Name: _____ Key _____   **Math 127 – Test 2A – Summer 2015**

Oath: *"I will not discuss the exam contents with anyone until it is returned to me by my instructor".*

Sign Name: _____

The penalty for cheating on this exam is a final grade of F for Math 127.

Testing Center Staff Instructions

1.   **One sheet of handwritten or typed notes is OK.**

        Students may <u>not</u> use the "pink sheet" or any copied or scanned answer keys or Math 127 department documents.

2.   **Collect the sheet of notes and staple it to the test when submitted.**

3.   **Testing Center issued TI calculator is OK. Student calculators are OK pending Testing Center rules.**

4.   **www.statcrunch.com is required. All other webpages are prohibited.**

5.   **Test must be completed in one sitting, but it is untimed. Very short bathroom breaks are permitted.**

Student Instructions

1. You will need to use www.statcrunch.com. This is the **only** permitted webpage.

2. 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.

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

4. Attach your sheet of notes to this exam on the way out.

5. The answer key will post Wednesday night, July 1st at some point.

6. Exams graded and returned by Monday July 6th. Grades should post to BB earlier than that.

7. Enjoy the holiday weekend with no assignments except for getting those videos for Lesson 14 watched.

Good luck,

Professor Kupe

Question 1: Regression, 4 points each.

1.   Use the "**Stat II Wal-Mart Supermarket**" dataset to answer the following questions.

1a.   Describe the relationship between "**Walmart Price**" and "**Supermarket Price**", hitting all the ~~supermarket~~ important points and include a measure of strength in your write up. We are predicting "~~Walmart Price~~" for this entire problem. No need to compute Studentized residuals or Cook's distances at this point – it will come later.

     Positive, linear, Strong with r = 0.9028

     Outliers: Visually there are 5 higher-priced Products away from the cloud.

1b.   Determine the linear equation using StatCrunch. Explain why the y-intercept is meaningless in the context of this problem – do not just say "because it is meaningless" or "it is just a point on the line".

     Equation: $\widehat{Supermarket} = 0.19 + 1.13(Walmart)$

     y-intercept: X = $0 at Walmart is nonsense since nothing is free. Also lowest price at Walmart is X = 69¢, So X = $0 is extrapolation.

1c.   For each dollar Wal-Mart increases a price, we'd expect the supermarkets to increase them by $1.13

     Something that costs $3.99 at Wal-Mart is predicted to cost $4.71 at the supermarket.

     The Kellogg's Fruit Loops in row 11 is expected to cost $3.53 at the supermarket, based on the equation.

     The equal spread condition is clearly (met) not met. (Circle)

1d.   Interpret the value of R^2 with a sentence in context. 81.5% of the variation in supermarket prices can be explained by knowing Walmart's Price. 18.5% is unexplained.

3

1e. How many products have unusual Studentized residuals and state how large a Studentized residual must be to be classified as "unusual":

How Many: _4 total_
(all positive)

"Unusual" begins at: ± _2_

1f. Calculate the Cook's Distance that is "large".

Large Cook's Distance: _4/95 = 0.0421_   How many products exceed it: _3_

1g. Interpret, in context, the residual (not the Studentized residual) for The Cap'n Crunch Cereal in row 31.

$e = \$0.98$. Cap'n Crunch was $\$0.98$ more expensive at the supermarket than predicted for its $\$2.49$ Walmart price

1h. Suppose a 3-pack of Orbit gum costs $2.49 at Walmart and $3.19 at the supermarket. Show the calculation to arrive at the residual. (This product is not in the dataset, so don't search for it.)

$x = 2.49$
$y = 3.19$
$\hat{y} = 3.01$

$e = y - \hat{y}$

$e = 3.19 - 3.01$

$\boxed{e = \$0.18}$

1i. Interpret the estimate of error standard deviation with a sentence in context.

On average, our predicted supermarket prices are off by about $\$0.61$.

Questions 2 – 20: Probability, 2 points each.

2. A stat professor always ends class early. The number of minutes class lets out from the proper time of 7:50 pm follows Uniform [–20, –5] distribution. Here $X = -20$ corresponds to class getting out 20 minutes early and $X = -5$ corresponds to class getting out 5 minutes early. $X = 0$ does not exist because he never goes the proper time and X cannot be positive because he never goes over the allotted time.

