Exam Simulation 2024

Statistics: Maura Mezzetti

Exercise

Let $(X_1, \ldots, X_n)$ be a random sample of i.i.d. random variables. Let $f_\theta(x)$ and $F_\theta(x)$ be the density function and cumulative distribution function respectively. Let $\hat{\theta}$ be the MLE of θ, θ_0 be the true parameter, $L(\theta)$ be the likelihood function, $logL(\theta)$ be the loglikelihood function, and $I(\theta)$ be the Fisher information matrix

1. Suppose $f_\theta(x)$ is the probability distribution function of a Possion with parameter λ. The Fisher information is

   (a) $I(\theta) = \frac{1}{\lambda}$
   (b) $I(\theta) = \lambda$
   (c) $I(\theta) = \frac{1}{\lambda^2}$
   (d) $I(\theta) = \lambda^2$

2. Suppose $\theta = (\theta_1, \theta_2)$, given the system of hypotheses: $H_0 : \theta_1 = 0$ and $\theta_2 = 2$ versus $H_1 : \theta_1 \neq 0$ and $\theta_2 \neq 2$, the LR test statistics $[-2log\Lambda]$ has an asymptotic distribution that is

   (a) $\chi(1)$, chi-squared with one degree of freedom,
   (b) $\chi(2)$ chi-squared with two degrees of freedom,
   (c) $F(2, n-2)$
   (d) $N(0, 1)$

3. Which statement is true about confidence intervals?

   (a) If we construct two 95% confidence intervals for a population mean μ (population with a Gaussian distribution with unknown variance) based on two different random samples with different sample size, the interval from a random sample with bigger sample size always gives a narrower confidence interval.
   (b) A confidence interval is an interval of values computed from sample data that is likely to include the true population parameter value.

(c) A confidence interval between 20% and 40% means that the population proportion definitely lies between 20% and 40%.

(d) A 99% confidence interval procedure has a lower probability of producing intervals that will include the population parameter than a 95% confidence interval procedure.

4. Assume the regularity conditions hold, which of the following term is zero?

   (a) $E(\hat{\theta} - \theta_0)$

   (b) $E(logL'(\theta_0))$

   (c) $Var(\hat{\theta})$

   (d) $E(L'(\theta_0))$

Exercise 2

Let $(X_1, \ldots, X_n)$ be independent identically distributed random variables with p.d.f.

$$f(x; \beta, \alpha) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \qquad x > 0$$

1. Find a sufficient statistics for (β, α)

Consider parameter α known and equal to 4:

1. Find $\hat{\beta}_{MLE}$ maximum likelihood estimator (MLE) for β

2. Compute the score function and the Fisher information.

3. Specify asymptotic distribution of $\hat{\beta}_{MLE}$.

4. Suppose form a random sample of 200 random variables you obtain $\sum_{i=1}^{200} x_i = 800$. Complete ALL of the following questions:

   (a) Find the Likelihood Ratio test statistic for testing $H_0 : \beta = 1.2 \quad versus \quad H_1 : \beta \neq 1.2$, specify the asymptotic distribution and verify the null hypothesis, with $\alpha = 0.05$.

   (b) Find the Wald test statistic for testing $H_0 : \beta = 1.2 \quad versus \quad H_1 : \beta \neq 1.2$, specify the distribution and verify the null hypothesis, with $\alpha = 0.05$.

   (c) Find the Score test statistic for testing $H_0 : \beta = 1.2 \quad versus \quad H_1 : \beta \neq 1.2$, specify the distribution and verify the null hypothesis, with $\alpha = 0.05$.

Exercise 3

Eggs are thought to be infected with a bacterium salmonella enteriditis so that the number of organisms, Y , in each has a Poisson distribution with mean μ. The value of Y cannot be observed directly, but after a period it becomes certain whether the egg is infected ($Y > 0$) or not ($Y = 0$). Out of n such eggs, r are found to be infected. Find the maximum likelihood estimator of μ and its asymptotic variance.

Exercise 4

The Economist collects data each year on the price of a Big Mac in various countries around the world. A sample of McDonald's restaurants in Europe in July 2018 resulted in the following Big Mac prices (after conversion to U.S. dollars).

4.44, 3.94, 2.40, 3.97, 4.36, 4.49, 4.19, 3.71, 4.61, 3.89

Assuming that the price of a Big Mac, X, is well modeled by a normal distribution

- Compute an estimate of $P(X < 4.2)$.

Exercise 5

1. Provide correct statement for Neyman Pearson Lemma
2. Provide correct statement for Factorization Theorem
3. Provide correct statement for Cramér-Rao inequality
4. Provide correct statement for Likelihood Principle and Describe method to reach Maximum Likelihood Estimation
5. Describe method to moments estimation
6. Provide correct statement for Likelihood ratio test
7. Define pivot quantities and explain their use for confidence intervals
8. Define and compare the three large sample tests (Wald, Score and Likelihood Ratio Test)