M3S3/M4S3 STATISTICAL THEORY II

WORKED EXAMPLE: TESTING FOR THE PARETO DISTRIBUTION

Suppose that $X_1, \ldots, X_n$ are i.i.d random variables having a Pareto distribution with pdf

$$f_{X|\theta}(x|\theta) = \frac{\theta c^\theta}{x^{\theta+1}} \qquad x > c$$

and zero otherwise, for known constant $c > 0$, and parameter $\theta > 0$.

(i) Find the ML estimator, $\widehat{\theta}_n$, of θ, and find the asymptotic distribution of

$$\sqrt{n}(\widehat{\theta}_n - \theta_T)$$

where θ_T is the true value of θ.

(ii) Consider testing the hypotheses of

$$\begin{aligned} H_0 &: \theta = \theta_0 \\ H_1 &: \theta \neq \theta_0 \end{aligned}$$

for some $\theta_0 > 0$. Determine the likelihood ratio, Wald and Rao tests of this hypothesis.

SOLUTION (i) The ML estimate $\widehat{\theta}_n$ is computed in the usual way:

$$L_n(\theta) = \prod_{i=1}^n f_{X|\theta}(x_i|\theta) = \prod_{i=1}^n \frac{\theta c^\theta}{x_i^{\theta+1}} = \frac{\theta^n c^{n\theta}}{s_n^{\theta+1}}$$

where $s_n = \prod_{i=1}^n x_i$. Then

$$\begin{aligned} l_n(\theta) &= n\log\theta + n\theta\log c - (\theta+1)\log s_n \\ \dot{l}_n(\theta) &= \frac{n}{\theta} + n\log c - \log s_n \end{aligned}$$

and solving $\dot{l}_n(\theta) = 0$ yields the ML estimate

$$\widehat{\theta}_n = \left[\frac{\log s_n}{n} - \log c\right]^{-1} = \left[\frac{1}{n}\sum_{i=1}^n \log x_i - \log c\right]^{-1}.$$

The corresponding estimator is therefore

$$\widehat{\theta}_n = \left[\frac{1}{n}\sum_{i=1}^n \log X_i - \log c\right]^{-1}.$$

Computing the asymptotic distribution directly is difficult because of the reciprocal. However, consider $\phi = 1/\theta$; by invariance, the ML estimator of ϕ is

$$\widehat{\phi}_n = \frac{1}{n}\sum_{i=1}^n \log X_i - \log c = \frac{1}{n}\sum_{i=1}^n (\log X_i - \log c)$$

which implies how we should compute the asymptotic distribution of $\widehat{\theta}_n$ - we use the CLT on the random variables $Y_i = \log X_i - \log c = \log(X_i/c)$, and then use the Delta Method.

To implement the CLT, we need the expectation and variance of $Y = \log(X/c)$. Now

$$F_X(x) = 1 - \left(\frac{c}{x}\right)^\theta \qquad x > c$$

so that

$$F_Y(y) = P[Y \le y] = P[\log(X/c) \le y] = P[X \le c\exp\{y\}] = 1 - \exp\{-\theta y\} \qquad y > 0$$

and hence $Y \sim \mathit{Exponential}(\theta)$. By standard results

$$E_{f_Y}[Y] = \frac{1}{\theta} = \phi \qquad Var_{f_Y}[Y] = \frac{1}{\theta^2} = \phi^2,$$

and therefore, by the CLT,

$$\sqrt{n}(\widehat{\phi}_n - \phi_T) \xrightarrow{\mathfrak{L}} N(0, \phi_T^2)$$

where $\phi_T = 1/\theta_T$.

