

LAST NAME (Please Print): **KEY**

FIRST NAME (Please Print): _____

HONOR PLEDGE (Please Sign): _____

Statistics 111

Midterm 1

- This is a closed book exam.
- You may use your calculator and a single page of notes.
- The room is crowded. Please be careful to look only at your own exam. Try to sit one seat apart; the proctors may ask you to randomize your seating a bit.
- Report all numerical answers to at least two correct decimal places or (when appropriate) write them as a fraction.
- All question parts count for 1 point.

1. For prenatal care, pregnant women could attend one of two clinics (A or B), and receive either basic care or extra care if their insurance covered it. The following table shows whether or not their children lived.

Clinic	Care	Lived	Died
A	Basic	176	3
A	Extra	293	4
B	Basic	197	17
B	Extra	23	2

0.032 What is $p_B - p_E$, where p_B is the mortality rate for those with basic care and p_E is the mortality rate for those with extra care? Round to three accurate decimal places.

$$p_B = (3 + 17)/(179 + 214) = 0.0509 \text{ and } p_E = (4 + 2)/(297 + 25) = 0.1863 \text{ so } p_B - p_E = 0.0323.$$

What is $p_B - p_E$ just for infants at Clinic A? For just the infants at B? Round to three accurate decimal places.

$$\text{A: } 0.003 \qquad \text{B: } -0.001$$

$$\text{A: } (3/179 - 4/297 = 0.003291 \qquad \text{B: } 17/214 - 2/25 = -0.00056$$

0.038 Using weighted averages, find the adjusted fatality rate for infants whose mothers received basic care. Round to three decimal places.

For the Berkeley example, life/death is accept/reject, basic/extra is male/female, and clinic is the confounding course. So

$$\frac{179 + 297}{179 + 297 + 214 + 25} \frac{3}{179} + \frac{214 + 25}{179 + 297 + 214 + 25} \frac{17}{214} = 0.0377$$

A similar adjusted mortality rate is found for those receiving extra care. Explain what is going on.

There is a much higher mortality rate at Clinic B, and the people at Clinic B were less likely to have insurance. So there is no important difference in health outcome between basic and extra care. Simpson's paradox.

2. **0.64** A professor says that about 30% of the class will get an A- or better for the final grade. In a class of 90, what is the probability that more than 25 people will get some kind of A?

The binomial has mean $0.3 * 90 = 27$ and sd $\sqrt{0.3 * 0.7 * 90} = 4.3474$. So the continuity-corrected z-transformation is $z = (25.5 - 27)/4.3474 = 0.34503$. From the table, the area above 0.35 is 0.3632.

3. Consider the density function $f(x) = 3(4x - 2x^2)/8$ on the interval $[0, 2]$ and 0 elsewhere.

0.32 What is the probability that $X \leq 0.75$?

$F(x) = \int_0^x f(y) dy = \frac{3}{8(2x^2 - \frac{2}{3}x^3)}$ and evaluating this at 0.75 gives 0.3164.

1 What is $\mathbb{E}[X]$?

$$\int_0^2 xf(x) dx = \frac{3}{8} * \left[\frac{4}{3}x^3 - \frac{1}{2}x^4 \right]_0^2 = 1$$

0.45 What is the standard deviation of X ?

$\mathbb{E}[X^2] = \int_0^2 x^2 f(x) dx = (3/8) * \left[x^4 - \frac{2}{5}x^5 \right]_0^2 = 1.2$. So the variance is $1.2 - (1)^2 = 0.2$ and thus the sd is $\sqrt{0.2} = 0.4472$.

4. From a class of 100, I randomly select eight students and ask how many pennies they have in their pocket. The results are:

5, 0, 9, 10, 7, 8, 4, 2

5.5 What is the IQR?

The first quartile is any number between 2 and 4, so we take 3. The third quartile is any number between 8 and 9, so we take 8.5.

3.50 What is the standard deviation?

It is a sample so divide by $n - 1$ in sd formula. Answer is 3.5026.

24 Suppose a sample has standard deviation equal to 4. You first subtract 2 from each number and then multiply by -6. What is the new standard deviation?

$$|-6| * 4 = 24$$

30 Suppose a set of numbers has median equal to -3. You first subtract 2 from each number and then multiply by -6. What is the new median?

