Far East Journal of Theoretical Statistics
Volume 32, Number 2, 2010, Pages 175-182
Published Online: October 28, 2010
This paper is available online at http://pphmj.com/journals/fjts.htm

THE SURVIVAL DISTRIBUTION BEYOND A PARTICULAR AGE

JOHN E. GOULIONIS*, P. MATZIOROS and D. NIKOLAKIS

Department of Statistics and Insurance Science
University of Piraeus
80 Karaoli and Dimitriou Street
18534 Piraeus, Greece
e-mail: jgouli@unipi.gr
goulionisyannis@yahoo.com
sondi@otenet.gr

Abstract

Old-age mortality for populations of developed countries has been improving rapidly since 1950's. We need a method that can extrapolate a survival distribution to extreme ages without requiring accurate mortality data for the centenarian population. More specifically, we use the asymptotic distribution of the excesses over threshold to model the survival distribution beyond a particular age. This age, which is known as the threshold age, is chosen to ensure that the tail of the fitted distribution is consistent with the parametric graduation for earlier ages. We estimate finally, the parameter of this model.

1. Introduction

The asymptotic behavior of the residual life time is investigated $(\text{for } t \to \infty)$. Balkema and Haan [1] gives fairly complete answers for the asymptotic behavior of residual life time distributions for $t \to \infty$. For example, we consider a light bulb. It

2010 Mathematics Subject Classification: 62C10, 91B30, 91B16.

Keywords and phrases: extreme value theory, old-age mortality, Pareto distribution.

*Corresponding author

Received May 13, 2010

has a certain life time X, which is a random variable. After having burned t hours there remains a residual life time. It is considerable interest to know the asymptotic behavior of these residual life time distributions. Natural questions: What are the possible limit distribution types? For each such limit distribution G, what is the domain of attraction (the set of all distribution functions, suitably normed, converges to G)? What is the speed of convergence?

In this paper, we consider a general situation. The paper is organized as follows: In Section 2, we give the model specification and in Section 3, we shall derive the possible limit distribution a (generalized Pareto distribution) and we shall estimate the parameters. Also, in Section 4, we provide a length-biased distribution for the real life application.

2. Model Specification

Let us define the following notation:

- X: The age at death random variable, assumed to be continuous.
- $f(X)$: The probability density function of the continuous random variable X.
- $F(X)$: The distribution function of the continuous random variable X.
- $S(X) = 1 - F(X)$: The survival function of the continuous random variable X.
- $\mu(X) = \dfrac{f(X)}{1 - F(X)}$: The force of mortality for the continuous random variable X.
- d_x: The number of deaths between the ages x and $x + 1$.
- E_x: The number of exposures-to-risk between the ages x and $x + 1$, in practice, E_X is approximated by the mid-year population at age x.
- l_x: The number of survivors to the age x.
- $m_x = \dfrac{d_x}{E_x}$: The central rate of death at the age x.
- $q_x = \dfrac{d_x}{e_x}$: The probability of death between the ages x and $x + 1$, conditioning on survival to the age x.

Let $Z = \{Y - d \mid Y > d\}$ be the conditional excess of Y over a threshold d. The Balkema-De Haan-Pisckands theorem (Balkema and De Haan [1]) states that, under certain regular conditions, the limiting distribution of Z is a generalized Pareto distribution, as the threshold d approaches the right-hand end support of Y. This important result, in extreme value theory, provides a theoretical formula of the threshold life table, which is defined as follows:

$$F(x) = \begin{cases} 1 - \exp\left[-\dfrac{B}{\ln C}(c^x - 1)\right], & x \le N, \\ 1 - p\left[1 + \gamma\left(\dfrac{x - N}{\theta}\right)\right]^{-\frac{1}{\gamma}}, & x > N, \end{cases}$$

