DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, IIT ROPAR

# Tutorial-1

CSL-471, Probability and Computing

Course Instructor- Dr. Sudarshan Iyengar

August 5, 2016

## 1 INSTRUCTIONS

1. This tutorial sheet consists of two parts- i) Instructor picked questions. ii) TA picked questions. You are supposed to attempt both of the parts.
2. Every question is marked with a rating- 1*, 2*, 3* or exploratory. * levels indicate the degree of complexity of the question, with 1* questions being the easiest and 3* questions demanding good time of intellectual thinking. Exploratory questions may take any number of days or weeks to be thoroughly attempted.
3. The tutorial is to be discussed in the class on August 11, 17:10- 18:00 hrs.
4. In case of any doubt or query regarding any question, please open a thread in the google group CSL-471 and discuss it.
5. Please do not get confused. The rating for every question has been mentioned after the question.

## 2 INSTRUCTOR PICKED QUESTIONS

1. Obtain the closed form for the following expression

$$1 + \frac{1}{2} + \frac{1}{3} + ..... + \frac{1}{n}$$

(Rating : 3*)

2. Assume an online hiring company is on a venture of choosing a secretary for the Managing Director of the company. There are $n$ number of applicants. The company follows the following procedure

   - Pick the first candidate from the applicants' queue (Please note that the applicants are randomly placed in the queue without any sort of criterion).
   - Hire this candidate. Then, the company looks at the next applicant in the queue. If the next applicant is inferior than the currently hired secretary, she is ignored. But if the next applicant is better than the current secretary, she is hired and the current employee is fired.
   - In the similar fashion, the company traverses through the entire queue firing the current secretary on seeing a better option later in the queue (Do you notice that in this scenario, the company will end up hiring the best candidate at the last).

   Find out the expected number of firings this algorithm leads to.

   (Rating : 2*)

3. We discussed the German tank problem in the class. We revisit the problem here. During the world war second, the western allies wanted to estimate the number of tanks possessed by the German army. Each of the tank produced by the Germans was marked with a number. The numbers were all consecutive, starting from 1. The alias could see some of the German tanks passing by and hence some of these numbers. The problem was to estimate the number of actual tanks Germans possessed by looking at the sample of the passing by tanks.

   The question can simply be stated as: Martha has $n$ numbers, She gives you $k$ numbers out of these uniformly at random. By looking at these $k$ numbers, can you find the expected value of $n$.

   We also discussed a simple solution to this problem in the class. If we pick the maximum number out of these $k$ randomly picked number say $a_k$, we have proved that $E[a_k] = \frac{k}{k+1}(n+1)$.

Can you now go ahead and find the standard deviation of $a_k$, i.e. $\sigma(a_k)$ ?

(Rating : Exploratory)

4. Recall the online hiring/ dating problem. The solution discussed in the class asks one to look at the first $k$ candidates, reject all of them, while keeping a note of the best one among them. Then, the company starts tracing from the $k+1$th candidate onwards. As soon as we find a person better than the best person seen among first $k$ candidates, we hire this better person and terminate the algorithm. We proved that finding the best candidate is likely if we choose $k = \frac{n}{e}$. This approach expects one to know the number of candidates $n$ beforehand. Give an algorithm to execute online hiring if the number of candidates $n$ is not known in advance.

(Rating : 1*)

5. We are aware with the Monty Hall problem we discussed in the first lecture of the course. In the original problem, there were 2 goats and 1 BMW. Assume now, there are $k$ doors hiding goats and 1 door having the BMW car. We use the same strategy to play, i.e. swapping after seeing a door with a goat. What is the probability of winning the BMW here?

(Rating : 1*)

6. Consider a language having a large number of alphabets (By large number, we mean some very big but finite number). Say $L = a_1, a_2, a_3, .....$ Let probability of occurrence of an alphabet $a_i$, $pr(occurrence\ of\ a_i) = \frac{1}{2^i}$. If we randomly pick two alphabets from a very big book of this language (obeying the above mentioned frequency distribution), what is the probability that both the picked alphabets are the same (In other words, what is the probability of a collision)?

(Rating : 2*)

7. This question is a very simple variant of the above question. If instead of two alphabets, we pick 3 alphabets uniformly at random, what is the probability that all these three alphabets are the same(again means collision only but between 3 alphabets) ?

(Rating : 1*)

8. Given a binary string of length $n$, what is the expected length of the longest streak of 0s that one can see?

(Rating : 3*)

9. The dual dating problem is a very interesting variant of the dating problem. Assume that both the boys and girls are given an equal chance to go ahead and make their choices, then the algorithm proceeds as follows.
   - Every girl and every boy goes ahead and checks out random $\frac{n}{e}$ boys and $\frac{n}{e}$ girls respectively and makes a note of the best person they have seen.
   - After looking at $\frac{n}{e}$ persons, they go on and keep looking at more random people (Not to mention, boys look for girls and vice versa). As soon as a person encounters a better person than their noted one, they select the person. But two persons settle down only if they both select each other.

   Let the best person in the boys be called the king and the best girl is called the queen. What is the probability that the king and queen
   a) Meet each other during the course of the game?
   b) Settle down with each other (marry)?

   (Rating : 2*)

10. Show that the congruence $x_i \equiv (ax_i - 1) mod n$ generates all numbers from 0 to $n-1$, given that $a$ and $n$ are relatively prime.

    (Rating : 3*)

11. In the one time pad algorithm we discussed in the class, the key length was considered equal to the length of the plain text. Assume that the key length is $n/2$ where $n$ is the length of the plaintext. We can replicate the key two times in order to perform ciphering. Comment on whether this encryption is perfectly secure or not ?

    (Rating : 2*)

12. Consider the k-permutation cipher. The cipher is very simple. Given a plain text, $p_1, p_2, p_3, p_4, p_5, p_6, ....., p_n$ of length $n$. We divide the plain text in the blocks of size $k$. While encrypting, every block is permuted randomly.

    Assume $k = 4$ in the below example.

    Plain text : $p_1, p_2, p_3, p_4, |p_5, p_6, p_7, p_8, |.....|p_{n-3}, p_{n-2}, p_{n-1}, p_n$

    Cipher text: $p_3, p_1, p_4, p_2, |p_8, p_7, p_5, p_6, |.....|p_{n-1}, pn-3, p_{n-2}, p_n$

    Prove that this cipher is not perfectly secure.

    (Rating : 2*)

## 3 TA PICKED QUESTIONS

1. Consider a game, where you throw a fair die. If you get a number from 1 to 5, you get the number of $ = The number that comes on the die. If die shows 6, you get 6 more $ and end the game. What is the expected amount of money you win? Also find the standard deviation.

   (Rating : 2*)

2. A die is rolled and a coin is tossed alternately. If the coin shows head, the die throw continues else stops. What is the expected sum of the numbers which have appeared on die throughout the game? Also find the standard deviation. Assume the die as well as the coin to be unbiased.

   (Rating : 2*)

3. Consider Caesar cipher. We know that it is breakable in a maximum of 26 attempts. Can one cipher text in the caesar cipher have 2 plain texts corresponding to it? What is the probability of this happening.

   (Rating : 2*)

4. Answer the above question for substitution cipher.

   (Rating : 3*)