SI 140A-02 Probability & Statistics for EECS, Fall 2024
Homework 12

Name: **Wenye Xiong**
Student ID: 2023533141

Due on Dec. 31, 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

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample from the distribution $\mathcal{N}\left(\mu, \sigma^2\right)$, where both μ and σ^2 are unknown constants. Suppose the observed data is $\boldsymbol{x} = (x_1, \ldots, x_n)$, find both $\hat{\mu}$ (estimate of μ) and $\hat{\sigma}^2$ (estimate of σ^2) through the MLE (Maximum Likelihood Estimation) rule.

Solution

The likelihood function is given by

$$f(\boldsymbol{x}|\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

And we have the log-likelihood function

$$\begin{aligned}\ell(\mu, \sigma^2; \boldsymbol{x}) &= \log L(\mu, \sigma^2; \boldsymbol{x}) \\ &= -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\end{aligned}$$

To find the MLE of μ, we take the derivative of ℓ with respect to μ and set it to zero:

$$\begin{aligned}\frac{\partial \ell}{\partial \mu} &= \frac{1}{2\sigma^2}\sum_{i=1}^{n} 2\left(x_i-\mu\right) \\ &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \\ &= \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i - n\mu\right)\end{aligned}$$

Setting the derivative to zero, we have

$$\sum_{i=1}^{n} x_i - n\mu = 0$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

Thus, the MLE of μ is $\hat{\mu} = \bar{x}$.
To find the MLE of σ^2, we take the derivative of ℓ with respect to σ^2 and set it to zero:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\left(\sigma^2\right)^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

Setting the derivative to zero, we have

$$\tfrac{n}{2\sigma^2} = \tfrac{1}{2(\sigma^2)^2}\textstyle\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Substitute $\mu = \bar{x}$ into the equation, we have

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Thus, the MLE of σ^2 is $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$.

Problem 2

In humans (and many other organisms), genes come in pairs. Consider a gene of interest, which comes in two types (alleles): type a and type A. The genotype of a person for that gene is the types of the two genes in the pair: AA, Aa, or aa (aA is equivalent to Aa). According to the Hardy-Weinberg law, for a population in equilibrium the frequencies of AA, Aa, aa will be $p^2, 2p(1-p), (1-p)^2$ respectively, for some p with $0 < p < 1$. Suppose that the Hardy-Weinberg law holds, and that n people are drawn randomly from the population, independently. Let X_1, X_2, X_3 be the number of people in the sample with genotypes AA, Aa, aa, respectively.

(a) What is the joint PMF of X_1, X_2, X_3 ?

(b) What is the distribution of the number of people in the sample who have an A ?

(c) What is the distribution of how many of the $2n$ genes among the people are A 's?

(d) Now suppose that p is unknown, and must be estimated using the observed data X_1, X_2, X_3. The maximum likelihood estimator (MLE) of p is the value of p for which the observed data are as likely as possible. Find the MLE of p.

(e) Now suppose that p is unknown, and that our observations can't distinguish between AA and Aa. So for each person in the sample, we just know whether or not that person is an aa (in genetics terms, AA and Aa have the same phenotype, and we only get to observe the phenotypes, not the genotypes). Find the MLE of p.

Solution

(a) By the story of the Multinomial, $(X_1, X_2, X_3) \sim \operatorname{Mult}_3\left(n, \left(p^2, 2pq, q^2\right)\right)$, where $q = 1-p$. The PMF is

$$P\left(X_1 = n_1, X_2 = n_2, X_3 = n_3\right) = \frac{n!}{n_1! n_2! n_3!} p^{2n_1} (2pq)^{n_2} q^{2n_3}$$

for $n_1 + n_2 + n_3 = n$.

(b) By the story of the Binomial (defining "success" as having an A and "failure" as not having an A), the distribution is $\operatorname{Bin}\left(n, p^2 + 2pq\right)$.

