Exercise 6: Statistical inference (III)

Siyue Yang

06/01/2022

Part 1: Wald, Score, and likelihood ratio test statistics

Write out the likelihood function, and derive the test statistics of the Wald, Score, and likelihood ratio test.

1. $X_i \overset{\text{i.i.d.}}{\sim} f(x \mid \theta)$

$$f(x \mid \theta) = \theta \exp(-x\theta)\mathbb{I}\{x > 0\}$$

2. $X_i \overset{\text{i.i.d.}}{\sim} f(x \mid \theta)$

$$f(x \mid \theta) = \theta c^{\theta} x^{-(\theta+1)}\mathbb{I}\{x > c\} \quad \text{(Pareto distribution)}$$

   where c is a known constant and θ is unknown.

Solution

1. The log-likelihood function is

$$l(\theta) = n\left(\log\theta - \theta\bar{X}_n\right)$$

   which yields

$$l'(\theta) = n\left(\frac{1}{\theta} - \bar{X}_n\right) \text{ and } l''(\theta) = -\frac{n}{\theta^2}$$

   The MLE $\hat{\theta}_n$ is obtained by setting $l'(\theta) = 0$,

$$\widehat{\theta}_n = \frac{1}{\bar{X}_n}$$

   and the Fisher information can be obtained by

$$I(\theta) = \theta^{-2}$$

   It follows that,

$$\begin{aligned} W_n &= \frac{\sqrt{n}}{\theta_0}\left(\frac{1}{\bar{X}_n} - \theta_0\right) \\ R_n &= \theta_0\sqrt{n}\left(\frac{1}{\theta_0} - \bar{X}_n\right) = \frac{W_n}{\theta_0\bar{X}_n} \\ \Delta_n &= n\left\{\bar{X}_n\left(\bar{X}_n - \theta_0\right) - \log\left(\theta_0\bar{X}_n\right)\right\} \end{aligned}$$

2. If let $S_n = \sum_{i=1}^n \log(x_i)$, the log-likelihood function is

$$l(\theta) = n(\log\theta + \theta\log c) - (\theta+1)S_n$$

which yields,

$$l'(\theta) = n\left(\frac{1}{\theta} + \log c\right) - S_n, \quad l''(\theta) = -\frac{n}{\theta^2}$$

The MLE can be obtained as

$$\widehat{\theta}_n = \frac{n}{S_n - n\log c}$$

and Fisher information as

$$I(\theta) = \theta^{-2}$$

Thus, the three test statistics are

$$\begin{aligned}
W_n &= \frac{\sqrt{n}}{\theta_0}\left(\frac{n}{S_n - n\log c} - \theta_0\right) \\
R_n &= \sqrt{n}\theta_0\left(\left(\frac{1}{\theta_0} + \log c\right) - S_n\right) \\
\Delta_n &= n\left(\log\frac{\widehat{\theta}_n}{\theta_0} + \left(\widehat{\theta}_n - \theta_0\right)\log c\right) - \left(\widehat{\theta}_n - \theta_0\right)S_n
\end{aligned}$$

Part 2: Test equivalence

Let θ be a scalar parameter and suppose we test

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0.$$

Let W be the Wald test statistic and let λ be the likelihood ratio test statistic. Show that these tests are equivalent in the sense that

$$\frac{W^2}{\lambda} \xrightarrow{\text{P}} 1$$

as $n \to \infty$. Hint: Use a Taylor expansion of the log-likelihood $\ell(\theta)$ to show that

$$\lambda \approx \left(\sqrt{n}\left(\widehat{\theta} - \theta_0\right)\right)^2\left(-\frac{1}{n}\ell''(\widehat{\theta})\right)$$

Solution

Throughout this proof, it is assumed that the density $f(x;\theta)$ appearing in the likelihood is sufficiently regular. A Taylor expansion reveals

$$\ell(\theta_0) = \ell(\hat{\theta}) + \left(\hat{\theta} - \theta_0\right)\ell'(\hat{\theta}) + \frac{1}{2}\left(\hat{\theta} - \theta_0\right)^2\ell''(\hat{\theta}) + O\left(\left(\hat{\theta} - \theta_0\right)^3\right).$$

Note, in particular, that $\ell'(\hat{\theta}) = 0$ since $\hat{\theta}$ is an MLE. Therefore,

$$\lambda = 2\log\left(\frac{\mathcal{L}(\hat{\theta})}{\mathcal{L}(\theta_0)}\right) = -\left(\hat{\theta} - \theta_0\right)^2\ell''(\hat{\theta}) + O\left(\left(\hat{\theta} - \theta_0\right)^3\right)$$

Moreover,

$$W^2 = \frac{\left(\hat{\theta} - \theta_0\right)^2}{\widehat{\text{se}}(\hat{\theta})^2} = nI(\hat{\theta})\left(\hat{\theta} - \theta_0\right)^2$$

It follows that

$$\frac{\lambda}{W^2} = \frac{n^{-1}\ell''(\hat{\theta})}{-I(\hat{\theta})} + O\left(\hat{\theta} - \theta_0\right)$$

Under the null hypothesis, $\hat{\theta} \xrightarrow{P} \theta_0$. Therefore, by two applications of Slutsky theorem, $1/I(\hat{\theta}) \to 1/I\left(\theta_0\right)$ where

$$I\left(\theta_0\right) = \mathbb{E}_{\theta_0}\left[\frac{\partial^2 \log f\left(X; \theta_0\right)}{\partial \theta^2}\right]$$

Since

$$\ell''(\theta) = \sum_n \frac{\partial^2 \log f\left(X_n; \theta\right)}{\partial \theta^2}$$

by the weak law of large numbers, $n^{-1}\ell''(\hat{\theta}) \xrightarrow{P} I\left(\theta_0\right)$ under the null hypothesis. The result now follows by Slutsky theorem.

Part 3: Omics

The p-value is uniformly distributed when the null hypothesis is true.

Let T denote the random variable with cumulative distribution function $F(t) \equiv \Pr(T < t)$ for all t. Assuming that F is invertible we can derive distribution of the random p-value $P = F(T)$ as follows:

$$\Pr(P < p) = \Pr(F(T) < p) = \Pr\left(T < F^{-1}(p)\right) = F\left(F^{-1}(p)\right) = p,$$

from which we can conclude that the distribution of P is uniform on $[0, 1]$.