Write your name on the appropriate line on the exam cover sheet. This exam contains 11 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	15	
2	10	
3	10	
4	15	
5	10	
6	10	
7	10	
8	10	
9	10	
Total:	100	

1. (15 points) Researchers studying changing viewpoints for youths in education took a survey of 1,001 students in middle school. The researchers asked the students what they believed was the most useful thing in school: having high grades, being popular, or athletic ability. The results, broken down by type of school, are found below.

	Public	Private	Charter	Total
Grades	87	115	82	284
Popularity	165	91	177	433
Athleticism	102	88	94	284
Total	354	294	353	1001

Assuming that this survey is representative of public, private, and charter schools as a whole, determine...

(a) ...the percentage of students that attend private schools.

(b) ...the percentage of students that believe grades are the most useful thing.

(c) ...the percentage of students that are in charter schools and believe popularity is the most useful thing.

(d) ...the percentage of students in public schools that believe athleticism is the most useful thing.

(e) ...the percentage of students in public school, private school, or believe that athleticism is the most important thing.

Solution.

(a) ...the percentage of students that attend private schools.

$$P(\text{private}) = \frac{294}{1001} \approx 0.2937 \rightsquigarrow 29.37\%$$

(b) ...the percentage of students that believe grades are the most useful thing.

$$P(\text{grades}) = \frac{284}{1001} \approx 0.2837 \rightsquigarrow 28.37\%$$

(c) ...the percentage of students that are in charter schools and believe popularity is the most useful thing.

$$P(\text{charter and popularity}) = \frac{177}{1001} \approx 0.1768 \rightsquigarrow 17.68\%$$

(d) ... the percentage of students in public schools that believe athleticism is the most useful thing.

$$P(\text{athleticism} \mid \text{public}) = \frac{102}{354} \approx 0.2881 \rightsquigarrow 28.81$$

(e) ... the percentage of students in public school, private school, or believe that athleticism is the most important thing.

$$P(\text{public or private or athleticism}) = \frac{354 + 294 + 284 - 102 - 88}{1001} = \frac{742}{1001} \approx 0.7413 \rightsquigarrow 74.13\%$$

2. (10 points) Several medical researchers believe that 'genetic drift' is causing the average cholesterol levels in adults to increase. Hence, previously used thresholds of 'high cholesterol' may be too low. The researchers measured the cholesterol levels of 265 individuals and found a sample average of 194.1 mg/dL. If cholesterol levels in adults have a standard deviation of 3.8 mg/dL, construct a 92% confidence interval for the average cholesterol level in adults.

Because the sample size of 265 individuals is sufficiently large, we know that the Central Limit Theorem applies. The underlying population standard deviation is $\sigma = 3.8$ mg/dL. We have a sample average of $\overline{x} = 194.1$ mg/dL on a sample size of $n = 265$ individuals. We require a z-score that corresponds to $92\% + \frac{100\% - 92\%}{2} = 92\% + 4\% = 96\%$. But then we have $z^* \approx 1.75$. But then we have...

$$\overline{x} \quad \pm \quad z^* \frac{\sigma}{\sqrt{n}}$$

$$194.1 \text{ mg/dL} \quad \pm \quad 1.75 \cdot \frac{3.8 \text{ mg/dL}}{\sqrt{265}}$$

$$194.1 \text{ mg/dL} \quad \pm \quad 1.75 \cdot 0.233 \text{ mg/dL}$$

$$194.1 \text{ mg/dL} \quad \pm \quad 0.408 \text{ mg/dL}$$

Note that 194.1 *mg/dL* $- 0.408$ *mg/dL* ≈ 193.7 *mg/dL and* 194.1 *mg/dL* $+ 0.408$ *mg/dL* ≈ 194.5 *mg/dL.*

Therefore, we are 92% confident that the average adult blood pressure is between 193.7 mg/dL and 194.5 mg/dL.

3. (10 points) Mike is playing a D&D game with a ten-sided die. He comes across a demogorgon—#FearTheDemogorgon. Unfortunately, the demogorgon has cloned itself, creating 12 copies. Mike needs to roll a 10 to defeat any single demogorgon.

  (a) Find the probability that Mike defeats exactly two demogorgons.

  (b) Find the probability that Mike defeats four or more demogorgons.

  (c) Find the probability that Mike defeats all the demogorgons.

  (d) Find the probability that Mike defeats at least one of the demogorgons.

