Stat 202
Summer 2015
Midterm Exam
7/16/15
Time Limit: 120 Minutes

Name (Print): Answers

This exam contains 6 pages (including this cover page) and 5 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

You may *not* use your books, or notes. You may use a computer—either your own or one from the lab—but only for StatCrunch, for accessing the data on my website, and for the calculator app. No other apps or websites may be open. You may *not* access the lecture notes or homework on my website. In addition to your computer, you may use a calculator, as long as it doesn't have internet. You may *not* use your cell phone during the exam.

You are required to show your work on each problem on this exam. The following rules apply:

- **Organize your work**, in a reasonably neat and coherent way, in the space provided.
- **Mysterious or unsupported answers will not receive full credit.** A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit. Graphing calculators should not be needed, but they can be used to check your work. If you use a graphing calculator to find an answer you must write the steps needed to find the answer, without the calculator.
- **For problems requiring StatCrunch, report all numbers you enter as input, otherwise you won't get partial credit for an incorrect answer.**
- If you need more space, use the back of the pages; clearly indicate when you have done this.
- Do not write in the table to the right.

Problem	Points	Score
1	20	
2	20	
3	20	
4	25	
5	15	
Total:	100	

1. (20 points) The results for the first two (hypothetical) statistics exams are in!

Name	Amy	Joe	Sue	Jan	Dan	Eva	Mia
1^{st} exam, S_1	98	85	62	84	73	98	83
2^{nd} exam, S_2	96	91	73	88	72	80	90

(a) (10 points) Find the correlation between S_1 and S_2.

Stat → Summary stats → Correlation
Pick columns S_1 and S_2
Compute

$r = 0.67537$

(b) (10 points) Write the equation for the regression line, showing one score as a function of the other. Use the first score as the explanatory variable and the second score as the response variable.

Stat → Regression → Simple linear

X variable = S1 (exam 1 score)

Y variable = S2 (exam 2 score)

$S_2 = 43.4198 + 0.4907\, S_1$

2. (20 points) A gumball machine contains a large number of candies. Of these 10 percent are red, 15 percent are purple, 30 percent are blue, and the rest are yellow. Also 40 percent are textured and the rest are smooth. Suppose that when a peice of candy is delivered, the machine stirs the candy in such a way that each piece has an equal chance of being served. Moreover suppose that the color delivered is independent of the texture delivered. What is the probability of getting either a red smooth candy or a yellow textured candy?

red	purple	blue	yellow	Sum
.1	.15	.3	?	1

$P(\text{yellow}) = 1 - .1 - .15 - .3 = .45$

$P(\text{smooth}) = .6 \qquad P(\text{textured}) = .4$

$P(\text{red and smooth}) = (.1)(.6)$ because independent

$= .06$

$P(\text{yellow and textured}) = (.45)(.4)$ because independent

$= .18$

$P((\text{red and smooth}) \text{ or } (\text{yellow and textured}))$

$= .06 + .18 = .24$

↑ because disjoint

3. (25 points) Simulate 1000 samples from the Normal Distribution with mean 1 and standard deviation 1. Use a fixed seed of 12. Call the column X. Create a new column X_{new}, using the following transformation:

$$X_{new} = \log(X^2)$$

(a) (5 points) What is the mean and standard deviation of the column X_{new}?

Use data → simulate → Normal { mean 1 std dev 1, rows 1000 cols 1, fixed seed 12
rename column to X
Use compute → expression
log(x^2)
Stat → summary stat → colums mean(-0.144902

(b) (5 points) Report the five-number summary of the data in X_{new}. std dev 2.18264

(min, Q_1, M, Q_3, max)
= (-12.5317, -1.269087, 0.05602, 1.10772, 2.7485

(c) (5 points) Report the interquartile range for X_{new}. What is the upper endpoint for range of diameters that will not be flagged as suspected outliers by the 1.5 × IQR rule?

$IQR = Q_3 - Q_1 = 2.286811$

upper fence (beyond which data flagged as potential outliers)

$Q_3 + 1.5\,IQR = 4.4479$

(d) (5 points) Are the X_{new} approximately normal? Justify your answer. What is the best (most sensitive) way of judging normality that we have learned in this class?

No, not normal

QQplot (most sensitive way learned to judge normality) is not a linear

Deviates substantially from linear

4. (25 points) Raw scores on a particular standardized test are approximately normal with mean 120 and the standard deviation 10.

(a) (5 points) What raw score is needed to be in the top 1 percentile?

Stat → calculators → Normal
mean 120; Std dev 10
$P(X \leq \_\_\_\_) = .99$ answer 143.263

(b) (5 points) On what percentile does a raw score of 150 fall?

$P(X \leq 150) = \_\_\_\_$

answer 99.865

(c) (5 points) What proportion of the population has a raw score between 85 and 115?

between tab

$P(85 \leq X \leq 115) = \_\_\_\_$

answer 0.308530

(d) (5 points) What z-score corresponds to a raw score of 100?

$$z = \frac{x-\mu}{\sigma} = \frac{100-120}{10} = -2$$

(e) (5 points) What raw score does a z-score of -1.2 correspond to?

$$-1.2 = \frac{x-\mu}{\sigma} = \frac{x-120}{10}$$

$$-12 = x - 120$$

$$x = 120 - 12 = 108$$

5. (15 points) Consider a 10-sided fair die, with faces labeled 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. Because the die is fair, each side has an equal probability of appearing when rolled.

   (a) (5 points) What is the probablity of 3 appearing at least twice in 6 successive rolls of the die?

Stat → Calculators → Binomial

$n=6 \quad p=.1$

$P(X \geq 2) =$ _____

answer = 0.114265

   (b) (5 points) What is the probability that 5 or 10 will appear at most twice is 4 successive rolls of the die?

$n=4 \quad p=.2$

$P(X \leq 2) =$ _____

answer = 0.9728

   (c) (5 points) What is the probablity that 3 will appear exactly once in 12 rolls of the die?

$n=12 \quad p=.1$

$P(X=1) =$ _____

answer = 0.37657