"In the Name of God"

Engineering probabilities and Statistics **CE 40181**
Sharif University of Technology

Homework 5
Wednesday, Dec 27 2017

Reminders:

- **Typing the solutions has extra marks.**
- **Collaboration is permitted, but you must write solutions *by yourself without* assistance.**
- **Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.**
- **Late submissions will be treated according to the course policy.**

Problems:

1. The amount of regular unleaded gasoline purchased every week at a gas station near Sharif University follows the normal distribution with mean 50000 gallons and standard deviation 10000 gallons. The starting supply of gasoline is 74000 gallons, and there is a scheduled weekly delivery of 47000 gallons.

   (a) Find the probability that, after 11 weeks, the supply of gasoline will be below 20000 gallons.

   (b) How much should the weekly delivery be so that after 11 weeks the probability that the supply is below 20000 gallons is only 0.5%?

2. Suppose someone gives you a coin and claims that this coin is biased; that it lands on heads only 48% of the time. You decide to test the coin for yourself. If you want to be 95% confident that this coin is indeed biased, how many times must you flip the coin?(Take 0.02 as your confidence interval.)

   (a) Estimate the answer using the law of large numbers and the Chebyshev inequality.

   (b) Estimate the answer using the central limit theorem.

3. A survey of 1500 people is conducted to determine whether they prefer iOS operating system or Android operating system. The results show that 27% of people prefer iOS while the remaining 73% favour Android. Estimate the Margin of error in the poll with a confidence of 90%.

4. A communication system designer is simulating a communication link. The link is being designed for binary transmission, i.e., bits $b_k \in \{0, 1\}$, are transmitted, and bits

$\hat{b}_k \in \{0, 1\}$, are received in the kth bit-period. Noise and other sources of channel impairments result in bit-errors, i.e., the event $\hat{b}_k \neq b_k$. The bit errors are assumed to occur independently from one bit-period to the next with a probability of error, or bit error-rate (BER), of p_e. The designer wishes to estimate p_e by running n bits through a simulation model, counting the number of errors E by comparing the transmitted bits b_k with recovered bits $\hat{b}_k$, and obtaining a BER estimate $\hat{p_e} = \frac{E}{n}$, where E is the error count, i.e., the total number of bits in error in a stream of n bits. Suppose that the designer knows that $p_e \leq 10^{-4}$. What is the minimum value of n required to achieve a confidence level of 99% and confidence interval of length 10^{-5}?

5. Let X_1, X_2, ... be independent and identically distributed random variables with mean μ and standard deviation σ. Set $S_n = X_1 + ... + X_n$. For a large n, what is the approximate probability that S_n is between $E[S_n] - k\,\sigma_{S_n}$ and $E[S_n] + k\,\sigma_{S_n}$ (the probability of being k deviates from the mean) for k being equal to 1, 2 ,3 and 4?
   (a) Using the Central Limit Theorem.
   (b) Using Chebyshev's inequality.

6. The number of items produced in a factory during a week is a random variable with mean 50.
   (a) What can be said about the probability that this weeks production is at least 100?
   (b) If the variance of a weeks production is known to equal 25, can we obtain a better bound for part (a)?

7. Prove the weak law of large numbers.(Hint: Use the inequalities you have learned throughout this chapter)

R Problem

A biologist gets a sample of n cells from human blood on each day and measures the density of Heterochromatin in each cell which is a random variable uniformly distributed between 0 and 1. And then calculates the average of density of Heterochromatin in the sample set as $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$.
Assume that she gets a sample of size 2, 3, 4, and 5 on each of the days 2nd, 3rd, 4th and 5th of December respectively.

1. Draw four separate plots showing the distribution of $\bar{X}_n$ in each day. Also show expected value and standard deviation of each day's distribution on the corresponding plot with whatever method you prefer.

2. Biologists' community derived an agreement that errors of less than $\epsilon = 0.005$ for average of density of Heterochromatin is accepted for each sample set. Write a code that calculates the probability that each sample on each day of December (from 2nd to 31st with sample size equals to n on nth day of month) is <u>not</u> acceptable. Show your results on a graphical table in which first column represents the date and second column represents the value of $P[|\bar{X}_n - \mu| \geq \epsilon]$ for that date.