COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

25,4 (1984)

ALTERNATIVE PARAMETRIC ESTIMATION IN THE EXPONENTIAL CASE UNDER RANDOM CENSORSHIP

Jan HURT

Abstract: An alternative parametric estimator for the parameter of the exponential distribution under random censorship is suggested and its asymptotic properties are derived. The proposed estimator has a smaller bias than the usual maximum likelihood estimator.

Key words and phrases: Parametric estimation, exponential distribution, random censorship, asymptotic normality, asymptotic expansions of moments.

Classification: Primary: 62F10, 62N05
Secondary: 62E20

Both in medical research and reliability testing we often deal with incomplete or censored observations. A useful tool for statistical inference in such situation is the model of random censorship.

Suppose that X is a random variable representing lifetime or time to failure. Together with this random variable consider random variable T which represents time censor. On any object which is put on trial, we can observe either X or T depending on what has happened sooner. We also get information whether we have observed X or T. Observation of n objects thus results in pairs

$$(W_1, I_1), \dots, (W_n, I_n)$$

where $W_j = \min(X_j, T_j)$, $I_j = 1$ if $X_j < T_j$ (j^{th} observation is uncensored), and $I_j = 0$ in the opposite case (j^{th} observation is censored at time T_j), $j=1,\dots,n$.

A usual statistical problem is to estimate the distribution of X or, in case of the known analytical form of the distribution, to estimate its parameters. The most popular non-parametric estimator of the distribution function is the well-known Kaplan-Meier product limit estimator which has been intensively studied recently. Parametric estimators are usually based on the method of maximum likelihood.

In this note we deal with the simplest case when X and T are independent exponential random variables with expected values Θ and ω, respectively. Let us denote $\bar{W} = \Sigma W_j/n$ and $\bar{I} = \Sigma I_j/n$. The maximum likelihood estimator of Θ is

$$(1) \qquad \hat{\Theta} = \bar{W}/\bar{I}$$

provided that $\bar{I} > 0$. We leave $\hat{\Theta}$ undefined in the opposite case. Let us define σ by the relation

$$\frac{1}{\sigma} = \frac{1}{\Theta} + \frac{1}{\omega} \ .$$

It has been shown in [1] that $\hat{\Theta}$ is asymptotically normally distributed as $N(\Theta, \frac{\Theta^3}{n\sigma})$ and

$$(2) \qquad E(\hat{\Theta} \mid \Sigma I_j > 0) = \Theta + \frac{1}{n}\sigma\left(\frac{\Theta}{\sigma}\right)^2\left(1 - \frac{\sigma}{\Theta}\right) + O(n^{-2}).$$

Note that $\sigma/\Theta = P(X < T)$ expresses the expected proportion of uncensored observations. It is not difficult to see that the leading term of the bias is always positive and grows rapidly with increasing proportion of censored observations. Sometimes the bias of (1) may cause troubles. An alternative estimator of Θ is

(3) $\tilde{\theta} = \hat{\theta}\left(\frac{n}{n-1} - \frac{1}{(n-1)\,\bar{\bar{I}}}\right)$

which is unbiased up to $O(n^{-2})$. Asymptotic properties of (3) are summarized in the following Theorem.

Theorem. For $n \longrightarrow \infty$, we have

(4) $\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{D} N(0, \theta^3/\sigma)$,

(5) $E(\tilde{\theta} \mid \bar{I} > 0) = \theta + \frac{\theta}{n^2}\left[1 - \left(\frac{\theta}{\sigma}\right)^2\right] + O(n^{-3})$,

(6) $\mathrm{var}(\tilde{\theta} \mid \bar{I} > 0) = \frac{1}{n}\frac{\theta^3}{\sigma} + O(n^{-2})$.

