REVSTAT – Statistical Journal
Volume 13, Number 3, November 2015, 227–231

FAILURE-TIME WITH DELAYED ONSET

Authors: Man Yu Wong
– Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, China
mamywong@ust.hk

D.R. Cox
– Nuffield College, Oxford OX1 1NF, United Kingdom
david.cox@nuffield.ox.ac.uk

Received: February 2014 Accepted: May 2014

Abstract:

- A special distribution is suggested for the analysis of survival data in which there is a long random delay before the onset of the terminal process. Estimation by the method of moments and by maximum likelihood is compared.

Key-Words:

- *survival data; unobserved time origin; moment estimation; maximum likelihood estimation.*

AMS Subject Classification:

- 62F12, 62N02.

1. INTRODUCTION

Lai [1] has reviewed the rich variety of parametric families of distributions that have been suggested for the analysis of survival data. With the exception of the Weibull distribution and its connection with extreme value theory, most are essentially empirically-based flexible representations covering the rich variety of distributional shapes encountered in such contexts. Here we give a distribution extremely restricted in form but motivated by a very simple model of a data generating process.

A key element in the analysis of survival data is often the choice of time origin, for example, for patients, birth, entry into a study, first report of symptoms, etc. Appropriate choice may greatly clarify interpretation. Sometimes, however, the natural time origin is unobserved. In this paper we outline a very special model for such a situation. Each study individual has an observed time origin and we assume that for a long time following that there is no possibility of a critical event. Then an unobserved transition occurs and following that the critical event rate becomes high. If the processes of transition and occurrence arise in independent Poisson processes of respectively small rate ρ_1 and large rate ρ_2, it follows that the failure-time has the form $T = T_1 + T_2$, where T_1 and T_2 are independently exponentially distributed with rates ρ_1 and ρ_2, with $\rho_1 << \rho_2$. This inequality is crucial both for separate estimation of the two parameters and indeed for the interpretation of the model.

Then the probability density function of T can be written as

$$\frac{e^{-t/\mu_1} - e^{-t/\mu_2}}{\mu_1 - \mu_2} \ ,$$

where it is convenient to parameterize in terms of the means $\mu_i = 1/\rho_i$ for $i = 1, 2$.

2. METHODS OF ESTIMATION

One simple method of estimating the two parameters is from the first two moments. This requires solving the equations

$$\begin{aligned} \bar{t} &= \tilde{\mu}_1 + \tilde{\mu}_2 \ , \\ s_t^2 &= \tilde{\mu}_1^2 + \tilde{\mu}_2^2 \ , \end{aligned}$$

where $\bar{t}$ and s_t^2 are sample mean and sample variance of observed values of T and $\tilde{\mu}_1 > \tilde{\mu}_2$, thus defining the moment estimators. Provided that $1 \geq s_t^2/\bar{t}^2 \geq 1/2$, as dictated by the special construction of the distribution,

$$\tilde{\mu}_1 = \frac{\bar{t}}{2} + \frac{1}{2}\sqrt{(2s_t^2 - \bar{t}^2)}\ ,$$
$$\tilde{\mu}_2 = \frac{\bar{t}}{2} - \frac{1}{2}\sqrt{(2s_t^2 - \bar{t}^2)}\ .$$

Notionally more efficient estimates will be given by the method of maximum likelihood. This involves numerical solution of the following two equations:

$$\sum_{i=1}^{n} \frac{t_i \exp(-t_i/\hat{\mu}_1)}{\{\exp(-t_i/\hat{\mu}_1) - \exp(-t_i/\hat{\mu}_2)\}} = \frac{n\hat{\mu}_1^2}{\hat{\mu}_1 - \hat{\mu}_2}\ ,$$
$$\sum_{i=1}^{n} \frac{t_i \exp(-t_i/\hat{\mu}_2)}{\{\exp(-t_i/\hat{\mu}_1) - \exp(-t_i/\hat{\mu}_2)\}} = \frac{n\hat{\mu}_2^2}{\hat{\mu}_1 - \hat{\mu}_2}\ ,$$

with $\hat{\mu}_1 > \hat{\mu}_2$. Both sets of estimates are sensible only if the data are consistent with the constraints on the squared coefficient of variation implied by the model and if one mean is substantially greater than the other.

