**Assignment – III**

**(k-mer Analysis)**

**Deadline: 2nd Sept**

1. Why is it important to find the origin of replication?
2. Simulate observations having the binomial distribution with p = 0.25 and *n* = 1000. What is the probability of observing at least 280 A’s in such a sequence? [Hint: Obtain 10,000 simulations and compute the number of A’s in each run]. Compare your result with the normal approximation to binomial distribution.
3. Suppose X has a binomial distribution with p = 0.3 and *n* = 10. Compute P(X=0), P(X=2), E(X) and VarX.
4. Give one application of finding k-mers: (i) k=1, (ii) k=2, (iii) k=3, (iv) k=4 – 8, (v) k=9.
5. Give a strategy to find cluster of k-mer patterns in a newly sequenced genome.
6. How can the asymmetry of replication help us locate the origin of replication, *ori*?