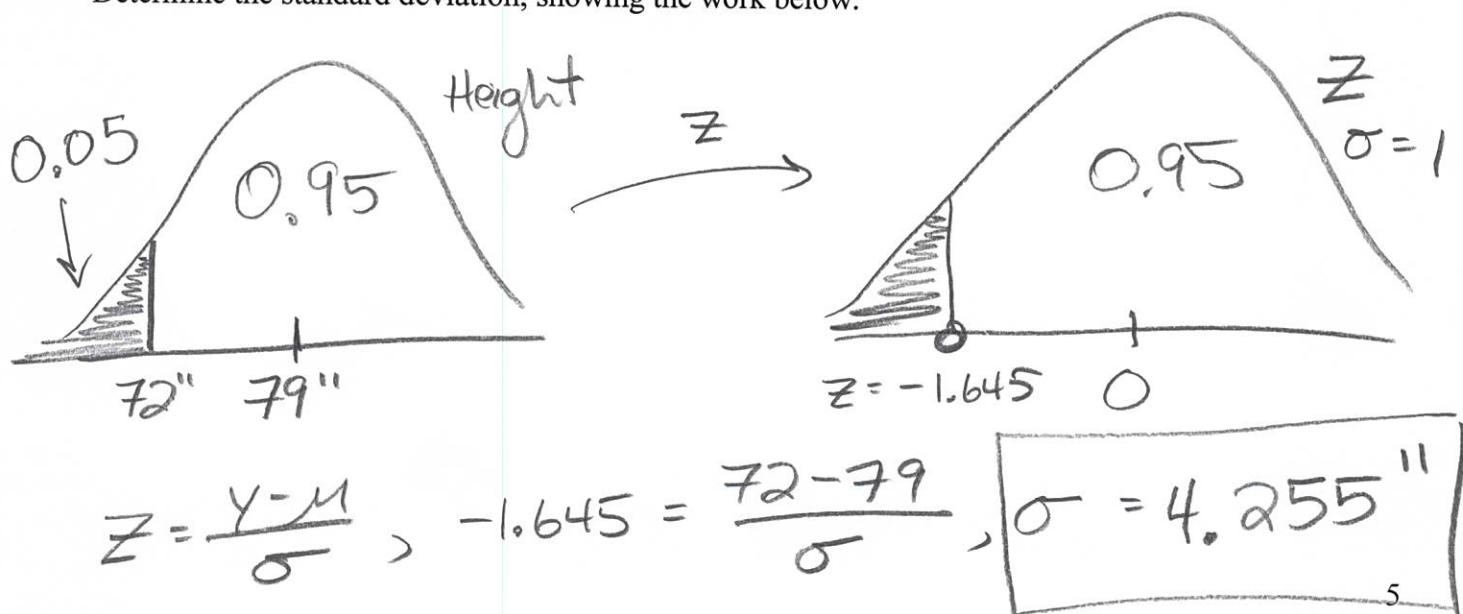2a. Draw the probability distribution graph and clearly state the probability function $f(x)$:

$$f(x) = \underline{\frac{1}{15}, \quad -20 \leq x \leq -5}$$
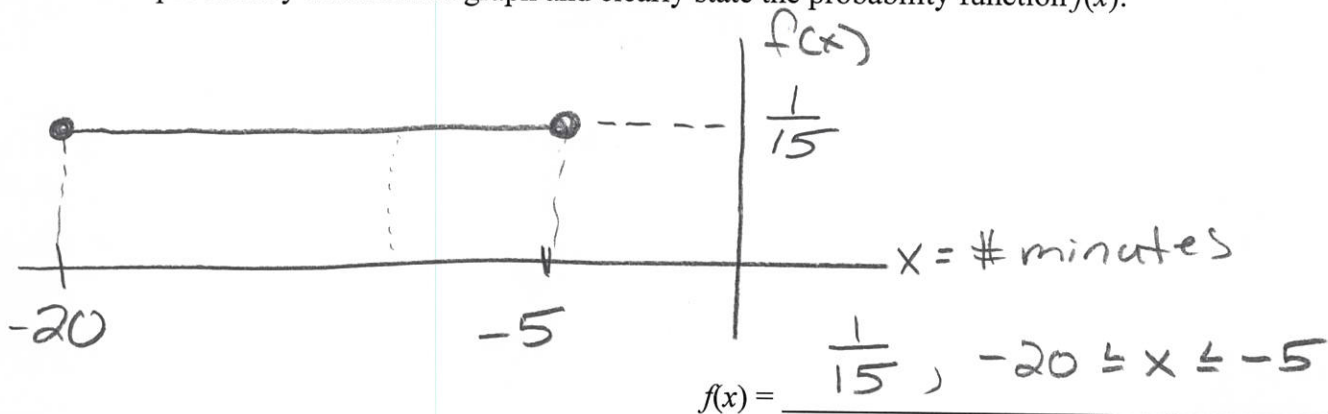
2b. P(Get out of class at least 10 minutes early) = $\underline{\dfrac{2}{3} \approx 0.667}$

