Stat. Inf. Open book part

January 19, 2017

Q1.

Consider the following regression model:

$$Y_i = f(x_i) + \epsilon_i, \text{ for } i = 1, \dots, n \qquad \text{(M)}$$

where $n \in \mathbb{N}$, $x_1, \dots, x_n$ are known constants and $\epsilon_1, \dots, \epsilon_n$ are i.i.d. with $E(\epsilon_1) = 0$ and $Var(\epsilon_1) = \sigma^2$ for $\sigma^2 > 0$ unknown. Moreover, we assume that $\min_{1 \le i \le n} x_i \le \frac{1}{2}$ and $\max_{1 \le i \le n} x_i > \frac{1}{2}$.
We know that the unknown function $f : \mathbb{R} \to \mathbb{R}$ is linear on the intervals $\left(-\infty, \frac{1}{2}\right]$ and $\left(\frac{1}{2}, +\infty\right]$. At $x = \frac{1}{2}$, the function may be discontinuous. Furthermore, we know that $f(0) = 0$, and $f(1) = 1$. We want to estimate the function f.

(a).

Describe the problem (M) as a linear regression model using matrix notations, and define every notation you use.
Hint: Verify that there exist (β_1, β_2) such that

$$f(x) = \begin{cases} \beta_1 x & \text{for } x \le \frac{1}{2}, \\ \beta_2 + (1 - \beta 2)x & \text{for } x > \frac{1}{2}. \end{cases}$$

(b).

Find the LSE $\hat{\beta}_1$ and $\hat{\beta}_2$ for β_1 and β_2 from part (a).

(c).

Based on part (b), give $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_2)$.
Consider a sub-model (M0), where the function f is as above, yet continuous. That is, f is linear on $\left(-\infty, \frac{1}{2}\right]$ and $\left(\frac{1}{2}, +\infty\right]$, $f(0) = 0$, $f(1) = 1$, and continuous on $\mathbb{R}$.

(d).

Parametrize the function f, denoting its parameters by γ, and formulate the model using matrix notation. Define every notation you use.

(e).

Find the LSE of $\hat{\gamma}$ of γ and compute its variance $Var(\hat{\gamma})$.

(f).

Compare $\hat{\gamma}$ form part (e) with $(\hat{\beta}_1, \hat{\beta}_2)$ form part (b). Comment on the results.

(g).

Construct a test for the problem $H_0 : (M0)$ against $H_1 : (M)$ of size 0.05.

(h).

Briefly discuss if additional assumptions are needed for the test from part (g).

Q2.

Let $X_1, \ldots, X_n$ be a random sample from X having probability density function

$$f_X(x;\theta) = \frac{1}{2}(1-\theta^2)\exp(\theta x - |x|), \text{for } x \in \mathbb{R}$$

with $\theta \in \Theta = (-1,1)$ is unknown and where $E(X) = \frac{2\theta}{1-\theta^2}$.

(a).

Prove that the MLE $\hat{\theta}$ of θ is

$$\hat{\theta} = \frac{-1+\sqrt{1+\bar{X}^2}}{\bar{X}},$$

where $\bar{X}$ denotes the population mean. In particular, verify that $\hat{\theta}$ is in Θ.

Hint: the functions

$$h_1 : t \mapsto \frac{-1-\sqrt{1+t^2}}{t}, h_2 : t \mapsto \frac{-1+\sqrt{1+t^2}}{t}$$

are strictly increasing on their domains.

(b).

State an asymptotic normality result for $\hat{\theta}$.

(c).

Find the estimator of θ using the method of moments, and compare it with the MLE from (a).

(d).

Based on the MLE, construct an (approximate) $100(1-\alpha)\%$ confidence interval for θ.

(e).

Use (b) to construct an (approximate) $100(1-\alpha)\%$ confidence interval for $E(X)$.

(f).

Construct the UMP test of size α for the testing problem $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$. Approximate the distribution of the test statistic using the result from part (e).