11/27 Kengo Kato

STSCI 5080 Practice Final Exam

Instructions:

- The actual exam is 2 and half hours long. In the exam, you can bring notes and textbooks, but can not bring any electronic device.
- When you are asked to find a confidence interval with (asymptotic) level $1-\alpha$, you need to find a confidence interval whose coverage probability is (approaching) $1-\alpha$, where $0<\alpha<1$ (e.g., $\alpha=0.05$). Likewise, when you are asked to find a test with asymptotic level α, you need to find a test whose probability of the type I error is (approaching) α.
- You may use 1.96 for the 97.5%-quantile of the $N(0,1)$ distribution.

Problem 1. Circle the correct choice in each of the following questions.

(1) Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$ for some $-\infty < \mu < \infty$ and $\sigma^2 > 0$, where $n \geq 2$. The sample mean and variance are defined by $\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$ and $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$, respectively. What is the distribution of $\overline{X}$?

a. $N(\mu, \sigma^2)$ b. $N(\mu, \sigma^2/(n-1))$ c. $N(\mu, \sigma^2/n)$ d. $N(0,1)$

(2) In the previous question, what is the distribution of $(n-1)S^2/\sigma^2$?

a. $\chi^2(n-1)$ b. $\chi^2(n)$ c. $\chi^2(n+1)$ d. $N(\mu, \sigma^2/n)$

(3) Let $X_1, \ldots, X_n \sim Ex(\lambda)$ i.i.d. for some $\lambda > 0$. What is the likelihood function for λ?

a. $\lambda e^{-\lambda \sum_{i=1}^{n} X_i}$ b. $\lambda^n e^{-\lambda \sum_{i=1}^{n} X_i}$ c. $e^{-\lambda \sum_{i=1}^{n} X_i}$ d. $\lambda^n e^{-\lambda^n \sum_{i=1}^{n} X_i}$

(4) Suppose that you are given a statistical model $\{f_\theta \mid -\infty < \theta < \infty\}$, and you have an estimator $\widehat{\theta}_n$ such that $\widehat{\theta}_n \xrightarrow{P} \theta/2$ as $n \to \infty$ for any value of θ. Find a consistent estimator for θ from the following choices. Only one of them is consistent.

a. $\widehat{\theta}_n$ b. $10\sqrt{\widehat{\theta}_n}$ c. $2\widehat{\theta}_n$ d. $4\widehat{\theta}_n$

(5) You observe the numbers of failures occurring in machines in a factory in a year. Denote by X_i the number of failures of the i-th machine, and suppose that

$$X_1, \ldots, X_n \sim Po(\lambda) \text{ i.i.d.}$$

where n is the number of machines in the factory. If the MLE is $\widehat{\lambda} = 0.4$, what is the MLE of $P_\lambda(X_1 = 0)$?

a. $e^{-0.4}$ b. $0.4e^{-0.4}$ c. $e^{0.4}$ d. $0.4e^{0.4}$

(6) Suppose that the heights of women in a certain population follow the normal distribution with mean μ and variance 9. If the sample size $n = 36$ and the sample mean $\overline{X}$ are given, then what is a confidence interval for μ whose coverage probability is 95%?

a. $[\overline{X}-0.49, \overline{X}+0.49]$ b. $[\overline{X}-0.98, \overline{X}+0.98]$ c. $[\overline{X}-1.96, \overline{X}+1.96]$ d. $[\overline{X}-2.45, \overline{X}+2.45]$

Problem 2. Let f_θ be a pmf of the form

$$f_\theta(x) = \begin{cases} \frac{2}{3}\theta & \text{if } x = 0 \\ \frac{1}{3}\theta & \text{if } x = 1 \\ \frac{2}{3}(1-\theta) & \text{if } x = 2 \\ \frac{1}{3}(1-\theta) & \text{if } x = 3 \\ 0 & \text{elsewhere} \end{cases},$$

where $0 < \theta < 1$. We observe a random sample

$$(X_1, X_2, \ldots, X_{10}) = (3, 0, 2, 1, 3, 2, 1, 0, 2, 1) \tag{*}$$

from such a pmf.

(a) Find the joint pmf of $(X_1, X_2, \ldots, X_{10})$ at (*). You need to derive an explicit form of the joint pmf.

(b) Find the MLE of θ at (*).

Problem 3. Let

$$X_1, \ldots, X_n \sim N(0, \theta) \text{ i.i.d.}$$

where $\theta > 0$ is unknown.

(a) Find the log likelihood function for θ. You need to derive an explicit form of the log likelihood function.

(b) Verify that the MLE of θ is $\widehat{\theta} = n^{-1} \sum_{i=1}^{n} X_i^2$.

Problem 3 (continued).

(c) Find the limiting distribution of $\sqrt{n}(\widehat{\theta} - \theta)$ as $n \to \infty$. You may use the fact that $E_\theta(X_1^2) = \theta$ and $E_\theta(X_1^4) = 3\theta^2$ without derivations.

(d) Find a confidence interval for θ with asymptotic level 95% by estimating the asymptotic variance given in Part (c). You don't need any derivation in this question and so only have to state your confidence interval.

Problem 3 (continued).

(e) Show that $\sqrt{n}\left(\frac{1}{\sqrt{2}}\log\widehat{\theta} - \frac{1}{\sqrt{2}}\log\theta\right) \xrightarrow{d} N(0,1)$ as $n \to \infty$.

(f) Find, with a brief justification, a confidence interval for θ with asymptotic level 95% using the result of Part (e).

Problem 4. Let $X_1, \dots, X_n \sim Po(\lambda)$ i.i.d. where $\lambda > 0$ is unknown. You know that the MLE of λ is $\widehat{\lambda} = \overline{X} = n^{-1} \sum_{i=1}^{n} X_i$ and $\sqrt{n}(\widehat{\lambda} - \lambda) \xrightarrow{d} N(0, \lambda)$ as $n \to \infty$. Consider the following testing problem:

$$H_0 : \lambda = 2 \quad \text{vs.} \quad H_1 : \lambda \neq 2.$$

Find a test based on the MLE that has asymptotic level 5%. Justify that your test has asymptotic level 5%.

Problem 5. Suppose that you are a market researcher interested in the proportion p of a population that will buy a product, and ask n potential customers about their willingness to buy the product. Discuss how to determine the sample size n of the survey using a confidence interval for p. You need to derive an explicit way to determine the sample size.

(Comment). Summarize Lecture 22 slides pages 13–15. I will ask the same question in the actual exam.