Generalized confidence intervals for the survival function of the exponential distribution

Shafiei, S.
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran,
email: sobhan.shafiei@yahoo.com

Razmkhah, M.
Department of Statistics, Ferdowsi University of Mashhad, Mashhad, Iran

Abstract

The interval estimation of the survival function of the two-parameter exponential distribution on the basis of the progressively Type-II censored samples is investigated. Toward this end, the concept of the generalized confidence intervals (GCIs) is used and the lower and upper generalized confidence limits (GCLs) are obtained. It will be shown that the coverage probabilities of the GCLs are satisfactory using a simulation study. Finally, some concluding remarks are presented.

Keywords and Phrases: Generalized pivotal quantity, Order statistics, Progressively Type-II censored data

AMS Subject Classification 2000: primary 62N05; secondary 62F25.

1 Introduction

It is well-known that the exponential distribution is one of the commonly used models in several areas of statistical practice, including survival and reliability analysis. It is used to model data with a constant failure rate; for more details concerning the exponential model and related topics, one may refer to the book by Balakrishnan and Basu [2]. A random variable X is said to have a two-parameter exponential distribution if its cumulative distribution function (cdf) is

$$F(x;\mu,\sigma) = 1 - e^{-(x-\mu)/\sigma}, \quad x \geq \mu, \quad \sigma > 0, \tag{1}$$

where μ and σ are the location and scale parameters, respectively. A problem of interest in the reliability analysis is to investigate the confidence intervals (CIs) for the survival function at a specified point τ, which for the two-parameter exponential distribution is defined by

$$R(\tau;\mu,\sigma) = e^{-(\tau-\mu)/\sigma}, \quad \tau \geq \mu. \tag{2}$$

Engelhardt and Bain [5] suggested an approximate method based on Type-II censored data. See also, Roy and Mathew [7] and Fernández [6]. In this paper, we study the problem of constructing GCIs for $R(\tau;\mu,\sigma)$ on the basis of the progressively Type-II censored order statistics. The model of progressive Type-II censoring is of importance in the field of reliability and life testing. Suppose n units are simultaneously placed on a lifetime test. At the time of the ith failure, R_i surviving units are randomly censored from the experiment, $1 \leq i \leq m$. Thus, if m failures are observed, then $R_1 + \cdots + R_m$ units are progressively censored; here, $\mathbf{R} = (R_1, \ldots, R_m)$ denotes the progressive censoring scheme. The interested readers may refer to the book by Balakrishnan and Aggarwala [1]. See also, Balakrishnan *et al.* [3] and Burkschat *et al.* [4].

The rest of the paper is as follows: In Section 2, some preliminaries are presented. In Section 3, the GCIs for the survival function of the two-parameter exponential distribution are derived on the basis of the progressively Type-II censored order statistics. In section 4, some concluding remarks are stated.

2 Preliminaries

Let $\{Y_1, \ldots, Y_n\}$ be a random sample of size n from the two-parameter exponential distribution with cdf in (1). Denote the first m progressively Type-II censored order statistics by $Y^{\mathbf{R}}_{1:m:n} \leq \cdots \leq Y^{\mathbf{R}}_{m:m:n}$ (1 $\leq$

$m \leq n$), where $\mathbf{R} = (R_1, \ldots, R_m)$ stands for the corresponding progressive censoring scheme. The likelihood function of the parameters of the two-parameter exponential distribution with cdf in (1) based on the progressively Type-II censored order statistics can be written as

$$L(\mu, \sigma) = c\sigma^{-m} \exp\left\{ -\sum_{i=1}^{m} (R_i + 1) \frac{y_i - \mu}{\sigma} \right\},$$

where $c = n(n-R_1-1)\cdots(n-R_1-R_2-...-R_{m-1}-m+1)$ and y_i is the observed value of $Y^{\mathbf{R}}_{i:m:n}$. Assuming $m \geq 2$, the maximum likelihood estimators (MLEs) of μ and σ based on $\mathbf{Y} = \{Y^{\mathbf{R}}_{1:m:n}, ..., Y^{\mathbf{R}}_{m:m:n}\}$ are given by

$$\hat{\mu} = \hat{\mu}\,(\mathbf{Y}) = Y^{\mathbf{R}}_{1:m:n} \quad \text{and} \quad \hat{\sigma} = \hat{\sigma}\,(\mathbf{Y}) = \frac{1}{m} \sum_{i=2}^{m} (R_i + 1)\,(Y^{\mathbf{R}}_{i:m:n} - Y^{\mathbf{R}}_{1:m:n}). \tag{3}$$

