

Print Name: _____

Key

Math 127 – Exam 2 – Fall 2017

Version Macaroni

Oath: "I will not discuss the exam contents with anyone on planet Earth until the answer key is posted to Blackboard."

Sign Name: _____

Key

The penalty for cheating on this Exam is a grade of 0% for Math 127 Exam 1.

Student Instructions

1. This test is graded out of 100 points and counts for 20% of your Math 127 grade. Points are in parentheses for each question.
2. You can use a calculator, but you cannot use your phone. You can use the calculator on the computers if you wish.
3. You will need to use www.statcrunch.com. This is the only permitted webpage.
4. You are permitted to use one 8.5" by 11" sheet of notes, front and back. You will submit it with your test.

You may not use the pink sheet or copies of the pink sheet.

You must produce (handwritten or typed up) your own sheet of notes.

You may not use copies or scans of any instructor-created Math 127 content or answer keys.

5. Show work or points will be deducted. If you only report an answer and it is wrong, you will receive no credit.

Probability Part: 28 questions worth 2 points each = 56 points

Regression Part: 1 question worth 44 points

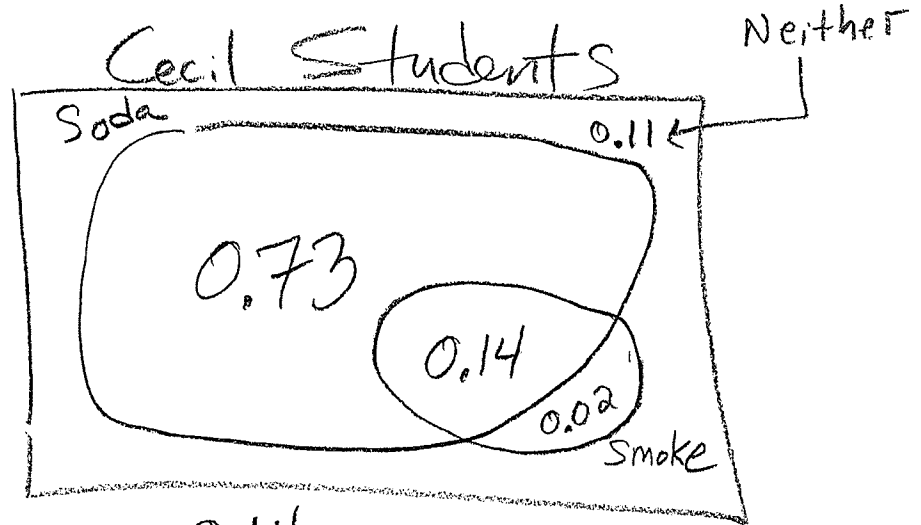
—

—

—

1. We have the following information about Cecil College students: 16% smoke cigarettes, 87% drink soda, and 14% do both.

1a. Draw and label a Venn diagram:



- 1b. Take a random student, $P(\text{Smoke} | \text{Drink Soda}) = \frac{0.14}{0.87} = 0.1609$
- 1c. Take four random students, $P(\text{No one smokes}) = (0.84)^4 = 0.4979$

2. An episode of "*Keeping Up With The Kardashians*" has a run time that follows a Uniform model on the interval [41.5 minutes, 42.5 minutes]. The rest of the hour is filled with commercials mixed in at random times.

- 2a. Give the probability function, $f(x) = 1, 41.5 \leq x \leq 42.5$
- 2b. What is the probability that the next episode's run time exceeds 42.2 minutes? 0.30
- 2c. What is the 19th percentile of the probability model? 41.69 minutes
- 2d. Presume the four episodes in November are all independent of each other. Determine the probability that at least one episode has a run time under 42 minutes. Show work or justify answer:

$$P(\text{at least one under } 42) =$$

$$1 - P(4 \text{ for } 4 \text{ go over } 42) =$$

$$1 - (0.5)^4 = 0.9375$$

Extra Credit (2 points): Flip on the TV to a random episode at a random time.

$$P(\text{Kardashians are on commercial break}) = \frac{18}{60} = 0.30$$

3. In its prime, the TV program "*Kourtney and Kim Take Miami*" averaged 2.09 million viewers with a standard deviation of 0.24 million viewers. A Normal model applies.

