Mathematics 1372/D552, Spring 2016 **Exam #2**
Instructor: Suman Ganguli **April 20, 2016** **Name:** ______________________

Question	Points	Score
1	20	
2	20	
3	20	
4	20	
5	20	
Total:	100	

Formulas:

- Addition Rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Conditional probability:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Multiplication rule:

$$P(A \cap B) = P(A)P(B|A)$$

- Binomial random variables:

$$P(X = i) = \binom{n}{i} p^i q^{n-i}$$

$$E[X] = np$$

$$\text{Var}(X) = npq$$

$$\text{SD}(X) = \sqrt{npq}$$

1. (20 points) Suppose a store estimates that the probability of a randomly selected customer buying item A is 0.1, and the probability of a customer buying item B is 0.15. But if a given customer has bought item A, then the probability of them buying item B increases to 0.4, i.e.

$$P(A) = 0.1, P(B) = 0.15 \text{ and } P(B|A) = 0.4$$

(a) How would the business estimate these probabilities? Why might they be interested in them, i.e., how might they make use of such information? (Hint: Amazon.com does makes use of this kind of information.)

Solution: Businesses can estimate these sorts of probabilities from existing customer data; they can use it to make recommendations and cross-promotions, e.g., direct ads about B to customers who have bought A

(b) What is the probability that a randomly selected customer buys item A *and* item B, i.e., what is $P(A \cap B)$?

Solution: By the Multiplication Rule:

$$P(A \text{ and } B) = P(B|A) \cdot P(A) = (0.4)(0.1) = 0.04$$

(c) What is the probability that a randomly selected customer buys item A *or* item B, i.e., what is $P(A \cup B)$?

Solution: By the Addition Rule:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) = 0.1 + 0.15 - 0.04 = 0.21$$

(d) What is the probability that a randomly selected customer buys item A given that he/she has bought item B?

Solution: We need to compute the conditional probability $P(A|B)$:

$$P(A|B) = P(A \text{ and } B)/P(B) = 0.04/0.15 = 0.267$$

2. (20 points) A bag contains 9 red marbles and 1 blue marble. Two marbles are selected from the bag *without replacement*. Let X denote the number of red marbles selected and let Y denote the number of blue marbles selected.

(a) What are the possible values of each random variable?

possible values of X:

possible values of Y:

Solution: $X : \{1, 2\}, Y : \{0, 1\}$

(b) Compute the following probabilities:

(i) that two red marbles are selected:

Solution:

$$\frac{9}{10}\frac{8}{9} = \frac{8}{10} = 0.8$$

(ii) that a red marble is selected first and then the blue marble is selected second:

Solution:

$$\frac{9}{10}\frac{1}{9} = \frac{1}{10} = 0.1$$

(iii) that the blue marble is selected first and then a red marble is selected second:

Solution:

$$\frac{1}{10}\frac{9}{9} = \frac{1}{10} = 0.1$$

(c) Use parts (a) and (b) to write down the probability distributions of X and Y.

Solution:

i	$P(X = i)$
1	0.2
2	0.8

i	$P(Y = i)$
0	0.8
1	0.2

(d) Calculate the expected values of X and Y. Show your calculations:

$E[X] =$

$E[Y] =$

Solution: $E[X] = 1(0.2) + 2(0.8) = 0.2 + 1.6 = 1.8$
$E[Y] = 0(0.8) + 1(0.2) = 0.2$

3. (20 points) Calculate the following. You can use a spreadsheet to carry out and/or check your calculations, but also write out all your calculations below:

   (a) Suppose you have to pick a 4-digit access code for a security system. Each digit can be any number from 0 through 9, except the first digit cannot be 0. How many different access codes are there?

   Solution: $9 * 10 * 10 * 10 = 9,000$

   (b) Consider a basketball team that has 15 players on its roster. The coach has to choose 5 players from the roster to play at any given time. How many different 5-player lineups does the coach have to choose from?

   Solution: $\binom{15}{5} = \frac{15*14*13*12*11}{5*4*3*2*1} = \frac{360,360}{120} = 3,003$

   (c) A club consisting of 12 members must select a president, vice-president and treasurer. In how many different ways can the positions be filled?

   Solution: $12 * 11 * 10 = 1,320$

   (d) Suppose restaurant has 12 different main dishes, 10 side dishes, and 6 different desserts. For a meal, you must choose one main dish, two different side dishes, and one dessert. How many different meals are possible?

