Biostat 202B Homework 4

Due May 9, 2024 @ 11:59PM

AUTHOR
Hanbei Xiong 605257780

Problem 1

Suppose $X_1, \ldots, X_n$ are iid with pdf $f(x;\theta) = e^{-x/\theta}/\theta$ for $0 < x < \infty$, zero elsewhere. Find the mle of $P(X \leq 2)$.

Answer:

$L(\theta) = \prod_{i=1}^n f(x_i;\theta) = \prod_{i=1}^n \frac{e^{-x_i/\theta}}{\theta} = \frac{e^{-\sum x_i/\theta}}{\theta^n}$

$l(\theta) = \log L(\theta) = -\sum x_i/\theta - n\log\theta$

$\frac{dl(\theta)}{d\theta} = \sum x_i/\theta^2 - n/\theta$

By setting $\frac{dl(\theta)}{d\theta} = 0$, we have $\hat{\theta}_{MLE} = \frac{\sum x_i}{n} = \bar{x}$

$F_X(x) = \int_0^x \frac{e^{-z/\theta}}{\theta} dz = -e^{-z/\theta}\Big|_0^x = 1 - e^{-x/\theta}$ for $0 < x \leq \theta$

$P(X \leq 2) = 1 - e^{-2/\theta} = 1 - e^{-2/\bar{x}}$

Problem 2

Suppose $X_1, \ldots, X_n$ are iid with pdf $f(x;\theta) = 2x/\theta^2$, $0 < x \leq \theta$, zero elsewhere.

a. Find the MLE $\hat{\theta}$ for θ.

b. Prove or disprove that $\hat{\theta}_{\text{MLE}}$ is unbiased.

c. Find the MLE for the median of the distribution.

(a) Answer:

$L(\theta) = \prod_{i=1}^n f(x_i;\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}}$ for $0 < x_i \leq \theta$

$L(\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}}$ for $0 < x_{(1)} < x_{(2)} < \ldots < x_{(n)} \leq \theta$

$L(\theta) = \frac{2^n \prod_{i=1}^n x_i}{\theta^{2n}} I(0 < x_{(1)}) \ldots I(x_{(n)} \leq \theta)$

So, $\hat{\theta}_{MLE} = x_{(n)}$ is the smallest value that maximizes the likelihood

(b) Answer:

$F_X(x) = \int_0^x \frac{2z}{\theta^2} dz = \frac{z^2}{\theta^2}\Big|_0^x = \frac{x^2}{\theta^2}$ for $0 < x \leq \theta$

$F_{X_{(n)}}(x) = P(X_{(n)} < x) = P(X_1 < x, X_2 < x, \ldots, X_n < x) = P(X_1 < x)^n = F_X^n(x)$

$E(x_{(n)}) = \int_0^\theta xnf_X(x)F_X^{n-1}(x)dx = \int_0^\theta xn\frac{2x}{\theta^2}(\frac{x^2}{\theta^2})^{n-1}dx = \frac{2n}{\theta^{2n}}\int_0^\theta x^{2n}dx = \frac{2n\theta}{2n+1} \neq \theta$

Hence, $\hat{\theta}_{MLE}$ is not unbiased

(c) Answer:

The median of distribution means $F_X(m) = 0.5$

$\frac{m^2}{\theta^2} = 0.5 \Rightarrow m = \frac{\theta}{\sqrt{2}}$

$\hat{m}_{MLE} = \frac{\hat{\theta}_{MLE}}{\sqrt{2}} = \frac{x_{(n)}}{\sqrt{2}}$

Problem 3

Suppose that $X_1, X_2, \ldots, X_{2n}$ are independently distributed as $X_i \sim N(0, \sigma^2)$ for $i = 1, \ldots, n$ and $X_i \sim N(0, 2\sigma^2)$ for $i = n+1, \ldots, 2n$. Find $\hat{\sigma}^2$, the maximum likelihood estimate of σ^2 based on the entire sample $X_1, \ldots, X_{2n}$. What is the **finite-sample** distribution of the appropriately normalized $\hat{\sigma}^2$? Justify your reasoning.