Solution. *Because each dice roll is independent, each roll is either a 'success' or 'failure' (rolling a 10 or not), there is a fixed number of rolls (12 rolls), and the probability of success is fixed ($p = \frac{1}{10}$), this is a binomial distribution, namely $B\left(12, \frac{1}{10}\right)$.*

(a)
$$P(X = 2) = 0.2301$$

(b)
$$P(X \geq 4) = P(X = 4) + P(X = 5) + \cdots + P(X = 11) + P(X = 12)$$
$$\approx 0.0213 + 0.0038 + 0.0005 + 0.0000 + \cdots + 0.0000$$
$$= 0.0256$$

(c)
$$P(X = 12) \approx 0.0000$$

(d)
$$P(X \geq 1) = 1 - P(X = 0) \approx 1 - 0.2824 = 0.7176$$

4. (15 points) Students in a class on quantitative methods in business and social science are studying for an exam. If a student studies, there is an 80% chance that they will pass the exam. If a student does not study, there is a 75% chance that they will fail the exam. After a brief survey, the instructors estimates that only 60% of the students will study for the exam.

   (a) Find the percentage of the students that will study for the exam and pass.
   (b) Find the percentage of students that will fail the exam.
   (c) Find the percentage of students that will pass the exam.
   (d) Find the percentage of students that either did not study or will fail the exam.
   (e) What percentage of students that did not study will pass the exam anyway?

   Solution.

   (a)
   $$P(\text{study and pass}) = 0.48$$

   (b)
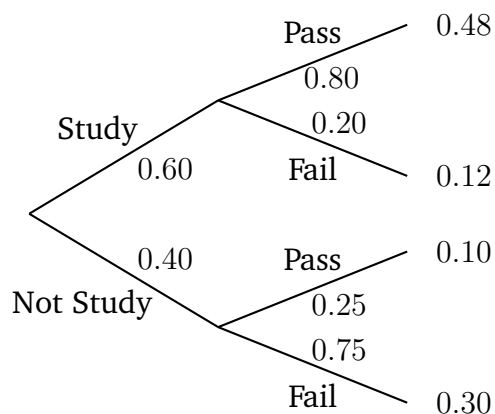 $$P(\text{fail}) = 0.12 + 0.30 = 0.42$$

   (c)
   $$P(\text{pass}) = 0.48 + 0.10 = 0.58$$

   (d)
   $$P(\text{not study or fail}) = 0.10 + 0.30 + 0.12 = 0.52$$

   (e)
   $$P(\text{pass} \mid \text{not study}) = \frac{0.10}{0.10 + 0.30} = \frac{0.10}{0.40} = 0.25$$

5. (10 points) Students in a class on quantitative methods in business and social science have just taken an exam. The exam scores were normally distributed with a mean of 45 and standard deviation 11.3.

    (a) Find the percentage of students that scored at least 65 on the exam.

    (b) Find the minimal score one would need to be in the top 10% of exam scores.

    Solution.

    (a) *We have...*
    $$z_{65} = \frac{65 - 45}{11.3} = \frac{20}{11.3} \approx 1.77 \rightsquigarrow 0.9616$$
    But because we want $X \geq 65$, we have $1 - 0.9616 = 0.0384$. Therefore, 3.84% of students scored at least a 65 on the exam.

    (b) *If you received a score, say x, to be in the top 10%, then 90% of students would have scored less than you. But then $z_x \rightsquigarrow 0.90$. Therefore, we know that $z_x \approx 1.28$. Then we have...*
    $$z_x = 1.28$$
    $$\frac{x - \mu}{\sigma} = 1.28$$
    $$\frac{x - 45}{11.3} = 1.28$$
    $$x - 45 = 14.464$$
    $$x = 59.464$$
    Therefore, a score of at least approximately 59.5 would put you in the top 10% of exam scores.

6. (10 points) Local government officials are trying to determine how to allocate public health funds from a recent grant. It is known that approximately 26% of adults in the US have some type of disability. One of the villages in the county has a population of 2,500 adults.

    (a) What is the probability that less than 620 of these adults will have a disability?

    (b) What is the probability that less than 720 of these adults will have a disability?

    (c) What is the probability that between 620 and 720 of these adults will have a disability?

Solution. Suppose that the sample is a simple random sample and that each selection for the sample is independent from the next. Because we have a fixed population of 2,500 adults, each person selected is a 'success' or not (has a disability or not), and the chance of 'success' is fixed with probability 0.26, this is binomial, namely $B(2600, 0.26)$. Because $np = 2600(0.26) = 676 \geq 10$ and $n(1 - p) = 2600(1 - 0.26) = 2600(0.74) = 1924 \geq 10$, we can approximate this binomial distribution with a normal distribution: $B(n, p) \approx N(np, \sqrt{np(1 - p)})$. Therefore, this distribution is approximately $N(2600(0.26), \sqrt{2600(0.26)(1 - 0.26)}) \approx N(676, 22.366)$.