(c) Let Y_j be how many A 's the j th person in the sample has. Then Y_j is 2 with probability $p^2, 1$ with probability $2pq$, and 0 with probability q^2, so $Y_j \sim \operatorname{Bin}(2, p)$. The Y_j are also independent. Therefore, $Y_1 + \cdots + Y_n \sim \operatorname{Bin}(2n, p)$.

(d) Let x_1, x_2, x_3 be the observed values of X_1, X_2, X_3. The MLE of p is the value of p that maximizes the function $L(p) = p^{2x_1} (pq)^{x_2} q^{2x_3} = p^{2x_1 + x_2} (1-p)^{x_2 + 2x_3}$ (we can omit factors which are constant with respect to p, since such constants do not affect where the maximum is). Equivalently, we can maximize the log:

$$\log L(p) = \left(2x_1 + x_2\right) \log p + \left(x_2 + 2x_3\right) \log(1-p)$$

Setting the derivative of $\log L(p)$ equal to 0 , we have

$$\frac{2x_1 + x_2}{p} - \frac{x_2 + 2x_3}{1-p} = 0$$

which rearranges to

$$p = \frac{2x_1 + x_2}{2\left(x_1 + x_2 + x_3\right)} = \frac{2x_1 + x_2}{2n}$$

This value of p does maximize $\log L(p)$ since the derivative of $\log L(p)$ is positive everywhere to the left of it and is negative everywhere to the right of it. Thus, the MLE of p, which we denote by $\hat{p}$, is given by $\hat{p} = \left(2X_1 + X_2\right)/(2n)$. Note that this has an intuitive interpretation: it is the fraction of A 's among the $2n$ genes.

Problem 2 continued on next page...

(e) Let $Y \sim \operatorname{Bin}\left(n, q^2\right)$ be the number of aa people, and let y be the observed value of Y. We need to maximize the function $L_2(q) = q^{2y}\left(1-q^2\right)^{n-y}$ (we will maximize over q and then find the corresponding value of p). Then

$$\log L_2(q) = 2y \log q + (n-y)\log\left(1-q^2\right)$$

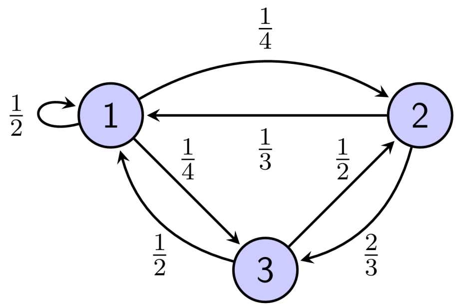
so

$$\frac{d \log L_2(q)}{dq} = \frac{2y}{q} - \frac{2q(n-y)}{1-q^2}$$

which simplifies to $y = q^2 n$. By looking at the sign of the derivative, we see that $\log L_2(q)$ is maximized at $q = \sqrt{y/n}$. Thus, the MLE of q is $\sqrt{Y/n}$, which shows that the MLE of p is $1-\sqrt{Y/n}$.

Problem 3

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



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

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

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

(d) Find $E(X_2 \mid X_0 = 2)$.

Solution

For the Markov chain, we have the transition matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

(a) According to the transition matrix, we have

$$P(X_3 = 3 \mid X_2 = 2) = \frac{2}{3}$$

$$P(X_4 = 1 \mid X_3 = 2) = \frac{1}{3}$$

(b) We have

$$P(X_0 = 2, X_1 = 3, X_2 = 1) = P(X_0 = 2)\, P(X_1 = 3 \mid X_0 = 2)\, P(X_2 = 1 \mid X_1 = 3) = \frac{2}{5} \times \frac{2}{3} \times \frac{1}{2} = \frac{2}{15}$$

(c) We are solving the probability of X_2 given X_0, so we can use the square of the transition matrix P to get the transition matrix from X_0 to X_2:

$$P^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} = \begin{bmatrix} \frac{11}{24} & \frac{1}{4} & \frac{7}{24} \\ \frac{1}{2} & \frac{5}{12} & \frac{1}{12} \\ \frac{5}{12} & \frac{1}{8} & \frac{11}{24} \end{bmatrix}$$

