Exercise 4

May 6, 2018

* This assignment can't be submitted in pairs - **individuals** only.

Part 1 - Probability Questions

For each one of the following question please supply your calculations along side your answers.

1. There are 2 boxes. The first has 3 green balls and 9 red balls.
   The second one has 5 green balls and 7 red balls.
   One ball is transfered from the first box to the second box.
   Now, A ball is drawn from the second box.
   Given that the drawn ball is green, find the probability that the transferred ball is red.

2. You and a friend are going to Shilat tomorrow for a day of climbing. In recent years, the conditions were good for climbing only 60 days each year. Dani Rup (the weatherman) has predicted good climbing conditions for tomorrow. When the climbing conditions are actually good, Dani correctly forecasts the conditions in 85% of the time. When the climbing conditions aren't good, he incorrectly forecasts good conditions 10% of the time. What is the probability that the climbing condition will be good tomorrow?

3. You toss 3 fair dice independently. Every die has 4 possible results (1, 2, 3, 4). Let A denote the event that the sum of the dice is 7. Let B denote the event that the sum of the dice is divisible by 3. Let X be a random variable denoting the outcome of the toss of the first die.

   (a) Compute $P(A|X = i)$ for each $i = \{1, 2, 3, 4\}$

   (b) Compute $P(A)$

   (c) Are the event A and $X = i$ independent for each $i = \{1, 2, 3, 4\}$

   (d) Compute $P(B|X = i)$ for each $i = \{1, 2, 3, 4\}$

   (e) Compute $P(B)$

   (f) Are the event B and $X = i$ independent for each $i = \{1, 2, 3, 4\}$

4. There a 1000 dice, which look identical. However, 999 of them are "fair" (i.e. when rolling the die, $P(X = k) = \frac{1}{6}$) and one of them is forged (specifically when rolling that die the probability of getting 1 is 0.9 and 0.02 for the rest).
   Assuming we selected a die randomly, rolled it ten times, and saw 8 "1" and 2 "6" (you don't know the sequence of events), what is the probability we choose the forged die?

5. In an ancient tribe, the chief decided he does not allow a family to have more than 1 boy as this will threaten his position. Thus, each family must stop child birth once a boy is born. As a result, families in this country are of the following type:

   - Boy

- Girl, Boy
- Girl, Girl, Boy
- Girl, Girl, Girl, Boy
- Girl, Girl, ... , Girl , Boy
- And so on...

Assuming the chance of having a boy is $\frac{1}{3}$, what is the expected value of the number of girls in a family in this tribe.

Part 2 - Naive Bayes Classifier

1. In a far away land where the climate is very strange, two researchers, Alice and Bob, are looking to film an exotic creature, the Randomamml.
   Since they are both lazy, and they know the probability of spotting a randommal is 0.2, they only go out of the camp when they think there is a good probability of finding a Randomamml. From previous observation they know something about the temperature and humidity in the days the Randomamml was spotted and the days he was not.
   When **spotted** they know :

   - Humidity $\sim N(0, 1)$
   - Temperature $\sim N(0, 1)$

   And when **not spotted** they know:

   - Humidity $\sim N(0.2, 1)$
   - Temperature $\sim N(0.2, 1)$

   This information is sufficient for naive Bob, but Alice wants to be more accurate so she searched for more data and managed to infer that tempartute and humidty follow a bivariate normal distribution in each one of the situations and further more evaluated the covariance of the humidity and temperature for each of the situations.
   When **spotted**:

   - $cov(humdity, temp)$ = 0.8

   When **not spotted**:

   - $cov(humdity, temp)$ = -0.8

   The forecast for tomorrow is :

   (a) temperature = 1, humidity = 1

   We know that the Randommal is coming out tomorrow (because we can see into the future). For both Bob(using a Naive Bayes classifier) and Alice(using full Bayes classifier with her inferred density function) say whether they manage to spot the Randommal or not.

   As a reminder here is the bivariate normal distribution density function:

   $$p(\bar{x}|A_i) = \frac{1}{2\pi \cdot \sqrt{|S|}} \cdot e^{-\frac{1}{2}\cdot(\bar{x}-\bar{\mu_i})^T \cdot S^{-1} \cdot (\bar{x}-\bar{\mu_i})}$$

   Where :

   - $|S|$ = the determinant of matrix S.
   - S^{-1} = The inverse of matrix S
   - T is the transpose operator.

   * Along side your answer provide full descriptive calculations of how you got your results.
   ** For the inverse matrix calculations you can use the following website: http://matrix.reshish.com/inverse.php.

2. The owner of the famous Randomistan restaurant comes to seek your help. After 10 themed nights he gathered some data, and wants to prepare dishes that are aligned with that data. He has the following information about the dishes: At what night they were served, were they spicy, the cooking method and whether they were returned or not. He seeks your advice with the following 2 dishes:

(a) A normal grilled dish in a mexican themed night.

(b) A spicy steamed dish in an indian themed night.

Here is the gathered dataset:

Theme	**Spiced**	**Cooking**	**Returned**
Mexican	Spicy	Fried	False
Mexican	Noraml	Grilled	True
Mexican	Spicy	Grilled	False
Mexican	Noraml	Fried	True
Mexican	Spicy	Grilled	True
Indian	Spicy	Fried	True
Indain	Noraml	Steamed	False
Indain	Spicy	Fried	False
Indain	Noraml	Grilled	False
Indain	Noraml	Grilled	False

Using a Naive Bayes Classifier with Laplace smoothing help the owner decide whether he should serve those dishes or not.
* Along side your answer provide full descriptive calculations of how you got your results.