3. COMPARISON OF ESTIMATORS

To compare the methods of estimation we compute asymptotic variances. For the moment estimates we have by local linearization, the delta method, that

$$\operatorname{var}(\tilde{\mu}_1) = \frac{\mu_2^2\left(2\lambda^4 - 2\lambda^3 + 2\lambda^2 + 1\right)}{n(\lambda-1)^2}\ ,$$
$$\operatorname{var}(\tilde{\mu}_2) = \frac{\mu_2^2\left(\lambda^4 + 2\lambda^2 - 2\lambda + 2\right)}{n(\lambda-1)^2}\ ,$$

where $\lambda = \mu_1/\mu_2$.

For the maximum likelihood estimates, the inverse of the Fisher information matrix gives

$$\operatorname{var}(\hat{\mu}_1) = \frac{\mu_2^2\lambda^4\left\{2\lambda^3\zeta - (\lambda-1)^2\right\}}{2n\left\{\lambda^3(\lambda^2+1)\zeta - (\lambda^2-\lambda+1)(\lambda-1)^2\right\}}\ ,$$
$$\operatorname{var}(\hat{\mu}_2) = \frac{\mu_2^2\{2\lambda^3\zeta + (\lambda-1)^2(\lambda^4 - 2\lambda^3 + 2\lambda - 2)\}}{2n\left\{\lambda^3(\lambda^2+1)\zeta - (\lambda^2-\lambda+1)(\lambda-1)^2\right\}}\ ,$$

where ζ is the generalized Riemann zeta function

$$\zeta\left[3,\, 2 + 1/(k-1)\right] = \sum_{k=0}^{\infty}\left\{k + 2 + 1/(k-1)\right\}^{-3}\ .$$

The efficiencies of the moment estimators for μ_1 and μ_2 depend only on the ratio of the two means, i.e., on λ. The efficiencies decrease as λ increases.

4. NUMERICAL COMPARISON

These results have been explored by simulation. Without loss of generality we set $\mu_2 = 1$ and generate 1000 sets of data for sample sizes 100 and 500 and for $\lambda = 2, 5, 10$. When $n = 100$ and $\lambda = 2$, approaching one-half the samples fall outside the range of validity of the above results, that is, are descriptively inconsistent with the model. This falls to about 15% when $n = 500$. In the more realistic case of larger λ, the incompatible samples are rare and theoretical and empirical variances agree reasonably closely, the variance of the maximum likelihood estimator being appreciably smaller than that of the moment estimate as λ becomes larger. When $\lambda = 10$, the asymptotic efficiency of $\tilde{\mu}_1$ is close to 50%, its limiting value for large λ, whereas that of $\tilde{\mu}_2$ is 25%, dropping slowly to its limiting value of zero.

5. DISCUSSION

The model could be generalized in various ways, for example to include uninformative censoring or a more complex transition process. With such generalizations simple estimation by the method of moments would typically not be possible. If the representation is plausible on general grounds and fits the data it would be very desirable to find a different type of observation predictive of the origin of the second component and study of such a marker would lead to a further generalization of the present model, which will, however, not be discussed here. If additional explanatory variables were available maximum likelihood estimation is likely to be needed. Finally note that the identification of the later stage parameter μ_2 with the smaller of the two estimates depends entirely on the prior specification.

A quite different approach to this kind of data is investigated in as yet unpublished paper by Peter McCullagh, University of Chicago to whom we are grateful for comments on an earlier version of the present note.

REFERENCES

[1] Lai, C.D. (2013). Constructions and applications of lifetime distributions, *Applied Stochastic Models in Business and Industry*, **29**, 127–140.