- 3a. Interpret the 80th percentile of the probability distribution with a sentence in context: 80% of the time, at most 2.29 million viewers watched.
20% of the time, more than 2.29 million tuned in.
- 3b. Take a random episode. $P(\text{under one-and-a-half million viewers tune in}) = \underline{0.006979}$
- 3c. Kourtney reports to Kim that last night's episode had a z-score of 2.45. Help Kim solve for the number of viewers. Show work.

$$z = \frac{y - \mu}{\sigma}$$

$$2.45 = \frac{y - 2.09}{0.24}$$

$$y = 2.09 + 2.45(0.24)$$

$$y = 2.678 \text{ million viewers.}$$

- 3d. Take two random episodes. Determine:

$P(\text{First episode exceeds 2 million BUT Second episode under 2 million}) =$

$$P(X > 2) \times P(X < 2) =$$

$$0.6462 \times 0.3538 = 0.2286$$

4. The number of Kardashians who show up to Thanksgiving dinner is modeled with this table:

Number of Kardashians	4	5	6	7	8	9	10	11	12
Probability	0.20	0.20	0.10	0.10	0.10	0.10	0.10	0.05	0.05

- 4a. $P(\text{at least 9 Kardashians at Thanksgiving}) = \underline{0.30}$
- 4b. $P(\text{at most 6 Kardashians at Thanksgiving}) = \underline{0.50}$
- 4c. Expected number of Kardashians at Thanksgiving (show calculation):

$$\mu = 4(0.20) + 5(0.20) + \dots + 12(0.05)$$

$$\mu = 6.95 \text{ Kardashians}$$

11 zzz Retired

5. Estimate the $P(\text{Cecil Student drinks "Alcohol"})$ using the "Calendar Year 2017 Large Survey".

Answer: $\underline{402/507 = 0.7929}$

11 zzz Retired

6. If we trust the self-reporting and accuracy on the "Calendar Year 2017 Large Survey", it appears that approximately 73.31% of Cecil College students have a GPA exceeding 3.0.

We will take a new random sample of $n = 13$ students and count up the number with GPAs over 3.0.

Use the Binomial model to answer the following questions. $n = 13, p = 0.7331$

- 6a. Calculate the mean and standard deviation of this probability model. Show work:

$$\mu = np = 13(0.7331) = 9.5303$$
$$\sigma = \sqrt{np(1-p)} = 1.5949$$

- 6b. Using your answers from above, how many students out of 13 would need to have a GPA exceeding 3.0 for that to be unusually high? Circle one answer:

0 1 2 4 5 6 7 8 9 10 11 12 13

- 6c. $P(\text{Under 10 have a GPA exceeding 3.0}) = \underline{P(X < 10) = 0.4723}$

- 6d. $P(\text{Everybody has a GPA exceeding 3.0}) = \underline{P(X = 13) = 0.0177}$

7. Time between logins to www.reddit.com follows an Exponential model with a mean of 2 seconds.

- 7a. $P(\text{Next login within the next second}) = \underline{P(X \leq 1) = 0.3935}$

- 7b. $P(\text{Next login not for at least 10 seconds}) = \underline{P(X \geq 10) = 0.006738}$

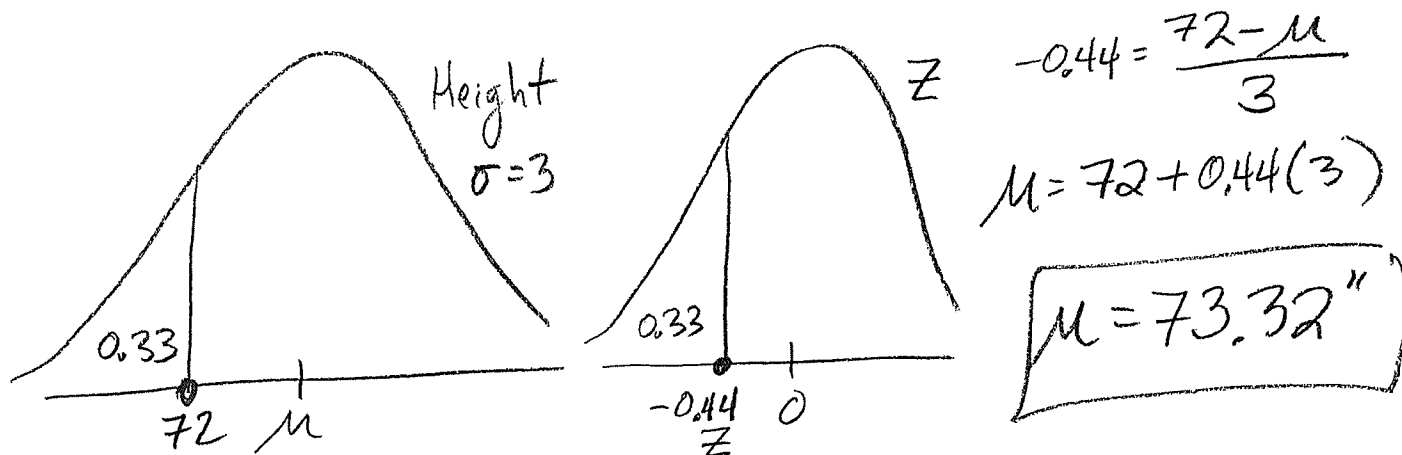
- 7c. 99.99th percentile = 18.42 seconds