Finally, let $g(t) = 1/t$ so that $\dot{g}(t) = -1/t^2$. Then, by the Delta Method

$$\sqrt{n}(g(\widehat{\phi}_n) - g(\phi_T)) \xrightarrow{\mathfrak{L}} N(0, \{\dot{g}(\phi_T)\}^2\phi_T^2)$$

so that, as $g(\phi_T) = 1/\phi_T = \theta_T$

$$\sqrt{n}(\widehat{\theta}_n - \theta_T) \xrightarrow{\mathfrak{L}} N(0, \{1/\phi_T^4\}\phi_T^2) \equiv N(0, \theta_T^2).$$

(ii) For the **likelihood ratio** test:

$$\lambda_n = 2\log\frac{L_n(\widehat{\theta}_n)}{L_n(\theta_0)} = 2\log\frac{\widehat{\theta}_n^n c^{n\widehat{\theta}_n}/s_n^{\widehat{\theta}_n+1}}{\theta_0^n c^{n\theta_0}/s_n^{\theta_0+1}} = 2n\left[\log(\widehat{\theta}_n/\theta_0) + (\widehat{\theta}_n - \theta_0)\log c - (\widehat{\theta}_n - \theta_0)m_n\right]$$

where $m_n = (\log s_n)/n$.

For the **Wald** test:

$$W_n = n(\widehat{\theta}_n - \theta_0)I(\widehat{\theta}_n)(\widehat{\theta}_n - \theta_0) = nI(\widehat{\theta}_n)(\widehat{\theta}_n - \theta_0)^2$$

where $I(\theta)$ is the Fisher information for this model. From first principles

$$\begin{aligned} l(\theta) &= \log\theta + \theta\log c - (\theta+1)\log x \\ \dot{l}(\theta) &= \frac{1}{\theta} + \log c - \log x \\ \ddot{l}(\theta) &= -\frac{1}{\theta^2} \end{aligned}$$

so $I(\theta) = 1/\theta^2$; in fact, we could have deduced this from the results derived in (i). Thus

$$W_n = n\left(\frac{\widehat{\theta}_n - \theta_0}{\widehat{\theta}_n}\right)^2.$$

Note that, although the Fisher Information is available in this case, it may be statistically advantageous in a finite sample case to replace $I(\widehat{\theta}_n)$ by $\widehat{I}_n(\widehat{\theta}_n)$, derived in the usual way from the first or second derivatives of l_n. That is, we might use

$$\widehat{I}_n(\widehat{\theta}_n) = \frac{1}{n}\sum_{i=1}^{n} S(x_i,\theta)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\widehat{\theta}_n} + \log c - \log x_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2$$

where $y_i = \log x_i - \log c$. Alternately, using the second derivative,

$$\widehat{I}_n(\widehat{\theta}_n) = -\frac{1}{n}\sum_{i=1}^{n}\ddot{l}_i(\theta) = \frac{1}{n}\sum_{i=1}^{n} 1/\widehat{\theta}_n^2 = 1/\widehat{\theta}_n^2 = I(\widehat{\theta}_n).$$

For the **Rao** test: in the single parameter case

$$R_n = Z_n^{\mathsf{T}}[I(\theta_0)]^{-1}Z_n = Z_n^2/I(\theta_0)$$

where

$$Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} S(x_i,\theta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{1}{\theta_0} + \log c - \log x_i\right)$$

so that if $y_i = \log x_i - \log c$ as before

$$R_n = \frac{\left\{\sum_{i=1}^{n}(y_i - 1/\theta_0)\right\}^2}{nI(\theta_0)} = \frac{\left\{\sum_{i=1}^{n}(y_i - 1/\theta_0)\right\}^2}{n/\theta_0^2}.$$

SUPPLEMENTARY EXERCISES: Suppose that c is also an unknown parameter. Find

- the ML estimator for c, $\widehat{c}$
- the weak-law (ie probability) limit of $\widehat{c}$
- an asymptotic (large n) approximation to the distribution of $\widehat{c}$.
- the profile likelihood for θ.

*Hint: Recall, when considering the likelihood for c, that $x_i > c$ **for all** i. Then, think back to M2S1 Chapter 5, and extreme order statistics.*