where $p = S(N) = P(X > N)$ and N is known as the threshold age. In other words, we assume that the survival distribution is Gompertzian before the threshold age and the excesses over the threshold age follow a Generalized Pareto distribution, according to the Balkema-De Haan-Pisckands theorem. To ensure that F is a proper distribution function, we require $B > 0$, $C > 0$ and $\theta > 0$. Such a specification ensures that F is continuous at the threshold age. However, it does not guarantee that F is smooth during the transition from graduation to extrapolation. To achieve smoothness, we require a careful choice of the threshold age. We notice that the excess $Z = \{X - N \mid X > N\}$ over the threshold age N follows a Generalized Pareto distribution, with parameter γ and θ. If $\gamma > 0$, then Z follows a Pareto distribution; If $\gamma = 0$, then Z follows an exponential distribution; If finally $\gamma < 0$, then Z follows a Beta distribution, which has a finite right-hand-end support $-\dfrac{\theta}{\gamma}$.

For the case $\gamma > 0$, we consider the case of estimating the single parameter θ of the Pareto distribution.

3. The Estimation of the Parameter for the Excess

We consider the case of estimating the single parameter θ of the Pareto distribution, with $0 < \theta < +\infty$ and

$$L(\theta \mid y) = \theta^n \prod_{i=1}^{n} y_i^{-(\theta+1)} = \theta^n \exp\left\{-(\theta + 1)\sum_{i=1}^{n} \ln y_i\right\} = \theta^n \exp\{-(\theta + 1)T\}, \quad (3.1)$$

where $T = \sum_{i=1}^{n} \ln y_i$ is a complete sufficient statistic for θ. A mathematically convenient prior density for the problem under consideration is conjugate prior, (Sampford [6])

$$\prod(\theta) \propto e^{-\beta\theta}\theta^{\alpha-1}, \quad \alpha, \beta \geq 0, \quad \theta \geq 0, \tag{3.2}$$

which is simply a member of the gamma family of distributions. The advantage of taking the prior distribution to be conjugate, lies in the fact that the likelihood function $L(\theta | y)$, the prior density $\prod(\theta)$ and the posterior density $\prod(\theta | y)$ are all of the same functional form, thus ensuring mathematical tractability.

If $\alpha = \beta = 0$ in (2.2), then we have the subclass of prior density, given by

$$\prod(\theta) \propto \frac{1}{\theta}, \tag{3.3}$$

which is a uniform density function. Substituting from (3.3) and (3.1), the optimum estimator $\hat{\theta}$ of θ is a solution of

$$\int \frac{\delta L}{\delta\hat{\theta}} \theta^{n-1} \exp\{-(\theta + 1)T\} d\theta = 0, \tag{3.4}$$

where $T = \sum_{i=1}^{n} \ln y_i$, is defined before.

If the loss function is Modified Linear Exponential (MLINEX), see (Cruz [2]), i.e.,

$$L(\hat{\theta}, \theta) = t\left[\left(\frac{\hat{\theta}}{\theta}\right)^{\gamma} - \gamma \ln\left(\frac{\hat{\theta}}{\theta}\right) - 1\right]; \quad \gamma \neq 0, \quad t > 0, \tag{3.5}$$

for the loss function given in (3.5), then it follows from (3.4) that:

$$\int_0^{\infty} (\theta^{-\gamma}\hat{\theta}^{\gamma-1} - \hat{\theta}^{-1})\theta^{n-1} \exp(-\theta T) d\theta = 0$$

$$\Rightarrow \hat{\theta}^{\gamma} = \frac{\int_0^{\infty} \exp(-\theta T)\theta^{n-1} d\theta}{\int_0^{\infty} \exp(-\theta T)\theta^{(n-\gamma)-1} d\theta},$$

from which follows that:

$$\hat{\theta} = \left[\frac{\Gamma(n)}{\Gamma(n-\gamma)}\right]^{\frac{1}{\gamma}} \frac{1}{T}$$

and same as

$$\hat{\theta}_B = [E_\theta(\theta^{-\gamma})]^{\frac{1}{\gamma}},$$

provided that this expected value exists.

