SI 140A-02 Probability & Statistics for EECS, Fall 2024
Homework 11

Name:
Student ID:

Due on Dec. 24, 2024, 11:59 UTC+8

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points of the problem.
- Write your homework in English; otherwise you will get no points of this homework.
- Any form of plagiarism will lead to 0 point of this homework.

Problem 1

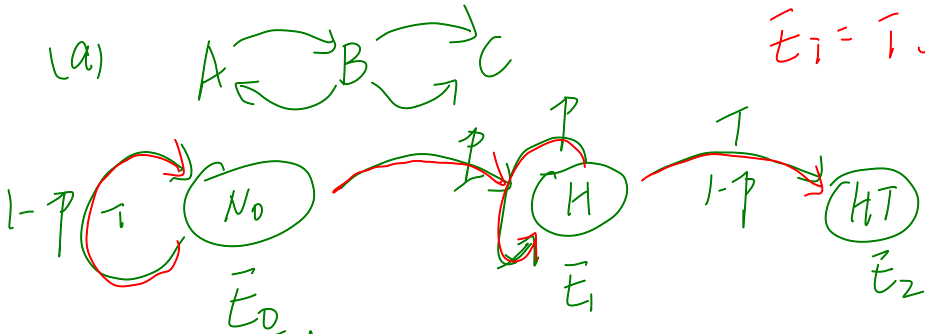
A coin with probability p of Heads is flipped repeatedly, where p is a known constant, with $0 < p < 1$.

(a) What is the expected number of flips until the pattern HT is observed?

(b) What is the expected number of flips until the pattern HH is observed?

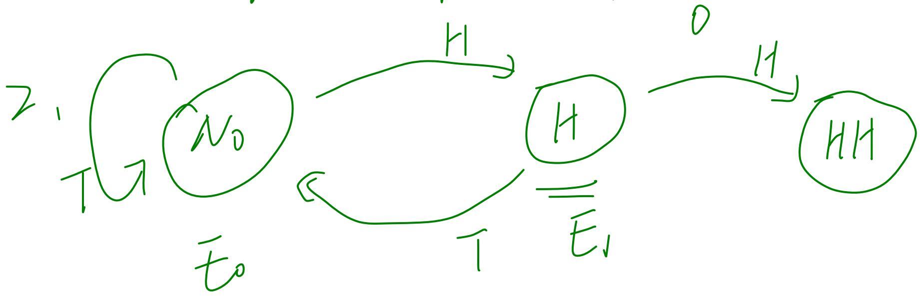
(a)

$\bar{E}_i$ = i steps to achieve the goal

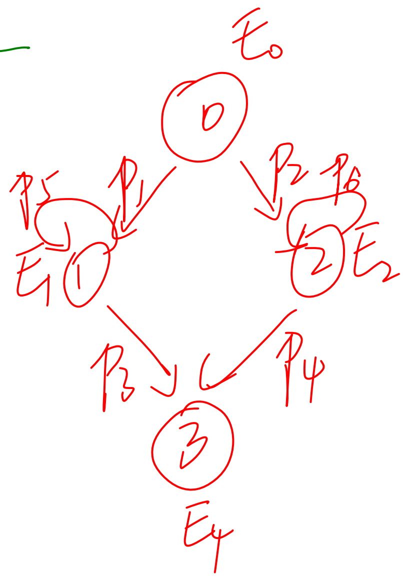


$$\begin{cases} \bar{E}_0 = 1 + p\bar{E}_1 + (1-p)\bar{E}_0 \\ \bar{E}_1 = 1 + p\bar{E}_1 + (1-p)\underbrace{\bar{E}_2}_{0} \end{cases} \Rightarrow \bar{E}_0 = \frac{1}{p} + \frac{1}{1-p}$$

2,



$$\begin{cases} \bar{E}_0 = 1 + p\bar{E}_1 + (1-p)\bar{E}_0 \\ \bar{E}_1 = 1 + p \times 0 + (1-p)\bar{E}_0 \end{cases} \Rightarrow \bar{E}_0 = \frac{1+p}{p^2}$$



$$\begin{cases} \bar{E}_0 = 1 + p_1\bar{E}_1 + p_2\bar{E}_2 \\ \bar{E}_1 = 1 + p_3\underbrace{\bar{E}_4}_{0} + p_5\bar{E}_1 \\ \bar{E}_2 = 1 + p_4\underbrace{\bar{E}_4}_{0} + p_6\bar{E}_2 \end{cases}$$

Problem 2

Given two random variables X and Y, the corresponding joint PDF is

$$f_{X,Y}(x,y) = \begin{cases} x+y & \text{if } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{Otherwise} \end{cases}$$

Find $E[Y \mid X]$ and $L[Y \mid X]$.