   Solution: $12 * \binom{10}{2} * 6 = 12 * \frac{10*9}{2} * 6 = 12 * 45 * 6 = 3,240$

4. (20 points) Cancer survival statistics are often quoted in terms of 5-year survival rates, which gives the percentage of patients alive 5 years after the diagnosis of their cancer.

   For example, according to the NIH the 5-year survival rate for all childhood cancers combined is 80% (via `http://report.nih.gov/nihfactsheets/viewfactsheet.aspx?csid=75`).

   A study will track 50 children diagnosed with cancer. Let the random variable X represent how many of the 50 children will be alive 5 years after their diagnosis.

   (a) If we interpret the 5-year survival rate as the probability that a child diagnosed with cancer will be alive 5 years after diagnosis, explain why this is an example of a binomial experiment. What are the values of n and p?

   Solution: This is a binomial experiment since there are two outcomes (survival or non-survival) for each of the 50 children ($n = 50$), the outcomes for each child is independent, and we assume the survival rate is the probability of survival for each child ($p = 0.8$).

   (b) Calculate the expected value and the standard deviation of X.

   $E[X] =$

   $\mathrm{SD}(X) =$

   Solution: $E(x) = np = 50(0.8) = 40$
   $\mathrm{SD}(X) = \sqrt{npq} = \sqrt{50(0.8)(0.2)} = \sqrt{8} \approx 2.83$

   (c) Calculate the following probabilities using the binomial probability formula; show all your calculations. (You can check your answers using the spreadsheet function `=binomdist(i,n,p,false)`.)

   (i) exactly 5 of the children will survive, i.e., calculate $P(X = 5)$:

   Solution:

   $$P(X = 5) = \binom{50}{5} * (0.8)^5 * (0.2)^{45} = (2,118,760)(0.32678)(3.52 \times 10^{-32}) = 2.44 \times 10^{-26}$$

   (ii) exactly 40 of the children will survive, i.e., calculate $P(X = 40)$:

   Solution:

   $$P(X = 40) = \binom{50}{40} * (0.8)^{40} * (0.2)^{10} = (10,272,278,170)(0.000132923)(1.02 \times 10^{-7}) = 0.139819$$

   (d) Use the spreadsheet function `=binomdist(i,n,p,true)` to calculate the probability that 40 or fewer children will survive, i.e., calculate $P(X \leq 40)$. Write down the exact spreadsheet command you use.

   Solution:

   `=binomdist(40, 50, 0.8, true)` $\Longrightarrow P(X \leq 40) \approx 0.5563$

5. (20 points) Advanced Placement exams are scored on a 5-point scale, where 1 is the lowest possible score and 5 is the highest possible. Two students Alice and Bob are preparing to take the statistics AP exam. Alice and Bob come up with subjective probability distributions to represent how they think they will do on the exam.

Alice's subjective probability distribution for her exam score A:

i:	1	2	3	4	5
$P(A = i)$:	0.05	?	0.15	0.4	0.3

Bob's subjective probability distribution for his exam score B:

i:	1	2	3	4	5
$P(B = i)$:	?	0.45	0.3	0.15	0.05

(a) Calculate the missing values in the distributions:

(i) $P(A = 2) =$

Solution: $P(A = 2) = 1 - (0.05 + 0.15 + 0.4 + 0.3) = 1 - 0.9 = 0.1$

(ii) $P(B = 1) =$

Solution: $P(B = 1) = 1 - (0.45 + 0.3 + 0.15 + 0.05) = 1 - 0.95 = 0.05$

(b) According to Alice's probability distribution, what is the probability that she will get a score of 3 of higher? What about for Bob?

$P(A \geq 3) =$

$P(B \geq 3) =$

Solution: $P(A \geq 3) = 0.15 + 0.4 + 0.3 = 0.85$; $P(B \geq 3) = 0.3 + 0.15 + 0.05 = 0.5$

(c) Find the expected value of each probability distribution.

(i) $E[A] =$

Solution: $E[A] = 1 * (0.05) + 2 * (0.1) + 3 * (0.15) + 4 * (0.4) + 5 * (0.3) = 0.05 + 0.2 + 0.45 + 1.6 + 1.5 = 3.8$

(ii) $E[B] =$

Solution: $E[B] = 1 * (0.05) + 2 * (0.45) + 3 * (0.3) + 4 * (0.15) + 5 * (0.05) = 0.05 + 0.9 + 0.9 + 0.6 + 0.25 = 2.7$