Let us take

$$Z_1 = (\hat{\mu} - \mu)/\sigma \quad \text{and} \quad Z_2 = \hat{\sigma}/\sigma, \tag{4}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are as defined in (3). It can be shown that $2nZ_1 \sim \chi^2_2$ and independently $2mZ_2 \sim \chi^2_{2(m-1)}$, where χ^2_m stands for a chi-square distribution with m degrees of freedom (see, Balakrishnan *et al.*, [3]).

3 Generalized confidence interval

In this section, we use the concept of the GCI to arrive the exact CIs for $R(\tau; \mu, \sigma)$. Let $\mathbf{X}$ be a random vector whose distribution depends on γ and ξ, a scalar parameter of interest and a nuisance parameter, respectively. Furthermore, let $\mathbf{x}$ denote the observed value of $\mathbf{X}$. The random variable $U(\mathbf{X}; \mathbf{x}, \gamma, \xi)$ is called a generalized pivotal quantity if it satisfies in the following two conditions:

$$\begin{array}{l} (i) \text{ The distribution of } U(\mathbf{X}; \mathbf{x}, \gamma, \xi) \text{ is free of unknown parameters, for fixed } \mathbf{x}, \\ (ii) \text{ The observed value of } U(\mathbf{X}; \mathbf{x}, \gamma, \xi), \text{ i.e., } U(\mathbf{x}; \mathbf{x}, \gamma, \xi), \text{ is equal to } \gamma. \end{array} \tag{5}$$

The CIs for γ obtained using the percentiles of $U(\mathbf{X}; \mathbf{x}, \gamma, \xi)$ are referred to as the GCIs. Therefore, the $U_\alpha(\mathbf{x})$ is a $100(1-\alpha)\%$ lower GCL for γ if

$$\mathrm{P}\left(U(\mathbf{X}; \mathbf{x}, \gamma, \xi) \geq U_\alpha(\mathbf{x})\right) = 1 - \alpha. \tag{6}$$

The quantiles $U_\alpha(\mathbf{x})$ and $U_{1-\alpha}(\mathbf{x})$ are the lower and upper $100(1-\alpha)\%$ GCLs for γ, respectively, whereas $[U_{\alpha/2}(\mathbf{x}), U_{1-\alpha/2}(\mathbf{x})]$ is the two-sided equi-tailed $100(1-\alpha)\%$ GCI for γ based on $U(\mathbf{X}; \mathbf{x}, \gamma, \xi)$. Notice that the coverage probability of such a confidence interval could depend on unknown parameters and hence it may not be exactly $1-\alpha$ (see, for details, [8] and [9]).

To construct a GCI for $R(\tau; \mu, \sigma)$, we first look for generalized pivotal quantities for μ and σ, denoted by U_μ and U_σ, respectively, satisfying the properties in (5). That is, the distribution of (U_μ, U_σ) is free of any unknown parameters, and the observed value of (U_μ, U_σ) is (μ, σ). Toward this end, let $\hat{\mu}_0$ and $\hat{\sigma}_0$ denote the observed values of $\hat{\mu}$ and $\hat{\sigma}$, respectively, where $\hat{\mu}$ and $\hat{\sigma}$ are as defined in (3). Consider a choice of U_μ and U_σ as follows

$$U_\mu = \hat{\mu}_0 - \frac{Z_1}{Z_2}\hat{\sigma}_0 \quad \text{and} \quad U_\sigma = \frac{\hat{\sigma}_0}{Z_2}. \tag{7}$$

