CS648 : Randomized Algorithms
CSE, IIT Kanpur

Practice sheet 1: **Elementary probability and linearity of expectation**

1. **Bins with i balls**

   We throw m balls randomly, uniformly, and independently into n bins. What is the expected number of empty bins ? What is the expected number of bins containing exactly i balls.

2. **Sum of samples**

   An urn contains n balls numbered 1,2,...,n. We remove k balls at random (without replacement) and add up their numbers. Find the expected value of this final number.

3. **Surviving couples**

   Of the $2n$ people in a given collection of n couples, exactly m die. Assuming that the m have been picked at random, find the expected number of surviving couples.

4. **Magnet blocks**

   A total of n bar magnets are placed end to end in a line with random independent orientations. Adjacent like poles repel, ends with opposite polarities join to form blocks. Find the expected number of blocks of joined magnets.

5. **Stick break**

   Given a stick with n joints. The stick is dropped from certain height. During the fall, each joint breaks with probability p independent of other joints. What is the expected number of pieces into which the stick breaks ?

6. **Memoryless Guessing**

   To amaze your friends, you have them shuffle a deck of n cards and then turn over one card at a time. Before each card is turned over, you predict its identity. Unfortunately, you don't have any particular psychic abilities - and you are not good at remembering what has been turned over already- so your strategy is simply to guess a card uniformly at random from the deck. What is the expected number of correct predictions that you would make with this strategy ?

7. **Guessing with memory**

   Now let us consider second scenario. Your psychic abilities have not developed any further since last time, but you have become very good at remembering which cards have already been turned over. Thus when you predict the next card now, you only guess uniformly from among the cards *not yet seen*. What is the expected number of correct predictions that you would make with this strategy ?

8. **Expected value of product of random variables**

   Let X and Y be two independent random variables defined over a probability space (Ω, P). Prove that

   $$\mathbf{E}[X \cdot Y] = \mathbf{E}[X] \cdot \mathbf{E}[Y]$$

   (you are advised to internalize the proof).

9. **Two urns with white and black balls**

   Urn 1 contains 1933 white and 2067 black balls, while urn 2 contains 191 white and 167 black balls. 43 balls are randomly selected from urn 1 and are then put in urn 2. If 46 balls are then randomly selected from urn 2, compute the expected number of white balls in the trio.

10. **Total displacement**
    Let $a_1, a_2, \ldots, a_n$ be a uniformly random permutation of $\{1, 2, \ldots, n\}$. When sorting list $a_1, a_2, \ldots, a_n$, the element a_i must move a distance of $|a_i - i|$ places from its current position to reach its position in sorted order. Find the expected total distance that all the elements will have to be moved during sorting.

11. **Conditional probability**
    Suppose that n balls are thrown independently and uniformly at random into n bins.

    (a) Find the conditional probability that bin 1 has one ball given that exactly one ball fell into the first three bins.

    (b) Find the conditional probability that the number of balls in bin 1 is 2 under the condition that bin 2 received $n/2$ balls.

    (c) What is the conditional probability that nth bin is empty given that the bins numbered 1 to $n/2$ are empty ?

12. **Dependence among random variables**
    Recall the randomized quick sort algorithm. Let e_i denote ith smallest element in the array. Let X_{ij} be the random variable which takes value 1 if e_i is compared with e_j, and zero otherwise.

    Prove that the random variables $\{X_{ij} | 1 \leq i < j \leq n\}$ are not independent. In more precise words, demonstrate a few random variables from this set which are not independent.

13. **Balls into bins**
    Recall the "balls into bins" experiment.

    - Let X_i be the random variable which denote the number of balls choosing ith bin. Are the random variables X_i's independent ?
    - Let Y_i be the random variable which denotes the destination bin for ith ball. Are the random variables Y_i's independent ?

14. Recall Frievald's algorithm for checking equality $A \times B = C$. We selected a random {0,1}-vector: each entry was selected randomly uniformlyindependently from $\{0, 1\}$. The error probability was bounded by 2 essentially because there were 2 choices. What is wrong with the following arguments ?

    Select a random vector wherein each entry is a real number selected randomly uniformly and independently from $[0, 1]$. *The error probability of the algorithm will be 0. In this manner we get a deterministic algorithm for the problem !*

15. **Red-blue balls**
    There is a bag containing r red balls and b blue balls. We take out balls one by one, uniformly randomly and throw them away. During each step, every ball present in the bag is equally likely to be removed. What is the expected number of red balls left after all the black balls have been taken out ?

16. **Randomized binary search tree**
    We know various height balanced binary search trees (Red-Black trees or AVL trees). Each of these trees keep an additional balance field in each node. We also know that the procedure to keep the tree balanced is quite tedious (it takes 1 or 2 lectures to explain all cases for insertions and deletions). Here is a randomized variant of binary search tree.

Let S be the set of elements for which we wish to build a binary search tree. Build the tree incrementally as follows: Select and remove a uniformly random element from S and make it the root. Now repeat the following step until S becomes empty: Select and remove a randomly and uniformly selected element from S and insert it into the present binary search tree.

- Prove that the expected depth (distance from the root) of each element in the tree will be $O(\log n)$.

17. **A random prime number**

    Design a Las Vegas algorithm that takes a number n as input and outputs a prime number from the interval $[n, 2n]$. The expected running time of the algorithm has to be polynomial of the input size. You might like to know the famous result from IITK about prime numbers to solve this problem.

18. **Problem of an equation**

    p is a prime number. Let $a_1, \ldots, a_n$ be some integers in the range $[1, p-1]$. Let $X_1, ..., X_n$ be n random variables each taking integer value uniformly and independently in the range $[0, p-1]$. What is the probability that the following equation holds ?

$$(\sum_i a_i X_i) \textbf{ mod } p = 0$$

Note: There are a few questions in this sheet which were asked during the lectures.