The order of operations matters. $(-3 - 2) * (-6) = 30$

5. 0.22 A meteor will destroy the world in 30 years, and a young statistician is in love and gets married. She knows that the duration of a marriage is exponentially distributed with mean 20 years. What is the probability that her marriage will last until The End?

$$1 - (1 - \exp(-30/20)) = 0.2231. \text{ Recall that the mean of an exponential is } 1/\lambda.$$

6. 0.33 You are dealt five cards from a standard deck. What is the probability of getting exactly three red cards?

$$\text{Hypergeometric. } \frac{\binom{26}{3} \binom{26}{2}}{\binom{52}{5}} = 0.3251$$

7. **0.06** You roll a fair die until you get your second six, and then you stop. What is the probability the you stop on roll 10?

The 10th roll must be a six, with probability $1/6$. The other nine rolls must contain one six and eight non-sixes. Any such arrangement has probability $(1/6)(5/6)^8$ and there are nine such arrangements. So multiplication gives 0.0581.

8. You have a box with ten tickets, and each ticket contains one letter from the word STATISTICS.

0.025 Make three draws without replacement. What is the probability that you obtain S, S, T in that order?

$$(3/10)(2/9)(3/8) = 0.025.$$

0.075 Make three draws without replacement. What is the probability that you obtain S, S, T in any order?

$$3 * 0.025 = 0.075$$

0.03 Make three draws with replacement. What is the probability that you obtain S, S, T in that order?

$$(3/10)(3/10)(3/10) = 0.027.$$

0.08 Make three draws with replacement. What is the probability that you obtain S, S, T in any order?

$$\text{Multinomial. } \frac{3!}{2!*1!*0!*0!*0!} (3/10)^2 * (3/10)^1 * (1/10)^0 * (2/10)^0 * (1/10)^0 = 0.081.$$

9. Suppose the joint density of X and Y is $f(x, y) = 8xy$ for $0 \leq x \leq y \leq 1$, and it is otherwise 0.

What is the marginal density of Y ?

$$f_2(y) = \int_0^y 8xy \, dx = 4y^3 \text{ for } 0 \leq y \leq 1 \text{ and } 0 \text{ otherwise.}$$

What is the conditional density of X given $Y = y$?

$$g_1(x | y) = f(x, y) / f_2(y) = 8xy / 4y^3 = 2x / y^2 \text{ for } 0 \leq x \leq y.$$

0.44 What is the probability that $X \leq 0.5$ when $Y = 0.75$?

$$G_2(x \leq 0.5 | Y = 0.75) = \int_0^{0.5} 2x / (0.75)^2 \, dx = 0.444.$$

10. A study by the CDC took a random sample of adults living in 1960 who were not morbidly obese and measured their weight. In 1970, they checked to see how many people in that sample were still alive, and found that people who weighed more were much less likely to survive. Based on this the CDC launched a campaign to encourage people to lose weight. (This is actually approximately what happened.)

Yes Is this an observational study?

Give a *sensible* explanation that does not support the CDC's conclusion.

As one gets older, one tends to gain weight. People who weighed more in 1960 were older, and more likely to die by 1970. Or, people who are thin tend to have more healthy lifestyles.

11. Suppose $f(x, y) = c$ for $(x - 1)^2 + (y - 3)^2 = 2$, and is otherwise 0.

$$c = 0.11 \quad \mathbb{E}[X] = 1 \quad \text{Cov}(X, Y) = 0$$

As in quiz 7. The $c = 1 / (2\sqrt{2}\pi)$.

12. List all, and only, the true statements. **B, C, D, H, I**
- A. Newton was the greatest mathematician ever.
 - B. If three events are mutually exclusive and one must happen, they form a finite partition.
 - C. Tchybeshev showed that $\mathbb{P}[|X - \mu| \leq a\sigma] \geq 1 - (1/a)^2$.
 - D. A confounding factor is one that is associated with both group membership and the outcome.
 - E. A placebo is often used in observational studies to ensure blinding.
 - F. If $Y = aX + b$ then X and Y have correlation 1.
 - G. If the correlation between X and Y is zero, then they are independent.
 - H. The marginal of X is $\int_{-\infty}^{\infty} f(x, y) dy$.
 - I. $\mathbb{P}[A \text{ and } B] = \mathbb{P}[A] * \mathbb{P}[B | A]$.
 - J. The mean of an exponential distribution is λ .