Note that a generalized pivotal quantity for any function of μ and σ, say $h(\mu, \sigma)$, is given by $h(U_\mu, U_\sigma)$. Here, the function $h(\mu, \sigma)$ can be quite arbitrary and could be rather complicated. Therefore, using the pivots in (7), a generalized pivotal quantity for $R(\tau; \mu, \sigma)$ in (2) is given by

$$U_R = U_R(\mathbf{Y}; \mathbf{y}, \mu, \sigma) = \exp\{-\frac{\tau - U_\mu}{U_\sigma}\} = \exp\{-Z_1 - \frac{\tau - \hat{\mu}_0}{\hat{\sigma}_0} Z_2\}. \tag{8}$$

Since a confidence limit (CL) for $R(\tau; \mu, \sigma)$ must be restricted to be 1, it is reasonable to use an alternative generalized pivot for $R(\tau; \mu, \sigma)$ as follows

$$U^* = U^*(\mathbf{Y}; \mathbf{y}, \mu, \sigma) = \min\{1, U_R(\mathbf{Y}; \mathbf{y}, \mu, \sigma)\}, \tag{9}$$

where $U_R(\mathbf{Y}; \mathbf{y}, \mu, \sigma)$ is as defined in (8). Clearly, the distribution of U^* is independent of (μ, σ) and $U^*(\mathbf{y}; \mathbf{y}, \mu, \sigma) = R(\tau; \mu, \sigma)$. The exact cdf of U^* is derived in the following subsection.

3.1 Distribution of the generalized pivot

To find the cdf of U^*, we consider two different cases if $\tau \leq \hat{\mu}_0$ or $\tau > \hat{\mu}_0$. Notice that in the case of $\tau \leq \hat{\mu}_0$, the U^* defined in (9) is a mixed random variable with probability function

$$f_{U^*}(x) = \begin{cases} f_{U_R}(x), & 0 < x < 1, \\ \pi, & x = 1, \end{cases}$$

such that

$$\pi = \mathrm{P}(U_R \geq 1) = \mathrm{P}\left(Z_1 + \frac{\tau - \hat{\mu}_0}{\hat{\sigma}_0} Z_2 \leq 0\right), \tag{10}$$

where Z_1 and Z_2 are as defined in (4). So, by some algebraic calculations, we get $\pi = 1 - \left(1 - \frac{n(\tau-\hat{\mu}_0)}{m\hat{\sigma}_0}\right)^{1-m}$. Therefore, the cdf of U^* is given by

$$F_{U^*}(x) = \begin{cases} 0, & x \leq 0, \\ x^n \left(1 - \frac{n(\tau-\hat{\mu}_0)}{m\hat{\sigma}_0}\right)^{1-m} \leq 1 - \pi, & 0 < x < 1, \\ 1, & x \geq 1. \end{cases} \tag{11}$$

It is obvious that in the case of $\tau > \hat{\mu}_0$, the U^* defined in (9) is a continuous random variable; that is, $U^* = U_R$. Hence, by performing some algebraic calculations, the cdf of U^* in this case is given by

$$F_{U^*}(x) = \begin{cases} 0, & x \leq 0, \\ \phi(x; \tau, \hat{\mu}_0, \hat{\sigma}_0), & 0 < x < 1, \\ 1, & x \geq 1, \end{cases} \tag{12}$$

such that

$$\begin{aligned} \phi(x;\tau,\hat{\mu}_0,\hat{\sigma}_0) &= \frac{\Gamma\left(m-1, -\frac{m\hat{\sigma}_0 \log x}{\tau - \hat{\mu}_0}\right)}{\Gamma(m-1)} \\ &\quad + \frac{x^n}{\Gamma(m-1)} \psi\left(-\frac{m\hat{\sigma}_0 \log x}{\tau - \hat{\mu}_0}, \left(1 - \frac{n}{m}\frac{\tau - \hat{\mu}_0}{\hat{\sigma}_0}\right), m-1\right), \end{aligned}$$

where $\Gamma(\alpha)$ stands for the complete gamma function, $\Gamma(\alpha, t)$ represents the incomplete gamma function (*i.e.*, $\Gamma(\alpha,t) = \int_t^\infty e^{-y} y^{\alpha-1} dy$) and

$$\psi(t,\beta,\alpha) = \begin{cases} \frac{t^\alpha}{\alpha}, & if\ \beta = 0,\ t > 0,\ \alpha > 0, \\ \frac{\Gamma(\alpha) - \Gamma(\alpha, \beta t)}{\beta^\alpha}, & if\ \beta \neq 0,\ t > 0,\ \alpha > 0. \end{cases}$$

