EE1080/AI1110/EE2102 Probability: HW 6

1st April, 2025

1. A statistician wants to estimate the mean height h (in meters) of a population, based on n independent samples $X_1, \cdots, X_n$ chosen uniformly from the entire population. He uses the sample mean $M_n = (X_1 + \cdots + X_n)/n$ as the estimate of h, and a rough guess of 1.0 meters for the standard deviation of the samples X_i.

   (a) How large should n be so that the standard deviation of M_n is at most 1 centimeter ?

   (b) How large should n be so that the Chebyshev's inequality guarantees that the estimate is within 5 centimeters from h, with probability atleast 0.99?

   (c) The statistician realizes that all the persons in the population have heights between 1.4 and 2.0 meters and revises the standard deviation figure that he uses. How should the values of n obtained in the earlier two parts be revised ?

2. Jensen's inequality: Let $f(x)$ be a convex function i.e., it satisfies $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$ for any $\lambda \in [0,1]$. Then show that:

$$f(E[X]) \leq E[f(X)]$$

   Proof Idea: Consider a tangent of $f(x)$ at point $(E[X], f(E[X]))$. Let $(x, \ell(x))$ be the points on this tangent. Since $f(x)$ is convex it follows that $\ell(x) \leq f(x)$ for all x.

3. Let f, g be two monotonic functions. Show that

$$E[f(X)g(X)] \geq E[f(X)]E[g(X)].$$

   Hint: Consider Y independent and identical to X and look at $E[(f(X) - f(Y))(g(X) - g(Y))]$. It should be ≥ 0 as both f and g are monotonic.

4. One sided Chebyshev bound: For random variable X with 0 mean show that:

$$P(X > a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

   Hint: Apply Markov inequality for $X + b$ and minimize over b

5. Using Chernoff bound find an upper bound for $P(X \geq i)$ where $X \sim \text{Poisson}(\lambda)$.

6. A set of 200 people consisting of 100 men and 100 women is randomly divided into pairs of two each. Given an upper bound to the probability that at most 30 of these pairs will consist of a man and a woman

7. Suppose a fair die is rolled 100 times. Let X_i be the value obtained at the ith roll. Compute an approximation for:

$$P(\prod_{i=1}^{100} X_i \leq a^{100}) \text{ where } 1 < a < 6$$

8. Let $X_1, X_2, \cdots$ be independent random variables that are uniformly distributed over $[-1, 1]$. Show that the sequence $Y_1, Y_2, \cdots$ converges in probability to some limit, and identify the limit, for each of the following cases:

   (a) $Y_n = X_n/n$;
   (b) $Y_n = (X_n)^n$;
   (c) $Y_n = X_1 X_2 \cdots X_n$
   (d) $Y_n = \max\{X_1, \cdots, X_n\}$

9. A factory produces X_n gadgets on day n, where the X_n are independent and identically distributed random variables with mean 5 and variance 9.

   (a) Find an approximation to the probability that the total number of gadgets produced in the 100 days is less than 440
   (b) Find (approximately) the largest value of n such that:

   $$P(X_1 + \cdots + X_n \geq 200 + 5n) \leq 0.05.$$

   (c) Let N be the first day on which the total number of gadgets produced exceeds 1000. Calculate an approximation to the probability that $N \geq 220$.

10. $\Omega = [0, 1]$. Let

$$X_n(\omega) = \begin{cases} 1 & \omega \in [0, \frac{1}{n}] \\ 0 & \text{otherwise}] \end{cases}$$

    Identify the limit RV X and the forms of convergence.

11. X_n be a normal random variable with mean 0 and variance $1/n$. Identify the limit and the forms of convergence.

12. $P(X_n = 1) = \frac{1}{n^2} = 1 - P(X_n = 0)$. Identify the limit and the forms of convergence.