Math 263

Exam II

March 3rd 2015

The University of Arizona

Name: _____

Answers without adequate justification will not receive full credit, including multiple choice. Include units with your answer when appropriate, and box all answers unless an answer line is provided. By signing below I am agreeing to abide by the University of Arizona academic integrity policies and that all work done on this test is my own.

Signature: _____

Tips for Success:

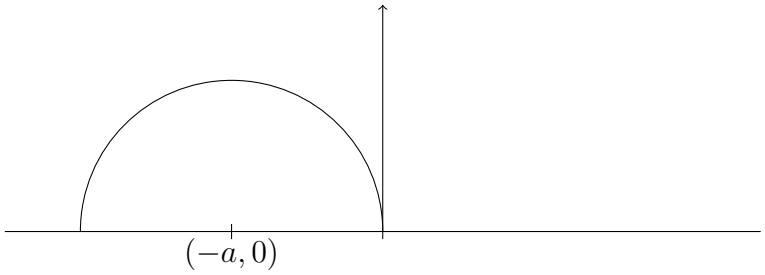
- Look through the entire test before starting to prioritize questions.
- If you get stuck on a question, move on and come back to it later.
- Do a quick reality check after each question: does my answer make sense? Did I include units? Did I show all my work?
- Read over the entire test at the end to make sure you didn't miss anything.
- For each question: take a deep breath, think slowly and deliberately at first, then work quickly once you see what to do.

Special notes for 263 exams:

- If you use a calculator to compute statistical quantities such as a 5 number summary, linear regression, normal proportion etc, please indicate which function you used i.e. 1-Var Stats, LinReg(ax+b), normalcdf().

- Round all numerical answers to 2 decimal places unless otherwise specified.

1. The graph below shows the probability density function of a continuous random variable X. The graph is a semi-circle with radius a.



   (a) What must a be in order for this to be a valid PDF? Leave your answer exact.

   The total area under the curve must be 1. The area of a semi-circle of radius a is $\pi a^2/2$, so we must have $a = \sqrt{2/\pi}$.

   (b) What is $P(-a < X < 0)$? Leave your answer exact.

   This region between $X = -a$ and $X = 0$ between the x-axis and the curve will have area exactly $1/2$, since it is half the probability distribution.

2. Suppose that a scratch ticket claims that you have a 1 in 10 chance of winning. What is a valid interpretation of this?

   ___ If you play the game 10 times, you will win exactly once

   ___ If you play the game 1000, you will win approximately 100 times

   ___ If you play the game once, you will win 1/10 of the maximum amount

   ___ If you play the game 10 times, you will win \$1.

   If you play 1000 times, you will win approximately 100 times.

3. Suppose we roll a 6-sided die and flip a coin.

   (a) List all the outcomes of this experiment

   The sample space is

   $$S = \{1H, 2H, 3H, 4H, 5H, 6H, 1T, 2T, 3T, 4T, 5T, 6T\}$$

(b) What is the probability that either the coin shows tails or the die is *not* a 4?

Let A be the event that the coin shows tails and let B be the event that the die is not a 4. Then,

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = \frac{1}{2} + \frac{10}{12} - \frac{5}{12} = \frac{11}{12}$$

(c) Are the following events disjoint? A = the die shows a 3, 4 or 5 and B = the coin is heads.

No, because for instance the outcome 3H lies in both events.

(d) What is the probability that the die shows a 3 *given* that the coin shows heads?

Let A be the event that the die shows a 3 and let B be the event that the coin shows heads. By the definition of conditional probability:

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{1/12}{6/12} = \frac{1}{6}$$

4. Suppose that the heights of every student on the UA campus are normally distributed. For the males, the mean is 68 inches and the standard deviation is 5 inches. For the females, the mean is 63 inches and the standard deviation is 4 inches.

(a) What is the probability that a randomly selected male student has a height greater than 62 inches? Hint: z-scores.

We can either convert to z-scores and use the z-table, or simply use the normalcdf() function on your calculator:

$$\text{normalcdf}(62,1000,68,5) \approx 0.885$$

(b) What is the probability that a randomly selected female has a height between 48 and 53 inches?

Again, we can use normalcdf:

$$\text{normalcdf}(48,53,63,4) \approx 0.00612$$

(c) What is the mean and standard deviation of the *sum* of the heights of males and females? I.e. if X is the height of a male, Y is the height of a female, then find the mean and standard deviation of $Z = X + Y$.

We can use the following facts:

$$E[X+Y] = E[X] + E[Y]$$
$$\sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y$$

(The second holds if we assume the heights of males and females are independent). Thus we have the mean of $X + Y$ is

$$\mu_Z = 68 + 63 = 131 \text{ inches}$$
$$\sigma^2_Z = \sigma^2_X + \sigma^2_Y = 25 + 16 = 41 \Rightarrow \sigma_Z = \sqrt{41} \approx 6.4 \text{ inches}$$

5. Consider a game where you roll two 4-sided dice. Each die is labeled with the numbers 1 to 4. Define a random variable X to be the sum of the two numbers on the dice.

X	2	3	4	5	6	7	8	
$P(X = x)$	1/16	1/8	3/16	1/4	3/16	1/8	1/16	

(a) Using the first row in the table above, list all the possible scores in this game. *Note: I may or may not have provided more boxes than necessary!*

See table

(b) Fill in the second row with the *probability* of getting each score. Hint: list all the possible pairs of rolls and compute the score for each one.

See table

(c) Compute the *expected value* of X, your score in the game. Show your work/formula used.

The formula for expected value:

$$\mu_X = \sum_{i=1}^{7} p_i x_i = 2 \cdot 1/16 + 3 \cdot 1/8 + 4 \cdot 3/16 + 5 \cdot 1/4 + 6 \cdot 3/16 + 7 \cdot 1/7 + 8 \cdot 1/16$$
$$= 5$$

(d) Compute the *variance* of X. If you were unable to find the expected value, you can use the symbol μ in its place and leave your formula unsimplified.

Use the formula for variance:

$$\sigma_X^2 = \sum_{i=1}^{7} p_i(x_i - \mu)^2 = (2-5)^2/16 + (3-5)^2/8 + \ldots + (8-5)^2/16 = 2.5$$

Then, we take the square root to get standard deviation:

$$\sigma_X = \sqrt{2.5} \approx 1.58$$

6. In an experiment on a new drug, subjects were randomly assigned to either a placebo or the active drug. In addition, the method of delivery of the drug (pill, skin patch, or nasal mist) was considered. In this experiment, there were how many factors?

There is only one explanatory variable (aka factor) here - the drug delivery mechanism. We can think of placebo as being a kind of drug delivery mechanism (i.e. they receive no drug).

7. A researcher is interested in the cholesterol levels of adults in the city in which she lives. A cholesterol screening program is set up in the downtown area during the lunch hour. Individuals can walk in and have their cholesterol determined for free. The service is used by 173 people, and their average cholesterol is 217.8. Mention some possible issues with this sampling procedure.

The sampling procedure is biased because it only samples from individuals who happen to be downtown. People who spend more time in their homes, or in different parts of the city, might have different cholesterol levels.

8. Consider a study performed by a medical center to determine which of two heart surgeries is most effective: angioplasty (running plastic tubes through the arteries) or bypass (rerouting arteries). The purpose of either procedure is to prolong the life of the patient. The study records the survival time of each patient (measured from the time of the surgery). What are the response and explanatory variables in this study?

The response variable is survival time and the explanatory variable is the surgery type.