(1) $f_X(x) = \int_0^1 f_{X,Y}(x,y)\,dy = x + \frac{1}{2}$

$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{x+y}{x+\frac{1}{2}}$, (x, y)

$E[Y|X=x] = \int_0^1 y \cdot \frac{x+y}{x+\frac{1}{2}}\,dy = \frac{3x+2}{6x+3}$

(2) $L[Y|X] = E[Y] + \frac{Cov(X,Y)}{Var(X)}(X - E(X))$

$E[Y] = \int_0^1 y f_Y(y)\,dy = \int_0^1 y \left[\int_{-\infty}^{+\infty} f(x,y)\,dx\right] dy = \frac{7}{12}$

$Var(X) = E[X^2] - E^2[X] \Rightarrow \frac{11}{144}$

$\int_0^1 x^2(\frac{1}{2}+x)\,dx = \frac{5}{12}$

$E[X] = \int_0^1 x(\frac{1}{2}+x)\,dx = \frac{7}{12}$

$Cov(X,Y) = E[X,Y] - E[X]E[Y]$

$= \int_0^1\int_0^1 xy f(x,y)\,dx\,dy - \underbrace{E[X]}_{\frac{7}{12}}\,\underbrace{E[Y]}_{\frac{7}{12}}$

$= -\frac{1}{144}$

$L = \frac{7}{11} - \frac{1}{11}X$

Problem 3

Let X be the height of a randomly chosen adult man, and Y be his father's height, where X and Y have been standardized to have mean 0 and standard deviation 1. Suppose that (X, Y) is Bivariate Normal, with $X, Y \sim \mathcal{N}(0, 1)$ and $\text{Corr}(X, Y) = \rho$.

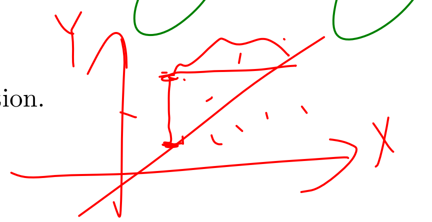
(a) Let $y = ax + b$ be the equation of the best line for predicting Y from X (in the sense of minimizing the mean squared error), e.g., if we were to observe $X = 1.3$ then we would predict that Y is $1.3a + b$. Now suppose that we want to use Y to predict X, rather than using X to predict Y. Give and explain an intuitive guess for what the slope is of the best line for predicting X from Y.

(b) Find a constant c (in terms of ρ) and an r.v. V such that $Y = cX + V$, with V independent of X. Hint: Start by finding c such that $\text{Cov}(X, Y - cX) = 0$.

(c) Find a constant d (in terms of ρ) and an r.v. W such that $X = dY + W$, with W independent of Y.

(d) Find $E(Y \mid X)$ and $E(X \mid Y)$.

(e) Reconcile (a) and (d), giving a clear and correct intuitive explanation.



(a) $Y = aX + b \Rightarrow X = \frac{1}{a}Y + \frac{b}{a}$

$$a = \rho \cdot \frac{\sigma_Y}{\sigma_X} \Rightarrow a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\rho = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\underbrace{\sqrt{\sum (x_i - \bar{x})^2}}_{\sigma_X^2} \cdot \underbrace{\sqrt{\sum (y_i - \bar{y})^2}}_{\sigma_Y^2}}$$

$$a = \rho \cdot \frac{\sigma_Y}{\sigma_X} = \rho$$

$$\frac{1}{a} = \frac{1}{\rho}$$

(b) Independent, $\text{Cov}(X, V) = 0 = \text{Cov}(X, Y - cX)$

$\Rightarrow c = \rho$

(c) Independent $\text{Cov}(Y, W) = 0$

$= \text{Cov}(Y, X - dY) \Rightarrow d = \rho$

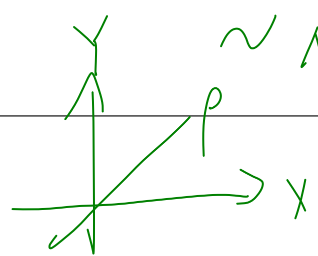
(d) $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$

$$= \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(y - \rho x)^2}{2(1-\rho^2)}\right)$$