3.2 Interval estimation for survival function

As previously mentioned, the percentiles of U^* construct the GCIs for the survival function at τ. Using (6), for given α, a $100(1-\alpha)\%$ lower GCL for $R(\tau;\mu,\sigma)$ is defined by $\inf\{x : F_{U^*}(x) \geq \alpha\}$. To derive the exact lower GCLs for $R(\tau;\mu,\sigma)$, we consider two cases whether $\tau \leq \hat{\mu}_0$ or $\tau > \hat{\mu}_0$.

Case I) Suppose that $\tau \leq \hat{\mu}_0$, then using (11), the α-quantile of U^* is defined by $F_{U^*}^{-1}(\alpha)$ if $\alpha < 1 - \pi$ and coincides with 1 otherwise. Hence, a $100(1-\alpha)\%$ lower GCL for $R(\tau;\mu,\sigma)$ is given by

$$R_1(\tau;\alpha) = \min\left\{1, \alpha^{1/n}\left(1 - \frac{n(\tau - \hat{\mu}_0)}{m\hat{\sigma}_0}\right)^{\frac{m-1}{n}}\right\}. \tag{13}$$

Case II) Now, suppose that $\tau > \widehat{\mu}_0$, then a $100(1-\alpha)\%$ lower GCL for $R(\tau;\mu,\sigma)$ is given by

$$R_2(\tau;\alpha) = F_{U^*}^{-1}(\alpha), \tag{14}$$

where $F_{U^*}(x)$ is as defined in (12). The lower GCLs for $R(\tau;\mu,\sigma)$ can be obtained by FindRoot command in Mathematica, using (14).

To illustrate the performance of the proposed procedure in this paper, we simulate the values of 95% lower GCLs for survival function at τ for $n = 10$, $m = 5$ and some selected choices of the progressive censoring schemes $\mathbf{R} = (R_1, \ldots, R_m)$. Furthermore, the values of μ and σ have been chosen to be $\mu = 0.5, 1$ and $\sigma = 0.4, 1.1$. The results are presented in Table 1. Similar results are tabulated in Table 2 for the associated coverage probabilities. The lower GCLs and the coverage probabilities are obtained using 5000 times simulations.

Table 1. Values of 95% lower GCLs for $R(\tau; \mu, \sigma)$ for $n = 10$ and $m = 5$.

		$\mu = 0.5$				$\mu = 1$		
		τ				τ		
σ	**R**	1	1.5	2	5	1.5	2	5
0.4	(1,1,1,1,1)	0.1017	0.0181	0.0042	0.0000	0.1030	0.0186	0.0000
	(0,0,0,0,5)	0.1015	0.0183	0.0044	0.0000	0.1030	0.0184	0.0000
	(0,5,0,0,0)	0.1050	0.0194	0.0047	0.0000	0.1028	0.0187	0.0000
	(5,0,0,0,0)	0.1053	0.0193	0.0047	0.0000	0.1031	0.0188	0.0000
	(0,0,3,2,0)	0.1037	0.0191	0.0046	0.0000	0.1041	0.0191	0.0000
1.1	(1,1,1,1,1)	0.3940	0.1760	0.0859	0.0031	0.3992	0.1797	0.0052
	(0,0,0,0,5)	0.3920	0.1747	0.0851	0.0031	0.3986	0.1790	0.0052
	(0,5,0,0,0)	0.3978	0.1789	0.0874	0.0032	0.3954	0.1777	0.0050
	(5,0,0,0,0)	0.3940	0.1752	0.0855	0.0031	0.3935	0.1753	0.0048
	(0,0,3,2,0)	0.3972	0.1779	0.0867	0.0031	0.3985	0.1783	0.0050

Table 2. Coverage probabilities of the lower GCLs for $R(\tau; \mu, \sigma)$ for $n = 10$ and $m = 5$.