2c. Expected number of minutes you will get out early:

$$\mu = \frac{a+b}{2} = \frac{-20 + (-5)}{2} = \frac{-25}{2} = -12.5$$

$$\boxed{12.5 \text{ minutes early}}$$

3. The height of NBA basketball players follows a Normal model with a mean of 79 inches, but suppose the standard deviation is unknown. We do know that 5% of all NBA players are under 6 feet tall. Determine the standard deviation, showing the work below.

$$Z = \frac{Y - \mu}{\sigma}, \quad -1.645 = \frac{72 - 79}{\sigma}, \quad \boxed{\sigma = 4.255''}$$

5

4. Life expectancy of a certain virus follows an Exponential distribution with a mean of 2 days.

4a. Determine the probability a virus lives at least a week. Draw the graph.

$f(x)$

$P(X \geq 7 \text{ Days}) = \underline{0.0302}$

Life Expectancy

2   7
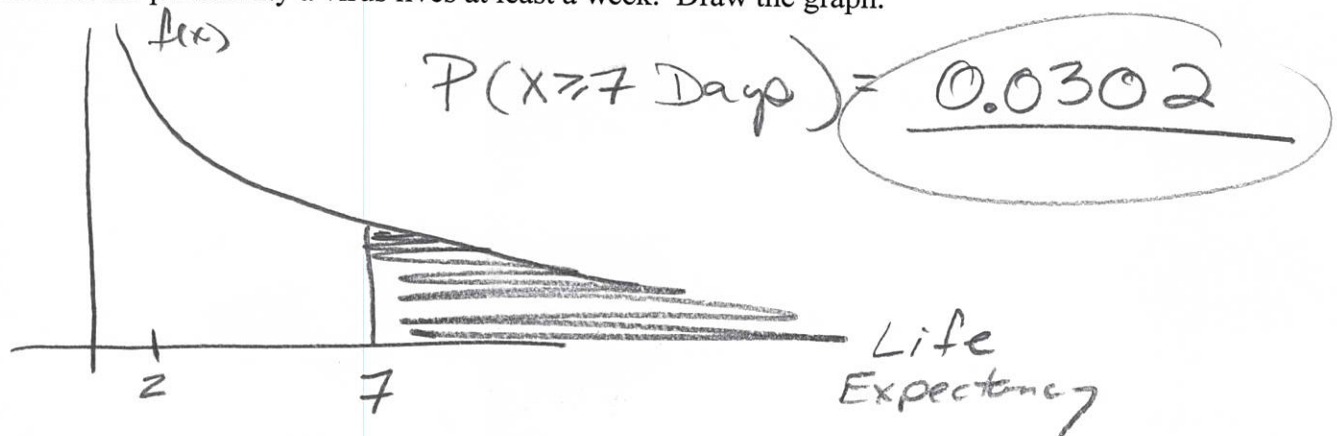
4b. Find P_{10} and P_{90}. Draw the graph with both together labeled.

