$k^k (n-k)^{n-k} \cdot \left(\frac{1}{k} p\right)^k \left(\frac{1-p}{n-k}\right)^{n-k}$

2023Fall Probability & Statistics for EECS $k(n-k)$ 2023/12/31

Homework 12

Professor: Ziyu Shao & Dingzhu Wen *Due: 2024/1/7 10:59pm*

$X \mid p$

1. 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 $p \sim Beta(a, b)$, 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 $p \sim Beta(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MMSE (Minimal Mean Squared Error) rule.

2. 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 $E(Y|X)$ and $E(X|Y)$.

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

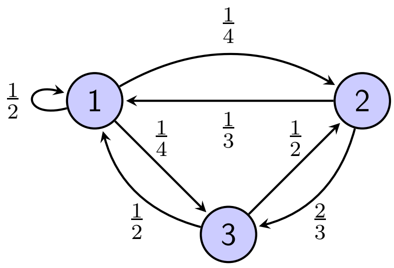
3. Two chess players, Vishy and Magnus, play a series of games. Given p, the game results are i.i.d. with probability p of Vishy winning, and probability $q = 1 - p$ of Magnus winning (assume that each game ends in a win for one of the two players). But p is unknown, so we will treat it as an r.v. To reflect our uncertainty about p, we use the prior $p \sim \text{Beta}(a, b)$, where a and b are known positive integers and $a \geq 2$.

   (a) Find the expected number of games needed in order for Vishy to win a game (including the win). Simplify fully; your final answer should not use factorials or Γ.

   (b) Explain in terms of independence vs. conditional independence the direction of the inequality between the answer to (a) and $1 + E(G)$ for $G \sim \text{Geom}(\frac{a}{a+b})$.

   (c) Find the conditional distribution of p given that Vishy wins exactly 7 out of the first 10 games.

4. Given a Markov chain with state-transition diagram shown as follows:



   (a) Is this chain irreducible?

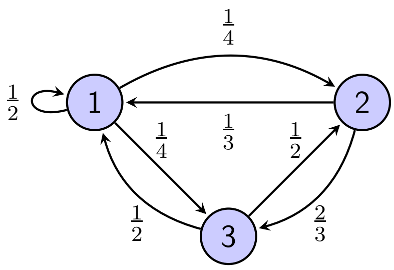
   (b) Is this chain aperiodic?

   (c) Find the stationary distribution of this chain.

   (d) Is this chain reversible?

$$\begin{matrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{matrix}$$

5. Given a Markov chain with state-transition diagram shown as follows:



   (a) Find $P(X_3 = 3 | X_2 = 2)$ and $P(X_4 = 1 | X_3 = 2)$.

   (b) If $P(X_0 = 2) = \frac{2}{5}$, find $P(X_0 = 2, X_1 = 3, X_2 = 1)$.

(c) Find $P(X_2 = 1|X_0 = 2)$, $P(X_2 = 2|X_0 = 2)$, and $P(X_2 = 3|X_0 = 2)$.

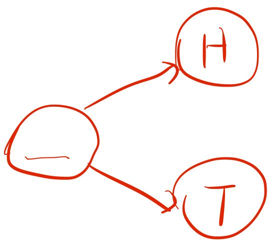
(d) Find $E(X_2|X_0 = 2)$.

6. (**Optional Challenging Problem**) A fair coin is flipped repeatedly. We use H to denote "Head appeared" and T to denote the "Tail appeared".

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

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

   (c) What is the probability that pattern HTHT is observed earlier than THTT?



HT