Hence, the optimum estimator $\hat{\theta}$ for the loss function (3.5), is given by

$$\hat{\theta}_B = \frac{\Lambda}{T},$$

where $\Lambda = \left[\frac{\Gamma(n)}{\Gamma(n-\gamma)}\right]^{\frac{1}{\gamma}}$ and $T = \sum_{i=1}^{n} \ln y_i$ is a complete sufficient statistic.

Since y is a Pareto variable with parameter θ, $T = \sum_{i=1}^{n} \ln y_i$ is distributed as gamma distribution with parameters θ and n, i.e., $T \sim G(\theta, n)$. The probability density function of T, is

$$p(T \mid \theta) = \frac{\theta^n}{\Gamma(n)} \exp(-\theta T) T^{n-1}; \quad T \geq 0, \quad \theta > 0. \tag{3.6}$$

Again, if the loss function is squared-error of the form

$$L(\hat{\theta}, \theta) = C(\hat{\theta} - \theta)^2, \tag{3.7}$$

where C is a positive constant.

For the loss function given by (3.7), it follows from (3.4) that the optimum estimator $\hat{\theta}$ is given by

$$\hat{\theta}_S = \frac{\int_0^\infty \theta^{(n+1)-1} \exp(-\theta T)\, d\theta}{\int_0^\infty \theta^{n-1} \exp(-\theta T)\, d\theta},$$

which follows that $\hat{\theta}_S = \frac{n}{T}$, where $T = \sum_{i=1}^{n} \ln y_i$ was defined before.

Now, we are interested in finding risk functions for the estimators $\hat{\theta}_B$ and $\hat{\theta}_S$, with respect to MLINEX and SE loss functions.

Hence, the risk function of the estimator $\hat{\theta}_B$, with respect to MLINEX, is given by

$$R_{ML}(\hat{\theta}_B, \theta) = E_B[L(\hat{\theta}_B, \theta)] = t\left[\frac{1}{\theta^\gamma} E(\hat{\theta}_B^\gamma) - \gamma E(\ln \hat{\theta}_B) + \gamma \ln \theta - 1\right]. \tag{3.8}$$

Here, $E(\hat{\theta}_B^\gamma) = \theta^\gamma$ and $E(\ln \hat{\theta}_B) = E\left[\ln\left(\frac{\Lambda}{T}\right)\right] = \ln \Lambda - E(\ln T)$.

For simplicity,

$$E(\ln T) = \frac{\theta^n}{\Gamma(n)} \int_0^\infty (\ln T) \exp(-\theta T) T^{n-1} dT.$$

Using a transformation $Z = \theta T$, then

$$E(\ln T) = -\ln \theta + \frac{1}{\Gamma(n)} \int_0^\infty (\ln y) \exp(-y) y^{n-1} dy = -\ln \theta + \frac{\Gamma'(n)}{\Gamma(n)},$$

where

$$\Gamma'(n) = \int_0^\infty (\ln y) \exp(-y) y^{n-1} dy$$

is the first differentiation of the $\Gamma(n)$ function, with respect to *n*.

Thus,

$$E(\ln \hat{\theta}_B) = \ln \Lambda + \ln \theta - \frac{\Gamma'(n)}{\Gamma(n)}.$$

By using the above results, (3.8) is given by

$$R_{ML}(\hat{\theta}_B, \theta) = t\left[\ln \frac{\Gamma(n-\gamma)}{\Gamma(n)} + \gamma \frac{\Gamma'(n)}{\Gamma(n)}\right]. \tag{3.9}$$

Similarly, the risk function of the estimator $\hat{\theta}_S$, with respect to MLINEX loss function, is

$$R_{ML}(\hat{\theta}_S, \theta) = t\left[n^\gamma \frac{\Gamma(n-\gamma)}{\Gamma(n)} - \gamma \ln(n) + \gamma \frac{\Gamma'(n)}{\Gamma(n)} - 1\right]. \tag{3.10}$$