0.10

0.10

0.80

$P_{10} = 0.211$ $P_{90} = 4.605$
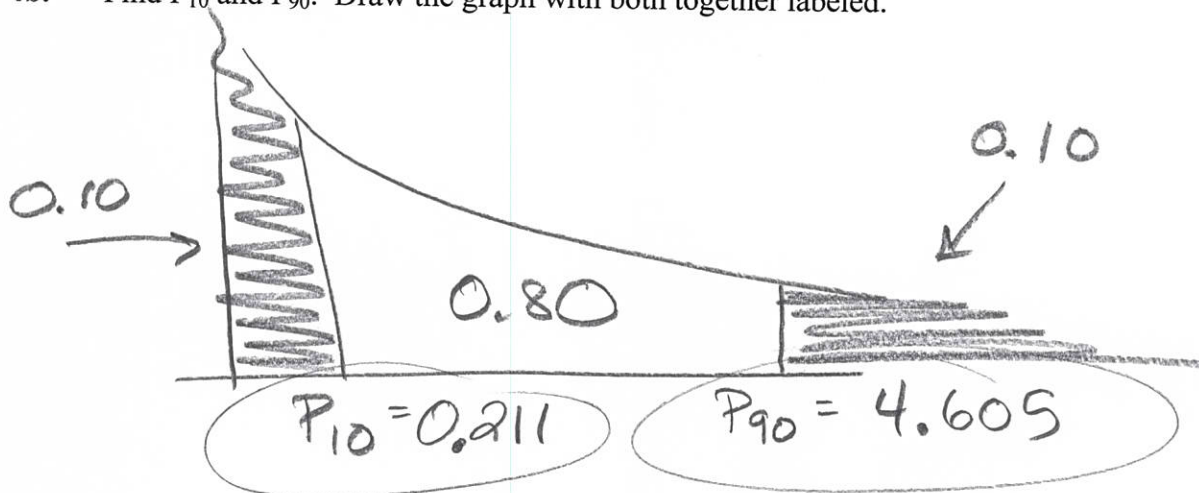
4c. Give a range of life expectancies that be the cutoffs for an unusually short or unusually long lifespan. The standard deviation of the probability model is ½.

$\mu \pm 2\sigma$

$2 \pm 2(\frac{1}{2})$

2 ± 1

1 to 3 days

OK too

$\mu \pm 3\sigma$

$2 \pm 3(\frac{1}{2})$

2 ± 1.5

0.5 to 3.5 days

5.   Below is the historical grade distribution for Math 127 students.

Grade	A	B	C	D	F
Probability	0.24	0.31	0.16	0.15	0.14

5a.   Determine the probability at least one student fails (Grade of F) in a class of 20 students. Show calculation.

$$P(\text{at least one fails}) =$$
$$1 - P(\text{all 20 pass}) =$$
$$1 - (0.86)^{20} = 0.9510$$

5b.   Three students are randomly selected – what's the chance all three earned Cs?

$$P(3 \text{ for } 3 \text{ get } Cs) = (0.16)^3$$
$$= 0.004096$$

6a.   Give the z-score that corresponds to the 1st percentile for a Standard Normal model: __-2.326__

6b.   Give the z-score that puts 8% of the area in the upper tail of Standard Normal model: __1.405__

7.   Suppose the starting salary for a bachelor's degree of chemical engineering is \$65,713 with a standard deviation of \$5,550 and that a Normal model applies. Professor Kupe negotiates a \$75,000 starting salary. Is that unusually good? Why or why not?

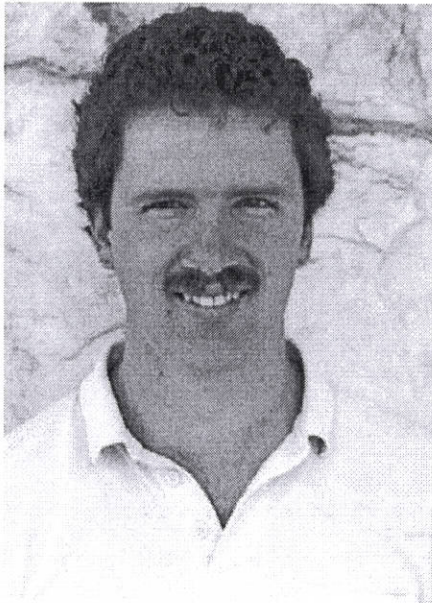Use z-scores: $Z = \dfrac{75000 - 65713}{5550} \approx 1.673$

(No)

Use probability: $P(X \geq 75,000) = 0.0471$

Not that unusual.

