Stat 202
Fall 2014
Exam 2 Practice
10/30/14
Time Limit: 75 Minutes

Name (Print): Answers

This exam contains 7 pages (including this cover page) and 7 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, or cell phone. A calculator is OK as long as it has no internet. You may use the browser on the lab computer (but not your computer) to access StatCrunch. You may not visit any other websites. You will not need to download any data from my website or anywhere else. No other computer use is allowed.

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.
- 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	10	
2	10	
3	10	
4	10	
5	30	
6	10	
7	20	
Total:	100	

Formulas that may or may not be useful:

- If X is a discrete random variable that takes on values $x_1, x_2, \ldots x_n$ with respective probabilities $p_1, p_2, \ldots, p_n$, then its mean is given by:

$$\mu_X = \sum x_i p_i$$

- The mean of a linear transformation of a random variable X is given by the following (where a and b are numbers, not random variables):

$$\mu_{a+bX} = a + b\mu_X$$

- The mean of a sum of two random variables X and Y is given by:

$$\mu_{X+Y} = \mu_X + \mu_Y$$

- A more general formula that combines the two above; here $X_1, \ldots, X_n$ are random variables and a and $b_1, \ldots, b_n$ are numbers:

$$\mu_{a+b_1X_1+b_2X_2+\ldots b_nX_n} = a + b_1\mu_{X_1} + b_2\mu_{X_2} + \ldots b_n\mu_{X_n}$$

- If X is a discrete random variable that takes on values $x_1, x_2, \ldots x_n$ with respective probabilities $p_1, p_2, \ldots, p_n$, then its variance is given by:

$$\sigma_X^2 = \sum (x_i - \mu)^2 p_i$$

- The variance of a linear transformation of a random variable X is given by the following (where a and b are numbers, not random variables):

$$\sigma_{a+bX}^2 = b^2\sigma_X^2$$

- The mean of a sum of two independent random variables X and Y is given by:

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

- A more general formula that combines the two above; here $X_1, \ldots, X_n$ are independent random variables and a and $b_1, \ldots, b_n$ are numbers:

$$\sigma_{a+b_1X_1+b_2X_2+\ldots b_nX_n}^2 = b_1^2\sigma_{X_1}^2 + b_2^2\sigma_{X_2}^2 + \ldots b_n^2\sigma_{X_n}^2$$

- Central limit theorem: For a population with mean μ and standard deviation σ, and for samples chosen of size n, the distribution of the sample mean is approximately normal:

$$\bar{x}_n \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

1. (10 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	90	87	78	82	64	77	95
2^{nd} exam, S_2	95	90	72	90	80	85	85

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

Stat → Summary Stats → Correlation

Pick exam 1, exam 2

Correlation between Exam 1 and Exam 2 is:

0.53467095 or 0.535

(b) (10 points) Write the equation for the regression line, showing one score as a function of the other. What variable did you choose for the response variable and what variable did you choose for the explanatory variable? (Note, there is only one way to make these choices correctly.)

Stat → Regression → Simple linear

Explanatory Var → X Variable: Exam 1

Response Var → Y Variable: Exam 2

Exam 1 explains Exam 2

Intercept 52.8 Slope 0.397 $\boxed{E_2 = 0.397 E_1 + 52.8}$

2. (10 points) (a) (5 points) What is the correlation between random variables X and Y where the two are related by $Y - X = 4$?

Solve for Y → $Y = 4 + X$ positive slope, perfect linear relationship

Correlation = 1

(b) (5 points) What is the correlation between random variables X and Y where the two are related by $X = 3 - \frac{1}{2}Y$?

Negative slope, perfect linear relationship

Correlation = −1

3. (10 points) Seven dice are thrown together. What is the probability that at least one die shows the number 3?

Stat → Calculators → Binomial Calculator

$n = 7 \quad p = 0.166667 = 1/6$

$P(X \geq 1) = 0.721$

4. (10 points) Plastic chips are placed into an box. The chips come in two colors: red and blue. Also the chips come in two shapes: square and circle. Thus there are 4 types of chips in the box (red square, blue square, red circle, blue circle). One chip is selected at random. The numbers of each type of chip in the box are such that the probability of selecting a red chip is 0.2. Also the probability that the selecting a square chip is 0.4. Finally, the event "select a blue chip" is independent from the event "select a circle chip". What is the probablity of selecting a blue circle?

$P(red) = .2 \quad P(blue) = .8$

$P(square) = .4 \quad P(circle) = .6$

$P(blue \text{ and } circle) = P(blue)P(circle)$

(because independent

$= (.8)(.6)$

$= .48$

5. (30 points) A population of seventh graders and their parents consists of 30 children and 60 adults. The mean weight of the children is 100 pounds, standard deviation 5 pounds. The mean weight of the adults is 165 pounds, standard deviation 10 pounds.

Three random adults and four random (unrelated) children get on the elevator; one of the adults has a backpack that weighs 20 pounds. Let X_1, X_2, X_3 be the weights of the adults. Let Y_1, Y_2, Y_3, Y_4 be the weights of the children. Let W be the weight of the backpack (which is 20 pounds, not random). Assume all of these random variables are independent.

Let Z be the total load on the elevator (total weight of all 7 people plus backpack).

(a) (10 points) Write and equation for Z in terms of the X's, Y's and W.

$$Z = X_1 + X_2 + X_3 + Y_1 + Y_2 + Y_3 + Y_4 + W$$

(b) (10 points) Find the mean of Z as a number not a formula. Write a formula first then derive the number.

$$\mu_Z = \mu_{X_1} + \mu_{X_2} + \mu_{X_3} + \mu_{Y_1} + \mu_{Y_2} + \mu_{Y_3} + \mu_{Y_4} + \mu_W$$

$$= 165 + 165 + 165 + 100 + 100 + 100 + 100 + 20$$

$$= 915$$

(c) (10 points) Find the standard deviation of Z as a number not a formula. Write a formula first then derive the number.

$$\sigma_Z^2 = \sigma_{X_1}^2 + \sigma_{X_1}^2 + \sigma_{X_3}^2 + \sigma_{Y_1}^2 + \sigma_{Y_2}^2 + \sigma_{Y_3}^2 + \sigma_{Y_4}^2 + \sigma_W^2$$

$$= 10^2 \cdot 3 + 5^2 \cdot 4 + 0$$

$$= 300 + 100$$

$$= 400$$

$$\sigma_Z = 20$$

6. (10 points) A casino is considering how much to charge to play a game that pays the player \$10 with probablity 0.1, and \$100 with probability 0.01, and otherwise pays nothing. What is the minimum the casino should charge to expect to break even (on average, pay out equal to what they charge).

mean payout is

$(.89)(0) + (.1)(10) + (.01)(100) = 2$

must charge \$2 per game to break even

Payoff Values	0	10	100
Probability	.89	.1	.01

mean $= .89 \cdot 0 + .1 \cdot 10 + .01 \cdot 100 = \$2$

7. (20 points) In a population of 2500 individuals, the heights of the people in the population have mean 70 inches and standard deviation 4 inches. Eighty-one random samples of 25 are chosen. For each of the 81 samples chosen, a statistic is calculated: the mean of the height of the people in the sample (i.e. the sample mean).

   (a) (10 points) What is the mean of these 81 sample means?

   70 inches

   $\bar{X}$ an unbiased estimate of μ

   (b) (10 points) What is the standard deviation of these 81 sample means?

   $\frac{\sigma}{\sqrt{n}} = \frac{4}{5}$ inches

   Careful: n is sample size (25)
   not number of samples.