?

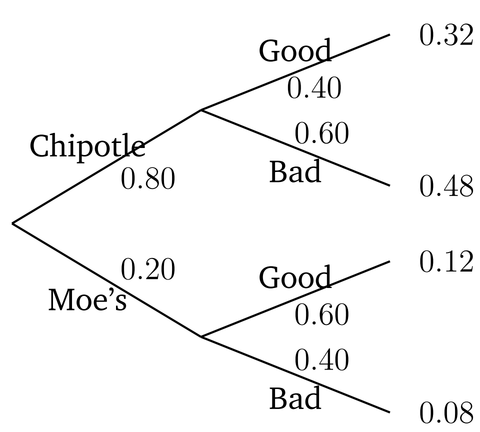
$$P(\text{Good at Chipotle/Moe's}) = 0.32 + 0.12 = 0.44$$

   (b) (3 points) What is the probability that a randomly selected person either preferred Moe's or had a bad experience at Chipotle?

$$P(\text{Prefer Moe's or Bad at Chipotle}) = 0.48 + 0.12 + 0.08 = 0.68$$

   (c) (4 points) What is the probability that a person that did not have a good experience had it at Chipotle?

$$P(\text{Chipotle} \mid \text{Bad Experience}) = \frac{0.48}{0.48 + 0.08} = \frac{0.48}{0.56} \approx 0.8571$$



7. "Step right up," yells a clown at a carnival. "Try your luck and see if fortune favors you today!" You walk up to the booth. The clown explains that you can pay \$1 to try your luck at a dice rolling game. If you roll either a one, two, three, or four, you win nothing. If you roll a five, you receive \$0.50. However, if you roll a six, then you win \$5.

   (a) (7 points) Compute the amount, on average, you can expect to win playing this game.

   The expected value, i.e. average amount you win per game, is...

$$\begin{aligned}
\mu &= \sum_i x_i p_i \\
&= 0 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 0.50 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} \\
&= 0 + 0 + 0 + 0 + \frac{1}{12} + \frac{5}{6} \\
&= \frac{11}{12} \approx \$0.92
\end{aligned}$$

   (b) (3 points) In the long run, should you play this game? Explain your reasoning using your computation in (a).

   No, you should not play this game. We know from (a) that you only win \$0.92 per game, on average. However, you are paying \$1 to play this game. Therefore, you make on average $\$0.92 - \$1 = -\$0.08$ *per game, i.e. you are losing money on average.*

8. High school students looking to go to college will often take either the SAT or ACT. The ACT parent company data suggests that ACT scores in New York State were normally distributed with mean 26.3 and standard deviation 2.3.

   (a) (3 points) What percent of students in NYS receive below 23.2 on the ACT?

   We have...

   $$z_{23.2} = \frac{23.2 - 26.3}{2.3} = \frac{-3.1}{2.3} \approx -1.35 \rightsquigarrow 0.0885$$

   Therefore, $P(X \leq 23.2) = 0.0885$.

   (b) (3 points) What percent of students in NYS receive above 32.7 on the ACT?

   We have...

   $$z_{32.7} = \frac{32.7 - 26.3}{2.3} = \frac{6.4}{2.3} \approx 2.78 \rightsquigarrow 0.9973$$

   Therefore, $P(X \geq 32.7) = 1 - P(X \leq 32.7) = 0.0027$.

   (c) (3 points) What percent of students in NYS receive between 23.2 and 32.7 on the ACT?

   $$P(23.2 \leq X \leq 32.7) = 0.9973 - 0.0885 = 0.9088$$

   (d) (4 points) If you surveyed 80 students that took the ACT in NYS, what is the probability that their average score was below 25?

   Because the underlying distribution is normally distributed (or because the sample size is $80 \geq 30$*), we know by the Central Limit Theorem that the sampling distribution for averages of 80 students is* $N(\mu, \sigma/\sqrt{n}) = N(26.3, 2.3/\sqrt{80}) \approx N(26.3, 0.2571)$. *We have...*

   $$z_{25} = \frac{25 - 26.3}{0.2571} = \frac{-1.3}{0.2571} \approx -5.06 \rightsquigarrow 0.00$$

   Therefore, $P(\overline{X} \leq 25) \approx 0$.

9. There are a plethora of quick dining options in the United States. A recent survey approximates that 1 in 5 people have been to a Sonic drive-in. Suppose that ten people are randomly surveyed.

   (a) (3 points) What is the probability that exactly three of them had been to a Sonic before?

   $$P(X = 3) = 0.2013$$

   (b) (3 points) What is the probability that less than four of them had been to a Sonic before?

   $$\begin{aligned} P(X < 4) = P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0.1074 + 0.2684 + 0.3020 + 0.2013 \\ &= 0.8791 \end{aligned}$$

   (c) (3 points) What is the probability that at least one of them had been to a sonic before?

   $$P(X \geq 1) = 1 - P(X = 0) = 1 - 0.1074 = 0.8926$$

   (d) (4 points) Using the normal approximation to the binomial distribution, approximate the probability that if 300 people were surveyed that less than fifty of them had been to a Sonic before?

   We have $np = 60 \geq 10$ *and* $n(1-p) = 240 \geq 10$. *By the Central Limit Theorem, we know that this binomial distribution is approximately* $N(np, \sqrt{np(1-p)}) \approx N(60, 6.9282)$. *We have. . .*

   $$z_{50} = \frac{50 - 60}{6.9282} = \frac{-10}{6.9282} \approx -1.44 \rightsquigarrow 0.0749$$

   Therefore, $P(X \leq 50) \approx 0.0749$.