		$\mu = 0.5$				$\mu = 1$		
		τ				τ		
σ	**R**	1	1.5	2	5	1.5	2	5
0.4	(1,1,1,1,1)	95.30	95.48	95.44	95.48	95.04	94.90	94.90
	(0,0,0,0,5)	95.00	95.06	95.06	94.90	94.90	95.12	95.04
	(0,5,0,0,0)	94.54	94.50	94.58	94.62	95.22	95.36	95.18
	(5,0,0,0,0)	94.72	94.58	94.70	94.76	95.36	95.28	95.30
	(0,0,3,2,0)	94.56	94.72	94.82	94.78	95.00	95.02	95.12
1.1	(1,1,1,1,1)	95.30	95.39	95.06	94.80	94.96	95.02	94.84
	(0,0,0,0,5)	95.62	95.54	95.22	94.96	95.16	94.72	94.54
	(0,5,0,0,0)	94.58	94.82	94.78	94.88	95.22	95.32	95.48
	(5,0,0,0,0)	95.16	95.20	94.98	95.04	95.08	95.16	95.38
	(0,0,3,2,0)	94.98	94.98	95.00	95.10	94.64	94.60	95.12

From Table 2, it is observed that the coverage probabilities of the GCLs are satisfactory.

4 Concluding remarks

The interval estimation of the survival function of the two-parameter exponential distribution on the basis of the progressively Type-II censored samples was studied in this paper. Toward this end, we obtained the GCI on the basis of a generalized pivotal quantity for the survival function. One may also derive a GCI which the associated expected width is minimum. The interval (L, U) is called a $100(1-\alpha)\%$ CI with the shortest expected width for unknown parameter θ on the basis of the pivotal quantity Q, if $F_Q(U) - F_Q(L) = 1 - \alpha$ and $\mathrm{E}(U - L)$ is minimum. Since the probability function of U^* defined in (9), for the case of $\tau \leq \hat{\mu}_0$, is increasing, the $100(1-\alpha)\%$ GCI with the shortest expected width for $R(\tau; \mu, \sigma)$ is $[R_1(\tau; \alpha), 1]$, where, the $R_1(\tau; \alpha)$ is as defined in (13). Now, suppose that $\tau > \widehat{\mu}_0$, then using (12), it can be shown that $f_{U^*}(0) = f_{U^*}(1) = 0$ and that the probability density function of U^* defined in (9) has a unique mode in $(0, 1)$. Therefore, by some algebraic calculations it is deduced that the interval $(\zeta_\alpha, \xi_\alpha)$ is a $100(1-\alpha)\%$ GCI with the shortest expected width for $R(\tau; \mu, \sigma)$ on the basis of the generalized pivot U^*, if

$$f_{U^*}(\zeta_\alpha) = f_{U^*}(\xi_\alpha) \quad \text{and} \quad F_{U^*}(\xi_\alpha) - F_{U^*}(\zeta_\alpha) = 1 - \alpha.$$

The exact values of ζ_α and ξ_α can be easily obtained using the FindRoot command in Mathematica.

References

[1] Balakrishnan, N. and Aggarwala, R. (2000). Progressive Censoring: Theory, Methods and Applications. Birkhauser, Boston.

[2] Balakrishnan, N. and Basu, A.P. (1995). Exponential Distribution: Theory, Methods and Applications. Taylor & Francis, New York.

[3] Balakrishnan, N., Cramer, E., Kamps, U. and Schenk, N. (2001). Progressive Type-II censored order statistics from exponential distributions. *Statistics*, **35**, 537–556.

[4] Burkschat, M., Cramer, E. and Kamps, U. (2006). On optimal schemes in progressive censoring. *Statistics and Probability Letters*, **76**, 1032–1036.

[5] Engelhardt, M. and Bain, L.J. (1978). Tolerance limits and confidence limits on reliability for the two-parameter exponential distribution. *Technometrics*, **20**, 37–39.

[6] Fernández, A. (2007). On calculating generalized confidence intervals for the two-parameter exponential reliability function. *Statistics*, **41**, 129–135.

[7] Roy, A. and Mathew, T. (2005). A generalized confidence limit for the reliability function of a two-parameter exponential distribution. *Journal of Statistical Planning and Inference*, **128**, 509–517.

[8] Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association*, **88**, 899–905.

[9] Weerahandi, S. (1995). Exact Statistical Methods for Data Analysis. Springer, New York.