Endsem: Probability and Statistics (100 marks)

Instruction:

- Please state reasons wherever applicable.
- Use precise mathematical arguments, no speeches.

Each question: 10 marks

1. Stochastic simulation: Suppose you want to generate samples from a discrete random variable X having pmf $\{p_j, j \geq 0\}$. Now assume that you have access to samples from another discrete random variable Y with pmf $\{q_j, j \geq 0\}$ with the property that $\frac{p_j}{q_j} \leq c$ for some constant c and for all j such that $p_j > 0$. The rejection method generates samples of X as follows.

   (a) Simulate/Generate the value of Y with mass function $\{q_j, j \geq 0\}$.

   (b) Generate random number U which is uniform in the interval $[0, 1]$.

   (c) If $U < \frac{p_Y}{cq_Y}$, set $X = Y$ and stop. Otherwise return to Step (a).

   Prove that samples of X generated using the above algorithm indeed have pmf $\{p_j, j \geq 0\}$.

2. MGF: Derive the expression for the MGF of a Gaussian $\mathcal{N}(\mu, \sigma^2)$ random variable and use the MGF to identify the first and the second moment. Furthermore, using MGF, show that sum of n independent Gaussian $\mathcal{N}(\mu, \sigma^2)$ random variables is also a Gaussian random variable. What are the resulting mean and variance parameters ?

3. MLE: Consider a Gaussian random variable X with unknown mean μ and unknown standard deviation σ. Suppose you observe k iid samples from this random variable which is denoted by $\mathcal{D} = \{x_1, x_2, \ldots, x_k\}$. Find the maximum likelihood estimate for μ and σ. Is the MLE for the standard deviation biased ? justify why.

4. Functions of random vectors: Let $X = [X_1, X_2]$ where X_1 and X_2 are independent exponential random variables with parameter 1. Find the probability density function of $U = [U_1, U_2]$ where $U_1 = X_1 + X_2$ and $U_2 = \frac{X_1}{X_1+X_2}$.

Each question: 15 marks

1. Bayesian Inference problem: Suppose $D = \{x_1, \dots, x_n\}$ is a data set consisting of independent samples of a Bernoulli random variable with unknown parameter θ, i.e., $f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$ for $x_i \in \{0,1\}$. Now assume a uniform $U[0,1]$ prior on the unknown parameter θ. Obtain an expression for the posterior distribution on θ. Using this obtain θ_{MAP} and the conditional expectation estimator θ_{CE}. (Hint: you may use the fact that $\int_0^1 \theta^m(1-\theta)^r d\theta = \frac{m!r!}{(m+r+1)!}$)

2. (a) (8 mks) Let $X_1, X_2, \dots$ be a sequence of random variables with density $f_{X_n}(x) = \frac{n}{2}e^{-n|x|}$. Show that X_n converges to 0 in probability and in distribution. (Do not use the result that convergence in probability implies convergence in distribution)
   (b) (7 mks) Let X_n be $Poisson(n\lambda)$ random variable for $n = 1, 2, 3, \dots$. Consider the sequence of random variables $Y_n = \frac{X_n}{n}$ for $n = 1, 2, 3, \dots$. Show that Y_n converges in mean square sense to λ.

3. (a) (8mks): Let $\mathcal{D} = \{x_1, \dots x_n\}$ denote i.i.d samples from a uniform random variable $U[0, a]$ where a is unknown. Find an MLE estimate for the unknown parameter a.
   (b) (7mks): Let $\mathcal{D} = \{x_1, \dots x_n\}$ denote i.i.d samples from a Poisson random variable with unknown parameter γ. Find an MLE estimate for the unknown parameter γ.

4. (a)(7mks): Consider a discrete time Markov coin with the following transition probabilities: $p_{ij} = 0$ when $i = j$. What is the probability of head and tail in the nth step and when the initial distribution is $\mu = [\mu_1, \mu_2]$.
   (b) (8mks): Find the limiting distribution and the stationary distribution π for Markov Chain with the following transition probability matrix
   $P = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Now obtain the values of F_{ii} (probability of ever returning to state i, having started in state i) for each of the 4 states and based on the values identify if each state is transient or recurrent.