Thus, we have $P(X_2 = 1 \mid X_0 = 2) = \frac{1}{2}$ $P(X_2 = 2 \mid X_0 = 2) = \frac{5}{12}$ $P(X_2 = 3 \mid X_0 = 2) = \frac{1}{12}$

(d) We have

$$E(X_2 \mid X_0 = 2) = 1 \times P(X_2 = 1 \mid X_0 = 2) + 2 \times P(X_2 = 2 \mid X_0 = 2) + 3 \times P(X_2 = 3 \mid X_0 = 2) = \frac{1}{2} + \frac{5}{6} + \frac{1}{4} = \frac{19}{12}$$

Problem 4

Let the random variable $X \sim \mathcal{N}\left(\mu, \tau^2\right)$. Given $X = x$, random variables $Y_1, Y_2, \ldots, Y_n$ are i.i.d. and have the same conditional distribution, i.e., $Y_i \mid X = x \sim \mathcal{N}\left(x, \sigma^2\right)$. Define the sample mean $\bar{Y}$ as follows:

$$\bar{Y} = \frac{Y_1 + \cdots + Y_n}{n}$$

1. Find the posterior PDF of X given $\bar{Y}$.
2. Find the MAP (Maximum a Posterior Probability) estimates of X given $\bar{Y}$.
3. Find the MMSE estimates of X given $\bar{Y}$. (We know that the MMSE of X given Y is given by $g(Y) = E[X \mid Y]$).

Solution

(a) Using Bayes' theorem, the posterior PDF is:

$$p(X \mid \bar{Y}) \propto p(\bar{Y} \mid X) p(X),$$

where $p(X)$ is the prior PDF of X, i.e., $X \sim \mathcal{N}(\mu, \tau^2)$, $p(\bar{Y} \mid X)$ is the likelihood, i.e., $\bar{Y} \mid X = x \sim \mathcal{N}(x, \sigma^2/n)$.

So we have

$$p(\bar{Y} \mid X = x) = \frac{1}{\sqrt{2\pi\sigma^2/n}} \exp\left(-\frac{n}{2\sigma^2}(\bar{Y} - x)^2\right).$$

$$p(X = x) = \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{1}{2\tau^2}(x - \mu)^2\right).$$

Combining the likelihood and prior, we have the unnormalized posterior

$$p(X \mid \bar{Y}) \propto \exp\left(-\frac{n}{2\sigma^2}(\bar{Y} - x)^2 - \frac{1}{2\tau^2}(x - \mu)^2\right).$$

Simplify the exponent, we have

$$-\frac{n}{2\sigma^2}(\bar{Y} - x)^2 - \frac{1}{2\tau^2}(x - \mu)^2 = -\frac{1}{2}\left[\left(\frac{n}{\sigma^2} + \frac{1}{\tau^2}\right)x^2 - 2\left(\frac{n\bar{Y}}{\sigma^2} + \frac{\mu}{\tau^2}\right)x + \text{const}\right].$$

This is quadratic in x, so the posterior is

$$X \mid \bar{Y} \sim \mathcal{N}(\nu, \rho^2),$$

where:

$$\rho^2 = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}, \quad \nu = \rho^2\left(\frac{n\bar{Y}}{\sigma^2} + \frac{\mu}{\tau^2}\right).$$

(b) The MAP estimate of X given $\bar{Y}$ is the mode of the posterior, which is the mean of the posterior. So the MAP estimate is:

$$\hat{X}_{\text{MAP}} = \nu = \rho^2\left(\frac{n\bar{Y}}{\sigma^2} + \frac{\mu}{\tau^2}\right) = \frac{\frac{n}{\sigma^2}\bar{Y} + \frac{1}{\tau^2}\mu}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}.$$

(c) The MMSE estimate is given by the conditional expectation

$$\hat{X}_{\text{MMSE}} = E[X \mid \bar{Y}] = \nu = \rho^2\left(\frac{n\bar{Y}}{\sigma^2} + \frac{\mu}{\tau^2}\right) = \frac{\frac{n}{\sigma^2}\bar{Y} + \frac{1}{\tau^2}\mu}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}.$$