$\begin{cases} X \sim N(\mu_X, \sigma_X^2) \\ Y \sim N(\mu_Y, \sigma_X^2) \\ \text{Corr} = \rho \end{cases}$

$\sim N(\rho x, 1 - \rho^2)$ $\begin{cases} E[Y|X] = \rho X \\ E[X|Y] = \rho Y \end{cases}$

(e)



Problem 4

Show the following orthogonality properties of MMSE:

(a) For any function $\phi(\cdot)$, one has

$$E[(Y - E[Y \mid X])\phi(X)] = 0$$

(b) If the function $g(X)$ satisfied

$$E[(Y - g(X))\phi(X)] = 0, \forall \phi(\cdot)$$

then $g(X) = E(Y \mid X)$.

(a) $$E[(Y - E[Y|X])\varphi(X)] = E[Y\varphi(X)] - E[E(Y|X)\varphi(X)]$$

$$= E[Y\varphi(X)] - E[E[Y\varphi(X)|X]]$$

$$= E[Y\varphi(X)] - E[Y\varphi(X)]$$

$$= 0$$

(b) $$g(X) = E[Y|X] + f(X)$$

$$E[(Y - E[Y|X] - f(X))\varphi(X)]$$

$$= \underbrace{E[Y\varphi(X)] - E[E(Y|X)\varphi(X)]}_{0} - E[f(X)\varphi(X)] = 0$$

$$E[f(X)\varphi(X)] = 0 \Rightarrow f(X) = 0$$

$$g(X) = E[Y|X]$$

Problem 5

Given a coin with the probability p of landing heads. p is unknown and we need to estimate its value through data. In our data collection model, we have n independent tosses, result of each toss is either Head or Tail. Let X denote the number of heads in the total n tosses. Now we conduct experiments to collect data and find $X = k$. Then we need to find $\hat{p}$, the estimation of p.

(a) Assume p is an unknown constant. Find $\hat{p}$ through the MLE (Maximum Likelihood Estimation) rule.

(b) Assume p is a random variable with a prior distribution $p \sim \text{Beta}(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MAP (Maximum a Posterior Probability) rule.

(c) Assume p is a random variable with a prior distribution $p \sim \text{Beta}(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MMSE (Minimal Mean Squared Error) rule.

(a) $L(p|k) = \binom{n}{k} p^k (1-p)^{n-k} \Rightarrow \mathcal{L}(p|k) = \log L = \log\binom{n}{k} + k\log p + (n-k)\ln(1-p)$

$\frac{\partial l}{\partial p} = \frac{k}{p} - \frac{n-k}{1-p} = 0 \Rightarrow \hat{p} = \frac{k}{n}$

(b) $f(p) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} \propto p^{a-1}(1-p)^{b-1}$ $\quad P(p|X=k) \propto \underbrace{P(X=k|p)}_{p^k(1-p)^{n-k}} \underline{P(p)}$

$f_{p|X=k} \propto p^{a+k-1}(1-p)^{b+n-k-1}$

$p^* = \underset{p}{\arg\max}\, f_{p|X} \Rightarrow$

$\log f_{p|X} = (a+k-1)\log p + (b+n-k-1)\log(1-p) = g(p)$

$g'(p) = -\frac{a+k-1}{p} - \frac{b+n-k-1}{1-p}$ $\quad g''(p) = -\frac{a+k+1}{p^2} - \frac{b+n-k-1}{(1-p)^2} < 0$

$g'(p) = 0 \Rightarrow p^* = \frac{a+k-1}{a+b+n-2}$

(c) Beta-Binomial conjugacy

$E = \frac{a+k}{a+b+n}$

Problem 6

(**Optional Challenging Problem**) Use two different methods to show that if X and Y are jointly Normal random variables, then

$$E[Y \mid X] = L[Y \mid X] = E(Y) + \frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X - E(X))$$

$$E[Y|X] = \alpha X + \beta = L[Y|X]$$

$$\text{Cov}(X,Y) = \text{Cov}(X, \alpha X + \beta) = \alpha \, \text{Var}(X)$$

$$\Rightarrow \alpha = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} \qquad \beta = E[Y] - \alpha E[X]$$

$$E[Y|X] - \alpha E[X]$$

$$E[Y|X] = \frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X - E(X)) + E(Y)$$