The risk functions of the estimators $\hat{\theta}_B$ and $\hat{\theta}_S$, with respect to SE loss function, are, respectively,

$$R_S(\hat{\theta}_B, \theta) = E\left(\frac{\Lambda}{T} - \theta\right)^2 = V\left(\frac{\Lambda}{T}\right) + \left[E\left(\frac{\Lambda}{T}\right) - \theta\right]^2$$

$$= \Lambda^2\left[V\left(\frac{1}{T}\right) + \left\{E\left(\frac{1}{T}\right) - \frac{\theta}{\Lambda}\right\}^2\right]$$

$$= \Lambda^2\left[\frac{1}{(n-1)^2(n-2)} + \left\{\frac{1}{n-1} - \frac{1}{\Lambda}\right\}^2\right]\theta^2, \tag{3.11}$$

$$R_S(\hat{\theta}_S, \theta) = E(\hat{\theta}_S - \theta)^2 = \frac{n+2}{(n-1)(n-2)}\theta^2. \tag{3.12}$$

MLINEX and SE risk functions, can be calculated for different values of these parameters. It is evident in every case considered, except of $\gamma = -1$, that the MLINEX risk function $R_{ML}(\hat{\theta}_B, \theta)$ is uniformly smaller than $R_{ML}(\hat{\theta}_S, \theta)$. This implies that in the case of the MLINEX loss function, the MLINEX estimator $\hat{\theta}_B$ is better compared to the SE estimator $\hat{\theta}_S$. If $\gamma = -1$, then the two estimators and hence their risk functions, are identical.

If now $\gamma < -1$, then the risk of $\hat{\theta}_B$, with respect to SE loss function is always greater than that of $\hat{\theta}_S$. Therefore, in this case, $\hat{\theta}_S$ is better compared to the estimator $\hat{\theta}_B$, when SE loss function is considered. If $\gamma = -1$, then the two risks are equal and either estimator is acceptable, whereas if $\gamma > -1$, then $R_S(\hat{\theta}_B) < R_S(\hat{\theta}_S)$, implying that $\hat{\theta}_B$ estimator is better and acceptable, with respect to SE loss function.

In Healthcare Organizations (HCOs) adverse events may provoke dangerous consequences on patients, such as death, a longer hospital stay, and morbidity. As a consequence HCO's department needs to manage legal issues and economic reimbursements. Governances and physicians are interested in Operational Risk (OR), (Goulionis [3-5]) and Clinical Risk (CR) assessment, mainly for forecasting and managing losses and for a correct decision making.

4. Conclusion

Old-age mortality rates are important in many actuarial applications. This paper considers a method which is based on the asymptotic distribution of the excesses over a threshold age. We estimate the parameter of the excess over the threshold age *N*. Also, we prove that sampling distributions that are subjected to weight functions other than *length-biased*, can also be seen in real life applications.

References

[1] A. A. Balkema and L. De Haan, Residual life times at great age, Ann. Probab. 2 (1974), 792-804.

[2] M. G. Cruz, Modelling, Measuring and Hedging Operational Risk, Wiley, New York, 2002.

[3] J. E. Goulionis, Medical treatments using partially observable Markov decision processes, JP J. Biostatist. 3(2) (2009), 77-97.

[4] J. E. Goulionis, A generalized model of optimal maintenance strategy, Pak. J. Statist. 26(2) (2010), 357-364.

[5] J. E. Goulionis and A. Vozikis, Medical decision making for patients with Parkinson disease under average cost criterion, Aust. New Zealand Health Policy 6 (2009c), 15.

[6] M. R. Sampford, On sampling without replacement with unequal probabilities of selection, Biometrika 54 (1967), 499-513.