Answer:

$L(\sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x_i^2}{2\sigma^2}} \prod_{j=n+1}^{2n} \frac{1}{2\sqrt{\pi}\sigma}e^{-\frac{x_j^2}{4\sigma^2}}$

If we take log of likelihood, take derivative and set it to 0,

We get $\hat{\sigma}^2_{MLE} = \frac{2\sum_{i=1}^n x_i^2 + \sum_{j=i+1}^{2n} x_j^2}{4n} = \frac{1}{2n}(\sum_{i=1}^n x_i^2 + \frac{1}{2}\sum_{j=i+1}^{2n} x_j^2)$

Since $\frac{\sum_{i=1}^n x_i^2}{\sigma^2} \sim \chi^2(n)$, $\frac{\sum_{j=n+1}^{2n} x_i^2}{2\sigma^2} \sim \chi^2(n)$ and X_i are independent r.v.s, we have

$\frac{\sum_{i=1}^n x_i^2}{\sigma^2} + \frac{\sum_{j=n+1}^{2n} x_i^2}{2\sigma^2} \sim \chi^2(2n)$

$\frac{2n\hat{\sigma}^2_{MLE}}{\sigma^2} = \frac{\sum_{i=1}^n x_i^2}{\sigma^2} + \frac{\sum_{j=n+1}^{2n} x_i^2}{2\sigma^2} \sim \chi^2(2n)$

Problem 4

Let $X_1, \ldots, X_n$ be iid Bernoulli(p). The object is to estimate $\theta = 1/p$.

a. Find the maximum likelihood estimator of θ.

b. What is the asymptotic distribution of the appropriately normalized MLE as $n \to \infty$?

(a) Answer:

$$L(p) = \prod_{i=1}^{n} p(x_i; p) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

$l(p) = \log L(\theta) = (\sum_{i=1}^{n} x_i)\log(p) + (n - \sum_{i=1}^{n} x_i)\log(1-p)$

$\frac{dl(p)}{dp} = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{n-\sum_{i=1}^{n} x_i}{1-\theta} = 0$

$p = \frac{\sum_{i=1}^{n} x_i}{n}$

Since $\theta = \frac{1}{p}, \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{n}{n\bar{X}} = \frac{1}{\bar{X}}$

(b) Answer:

By CLT, $\sqrt{n}(\bar{X} - p) \xrightarrow{d} N(0, p(1-p))$

By Delta method,

$\sqrt{n}(\frac{1}{\bar{X}} - \frac{1}{p}) \xrightarrow{d} N(0, \frac{1}{p^4}p(1-p)) = N(0, \frac{1-p}{p^3})$

Problem 5

Let $X_1, \ldots, X_n$ be a random sample from a gamma distribution with $\alpha = 4$ and $\beta = \theta > 0$. (Note that the pdf of $X \sim \Gamma(\alpha, \beta)$ is $f(x) = x^{\alpha-1}e^{-x/\beta}/(\Gamma(\alpha)\beta^{\alpha})$ for $0 < x < \infty$. In additon $EX = \alpha\beta$ and $VarX = \alpha\beta^2$.)

a. Find the mle of θ.

b. Find the limiting distribution of $\sqrt{n}(\hat{\theta} - \theta)$, where $\hat{\theta}$ is the mle of θ.

c. Derive an asymptotic 95% CI for θ. Show all your work and justify each step in the derivation.

(a) Answer:

$L(\theta) = \prod_{i=1}^{n} p(x_i; \theta) = \prod x_i^3 e^{-x_i/\theta}/(6\theta^4)$

$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} 3\log x_i - x_i/\theta - \log 6 - 4\log\theta$

$\frac{dl(\theta)}{d\theta} = \sum_{i=1}^{n} x_i/\theta^2 - 4/\theta = 0$

$\hat{\theta}_{MLE} = \frac{\sum_{i=1}^{n} x_i}{4n} = \frac{\bar{x}}{4}$

(b) Answer:

$\sqrt{n}(\bar{X} - E(X)) = \sqrt{n}(\bar{X} - 4\theta) \xrightarrow{d} N(0, 4\theta^2)$

$\sqrt{n}(\frac{\bar{X}}{4} - \theta) \xrightarrow{d} N(0, \frac{\theta^2}{4})$

$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \frac{\theta^2}{4})$

(c) Answer:

$SE(\hat{\theta}) = \sqrt{\frac{Var(\hat{\theta})}{n}} == \sqrt{\frac{\hat{\theta^2}}{4n}}$

$CI = (\hat{\theta} - 1.96SE(\hat{\theta}), \hat{\theta} + 1.96SE(\hat{\theta})) = (\hat{\theta} - 1.96\sqrt{\frac{\hat{\theta^2}}{4n}}, \hat{\theta} + 1.96\sqrt{\frac{\hat{\theta^2}}{4n}}) = (\frac{\bar{x}}{4} - 1.96\sqrt{\frac{\bar{x}^2}{64n}}, \hat{\theta} + 1.96\sqrt{\frac{x^2}{64n}})$