Lemma. Let $g = g(t,n)$ be a real function defined on $R_1 \times N$. Suppose that g admits a continuous $(q+1)^{st}$ derivative with respect to t on $[\theta - \sigma, \theta + \sigma] \times N$ for some $\sigma > 0$. Suppose that $g'(\theta,n), \ldots, g^{(q)}(\theta,n)$ are bounded functions of n and $g^{(q+1)}(t,n)$ is bounded on $[\theta - \sigma, \theta + \sigma] \times N$. Suppose that $\{T_n\}_{n \in N}$ is a sequence of statistics and that there exist $s \geq q + 2$, $\eta > 0$ such that $E|T_n - \theta|^s = O(n^{-\eta s})$ and $|g(t,n)|^2 \leq C_1 + C_2 n^{\eta(s-q-2)}$ where C_1, C_2 are constants. Then

(7) $$E\, g(T_n,n) = g(\theta,n) + \sum_{j=1}^{q} \frac{1}{j!} g^{(j)}(\theta,n)\, E(T_n - \theta)^j + O(n^{-\eta(q+1)}),$$

(8) $$\mathrm{var}\, g(T_n,n) = \sum_{\substack{j=1 \\ j+k \leq q+1}}^{q} \sum_{k=1}^{q} \frac{1}{j!\,k!} g^{(j)}(\theta,n) g^{(k)}(\theta,n)\, \mathrm{cov}\,[(T_n - \theta)^j, (T_n - \theta)^k] + O(n^{-\eta(q+2)}).$$

Proof of the Lemma. The proof will be given in [2].

Proof of the Theorem. First we prove (4). We can write (3) as

$$(9)\quad \sqrt{n}(\tilde{\theta} - \theta) = \sqrt{n}\ \frac{n\bar{I} - 1}{(n-1)\bar{I}}(\hat{\theta} - \theta) + \frac{1}{\sqrt{n}}\ \frac{n\bar{I} - 1}{(n-1)\bar{I}}\ \theta\ \frac{\bar{I} - 1}{\bar{I} - 1/n}$$

It is obvious that

$$\frac{n\bar{I} - 1}{(n-+)\ \bar{I}} \xrightarrow{P} 1,$$

$$\frac{1}{\sqrt{n}}\ \frac{n\bar{I} - 1}{(n-1)\bar{I}}\ \theta \xrightarrow{P} 0,$$

$$\frac{\bar{I} - 1}{\bar{I} - 1/n} \xrightarrow{P} \frac{E\bar{I} - 1}{E\bar{I}}$$

so that the second term in (9) converges to 0 in probability. From Theorem 2 in [1] it follows that

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N(0, \theta^3/\sigma).$$

Using 2c.4 (x) in [3] now gives the result.

To prove (5) and (6) we make use of the Lemma. Since $\bar{W}$ and $\bar{I}$ are independent, and $E\bar{W} = \sigma$, we only need to calculate

$$E\left(\frac{n}{(n-1)\ \bar{I}} - \frac{1}{(n-1)\bar{I}^2}\right).$$

We have

$$\frac{n}{n - 1} = 1 + \frac{1}{n} + \frac{1}{n^2} + O(n^{-3}),$$

$$\frac{1}{n - 1} = \frac{1}{n} + \frac{1}{n^2} + O(n^{-3}).$$

Define

$$g(t) = t^{-1},\quad t \geq n^{-1},$$
$$= 0 \quad \text{otherwise},$$

and put $T_n = \bar{I}$ in the Lemma. Since $n\bar{I}$ possesses a binomial distribution with parameters n and σ/θ, we have $\eta = 1/2$. Thus we take $q = 5$. Calculation of $E(\bar{I}^{-1})$ is now a tedious but straightforward application of the Lemma. Similarly we handle with $E(\bar{I}^{-2})$ and analogously with var $\tilde{\theta}$.

R e f e r e n c e s

[1] HURT J.: On estimation in the exponential case under random censorship, Proc. of the 3^{rd} Pannonian Symp. on Math. Stat., Visegrád, Hungary (eds.:Mogyoródi J., Vincze,I., Wertz W.), Akadémiai Kiadó, Budapest(1982), 115-124.

[2] HURT J.: Asymptotic expansions for moments of functions of stochastic processes and their applications (1984). Submitted to: Statistics and Decision.

[3] RAO, C.R.: Linear statistical inference and its applications, Wiley, New York (1973), 2^{nd} edition.

Matematicko-fyzikální fakulta, Univerzita Karlova, Sokolovská 83, 186 00 Praha 8, Czechoslovakia

(Oblatum 31.5. 1984)