8. Number of strikes thrown in a game by Professor Kupe's favorite PBA bowler Walter Ray Williams, Jr. is given by the following discrete probability table:

Strikes	6	7	8	9	10	11	12
Probability	0.04	0.09	0.46	0.30	0.09	0.015	0.005

8a. *P*(At least 8 strikes) = ___ 0.87 ___

8b. *P*(At most 6 strikes) = ___ 0.04 ___

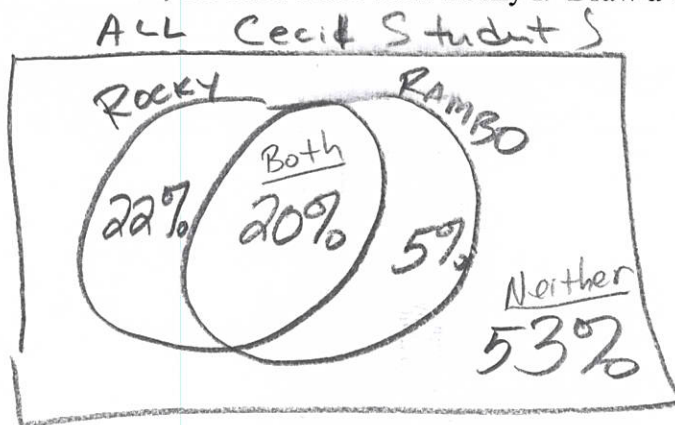**8c.** *P*(Two games in a row with at least 10 strikes):

$$P(2 \text{ for } 2, \; 10+ \text{ strikes}) = (0.11)^2$$
$$= 0.0121$$

8d. Expected number of strikes:

$$\mu = (6)(0.04) + (7)(0.09) + \cdots + (12)(0.005)$$
$$= 8.375 \text{ strikes}$$

8

9. Suppose that 42% of Cecil College students have seen Rocky I, and 25% have seen Rambo I. Of those who have seen Rambo I, 80% have also seen Rocky I. Draw a well-labeled Venn diagram.

ALL Cecil Students

ROCKY   RAMBO
        Both
22%   20%   5%
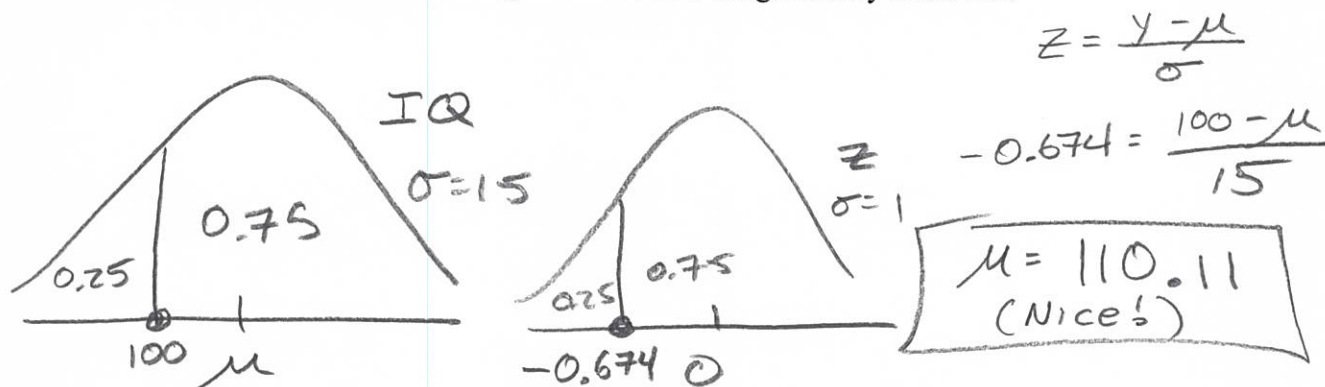   Neither
            53%

10. We have a Binomial probability model with $n = 15$ and $p = 0.50$. You could think of this as flipping a fair coin 15 times and counting up the number of heads.

10a. Determine the mean: $\mu = np = 15(0.5) = 7.5$

10b. Determine the standard deviation: $\sigma = \sqrt{np(1-p)} = \sqrt{15(0.5)(0.5)} \approx 1.936$

10c. P(All heads) $= (0.5)^{15} = 0.0000305$

11. Suppose faculty at Cecil College have IQ scores that follow a Normal model with an unknown mean and a standard deviation of 15 points. All 52 faculty members take an IQ test and 25% of them score 100 or below. Determine the mean IQ for all Cecil College faculty members.

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

IQ
$\sigma = 15$
0.75
0.25
100   μ

z
$\sigma = 1$
0.75
0.25
-0.674 0

$-0.674 = \dfrac{100 - \mu}{15}$

$\boxed{\mu = 110.11}$
(Nice!)

12. **(1 point)** If a randomization test for a statistically significant correlation were set up, what is the null hypothesis in all cases?

H_0: Two variables are not linearly related.

9