  (a) We have...
  $$z_{620} = \frac{620 - 676}{22.366} = \frac{-56}{22.366} \approx -2.50 \rightsquigarrow 0.0062$$
  Therefore, $P(X \leq 620) \approx 0.0062$.

  (b) We have...
 $$z_{720} = \frac{720 - 676}{22.366} = \frac{44}{22.366} \approx 1.97 \rightsquigarrow 0.9756$$
  Therefore, $P(X \leq 720) \approx 0.9756$.

  (c) Using (a) and (b), we know that...
  $$P(620 \leq X \leq 720) = P(X \leq 720) - P(X \leq 620) \approx 0.9756 - 0.0062 = 0.9694$$

7. (10 points) A researcher is collecting data on sleep and academic performance. They collect data on student's GPA, g, and plot it against the average number of hours of sleep, h, students had during a given week. They find a linear regression model of $\widehat{g} = 0.028h + 3.5$. A scatterplot of their data with the model is found below.
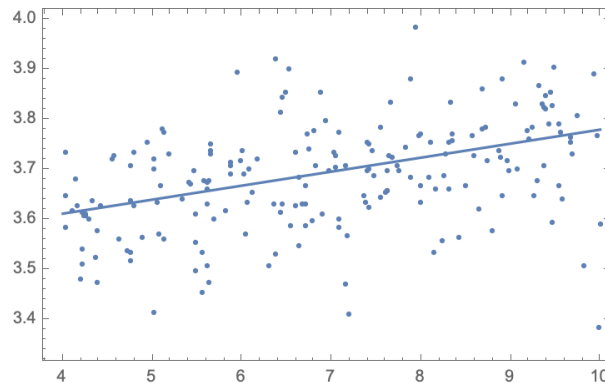
(a) Find the model's prediction of a student's GPA if they slept 8 hours a night.

(b) If a student that slept 8 hours a night had a GPA of 3.528, find the residual for this model's prediction of this datapoint.

(c) If the researchers had an R^2 value of 0.407263, does this imply a 'good' model, i.e. a strong correlation? Explain.

(d) Is (c) supported by the scatterplot above? Explain.

Solution.

(a) We have

$$\widehat{y}(8) = 0.028(8) + 3.5 = 0.224 + 3.5 = 3.724$$

Therefore, the model predicts that a student that sleeps 8 hours a night will have a GPA of 3.724.

(b) We have. . .

$$e_i = y - \widehat{y} = 3.528 - 3.724 = -0.196$$

(c) This is not a 'good' model. The R^2 value essentially tells you the 'percent linearity' of the data. Specifically, $R^2 = 0.407263$ implies that approximately 40.7% of the variation in the data is explained by this linear model. The closer R^2 is to 1, the better the explanatory power of the model. Typically, we want R^2 to be at least 0.60. Therefore, this R^2 value demonstrates a weak predictive power for the model.

(d) Yes. Observe that the line does not 'tightly fit' the data. Hence, we expect a rather low value for R^2. The larger the R^2-value, the 'closer' the points are to the linear model.

8. (10 points) A company claims that their warehouse workers are paid more than the national average of $23.75 per hour. A local journalist doubts the company's claim and surveys 15 of the workers at the warehouse about their hourly wage. The journalist assume that wages at the warehouse have the same standard deviation of $1.27 as wages nationally for this type of job.

   (a) Find the probability that the average salary of these workers is less than $23 per hour.

   (b) What assumption(s) about the workers' salary do your calculations in (a) require, if any.

   Solution.

   (a)

   (b)

9. (10 points) Read the following two scenarios:

   I. A politician is making claims about immigrants into their country. They cite data showing that 68.3% of these immigrants do not have higher education. Because studies have shown that 4.12% of those without higher education will commit violent crimes, they claim that $68.3 \cdot 4.12 \approx 2.81\%$ of these immigrants will commit a violent crime. They also claim that because there were $10,000$ new immigrants last month, there are now $10,000 \cdot 0.0281 \approx 281$ new violent criminals in the area.

   II. Efficiency managers have been hired at a local law firm to improve 'work flow' at their company. The efficiency managers gather data and find that 40% of workers handle some type of paperwork while 33% of workers handle some type of data entry. They conclude that at least $40\% + 33\% = 73\%$ of workers at the company either handle some type of paperwork or some type of data entry.

   Now respond to the following two questions:

   (a) Is the mathematical claim the politician made in (I) correctly reasoned? Explain.

   (b) Is the mathematical claim the efficiency managers made in (II) correctly reasoned? Explain.

   Solution.

   (a)

   (b)