- 7d. One amazing feature of an Exponential model is that the standard deviation is equal to the mean. Using that fact, what is an unusually long time in between logins to Reddit?

$$\mu + 2\sigma = 2 + 2(2) = \underline{\underline{6 \text{ seconds}}}$$

8. We have a Normal model with an unknown parameter. In America, males have an average height of 69.7 inches, but in Hollywood, it is believed males tend to be taller.

OK. So we survey male actors and it turns out that only 33% are under 72 inches tall. Presuming the standard deviation for male height is 3 inches, solve for the mean height of male actors in Hollywood. Show work.



9. We know the $P(\text{New marriage ends in divorce / annulment / separation within 10 years}) = 0.29$. We take five random, newlywed couples.

9a. $P(\text{all five couples don't make it}) = P(5 \text{ for } 5 \text{ end}) = (0.29)^5$

$= 0.002051$

9b. $P(\text{all five couples do make it}) = P(5 \text{ for } 5 \text{ last}) = (0.71)^5$

$= 0.1804$

9c. $P(\text{at least one couple doesn't make it}) = 1 - P(\text{all } 5 \text{ last})$

$= 1 - 0.1804 = 0.8196$

- 9d. How many couples do we expect to make it? (This is Binomial)

$\mu = np = 5(0.71) = 3.55$

couples

10. Use the "Skyscrapers in the U.S." dataset.

Run the linear regression, using "Number of Floors" to predict the "Height (meters)".

10a. (3) First things first – describe the relationship between "Number of Floors" and "Height".

Form - Linear

Direction - Positive

Strength - Strong with $r = 0.8281$

#4 Outliers – Omit for now, address at end of problem.

10b. (2) Second things second – Check for three conditions required to do a linear regression analysis:

#1 Linearity ✓

#2 Equal Spread ✓

#3 Outliers – Omit for now, address at end of problem.

10c. (1) Linear Equation:

$$\hat{\text{Height}} = 23.82 + 3.187(\text{Floors})$$

10d. (4) Interpret the slope with a sentence in context:

For each floor added, we expect Height to increase by 3.187 meters.

10e. (4) Interpret the value of S_e with a sentence in context:

On average, our predicted Heights are off by ~ 22.91 meters when using $X = \text{Floors}$.

10f. (4) Interpret the value of R^2 with a sentence in context:

68.57% of the variation in Height is explained by knowing $X = \text{Floors}$.
31.43% is still unexplained.

10g. (4) Give one solid reason as to why the y-intercept of 23.82 meters is just a point on the red line:

$X=0$ floors is nonsense. No data by
 $X=0$ floors. y-int of 23.82 is not reasonable
for these buildings.

10h. (4) A new 66-floor building is being built. Show the calculation for the predicted "**Height**":

$$\hat{\text{Height}} = 23.82 + 3.187(66) \\ \approx 234.18 \text{ meters.}$$

10i. (4) Interpret the residual for "**Leveque Tower**" with a sentence in context: L.T. is

4.32 meters shorter than expected for
its floors of $X=47$.

10j. (4) Here is a made up building (so don't waste time looking for it). "**Cecil College Luxury Dormitories**" with 22 floors and "**Height**" = 111 meters. Show the calculation to obtain the residual for this building.

$$\hat{\text{Height}} = 93.94 \text{ meters (predicted)} \\ e = \text{Height} - \hat{\text{Height}} = 111 - 93.94 = 17.06$$

10k. (2) How many ~~hurricanes~~ ^{buildings} have large positive residuals? 61

10l. (2) How many ~~hurricanes~~ have large negative residuals? 8

10m. (2) How many ~~hurricanes~~ have large Cook's Distances? 112 $\frac{4}{2093} = 0.0019111$

10n. (2) What is the minimum # of floors you are comfortable doing predictions for? 13

10o. (2) What is the maximum # of floors you are comfortable doing predictions for? 110