PA91050280( 7 .P)

行政院國家科學委員會專題研究計畫成果報告

# Exact inference based on exponential distribution for censored samples

## 指數分配雙邊型 II 有限樣本的完全線性推論與預測區間

計畫編號:NSC 90 - 2118- M- 032- 005

執行期限:90年8月1日至91年7月31日

主持人: 林千代　　執行機構及單位名稱: 淡江大學數學系

計畫參與人員: 吳正新及凌爾瑩　　執行機構及單位名稱: 淡江大學數學系

中文摘要

我們利用 Huffer and Lin (2001) 的演算法來取得指數分配雙邊型 II 受限資料的完全線性推論與預測區間.

關鍵詞: 最佳線性不偏估計式, 留間隔, 預測區間.

## Abstract

In this paper, we make use of an algorithm of Huffer and Lin (2001) in order to develop the exact interval estimation for the exponential scale parameter based on doubly Type-II and general progressively Type-II censored samples.

Keywords: best linear unbiased estimators, spacings, prediction intervals.

## 1 Introduction

It is well known that the normalized spacings from an exponential distribution are independent and identically distributed as exponential; see, Sukhatme (1937) and Rènyi (1953). Similarly, this property can be extended to the general progressively Type-II censored samples; see Viveros and Balakrishnan (1994). The property of normalized spacings allows the development of exact chi-square confidence intervals for the scale parameter of an exponential distribution based on Type-II right censored samples. However, when the available sample from the life-testing experiment is doubly Type-II censored and general progressively Type-II censored, such an elegant result does not exist in the literature.

In this work, we develop exact confidence intervals for the exponential scale parameter based on doubly Type-II censored and general progressively Type-II censored samples by utilizing an algorithm of Huffer and Lin (2001). Our results obtained in the case of the Type-II right censored samples (case $r = 0$) and the progressively Type-II right censored samples (case $r = 0$) all agree with those of

Lawless (1971), Balakrishnan and Aggarwala (2000), and Viveros and Balakrishnan (1994). Moreover, they can also serve as a criterion to evaluate the accuracy of a chi-square approximation proposed by Balakrishnan and Gupta (1998) for the distribution of the BLUE of the scale parameter. Finally, we present two examples to illustrate the exact methods of inference developed here.

## 2 Algorithm

Our work relies heavily on the algorithm proposed by Huffer and Lin (2001). For completeness, we now introduce the basic idea of this algorithm. Suppose that $Z_1, Z_2, \ldots, Z_{n+1}$ are i.i.d. exponentials with mean 1. We define $\boldsymbol{Z}^{(n)} = (Z_1, Z_2, \ldots, Z_{n+1})'$. Huffer and Lin (2001) recently developed an algorithm for evaluating the probabilities involving linear combinations of i.i.d. exponentials with arbitrary rational coefficients of the form

$$P(\boldsymbol{A}\boldsymbol{Z}^{(n)} \leq t\boldsymbol{b}), \tag{1}$$

where $\boldsymbol{A}$ is any matrix of rational values, $\boldsymbol{b}$ is any vector of rational values, and $t > 0$ is a real-valued scalar. This algorithm depends on the repeated, systematic use of the two recursions [given in Eqs. (3) and (4) below] which are stated in terms of a function $Q$ defined by

$$Q(\boldsymbol{A}, \boldsymbol{b}, \lambda, p) = p!\, R(p, \lambda) \times P\left((1 - \lambda t)\boldsymbol{A}\boldsymbol{Z}^{(n-p)} > t\boldsymbol{b}\right), \tag{2}$$

where $R(p, \lambda) = \frac{t^p}{p!} e^{-\lambda t}$, for integers $p \geq 0$ and real values $\lambda \geq 0$. In particular, $Q(\emptyset, \emptyset, \lambda, p) = p!\, R(p, \lambda)$.

**Result 1** *Let $\boldsymbol{A}$ be an arbitrary matrix. Let $r$ and $q$ be the number of rows and columns of $\boldsymbol{A}$. For any $r \times 1$ vector $\boldsymbol{x}$, we define $A_{i,x}$ to be the matrix obtained by replacing the $i^{th}$ column of $\boldsymbol{A}$ by $\boldsymbol{x}$. Let $\boldsymbol{c} = (c_1, \ldots, c_q)'$ be any $q \times 1$ vector satisfying $\sum_{i=1}^{q} c_i = 1$. Define $\boldsymbol{\xi} = \boldsymbol{A}\boldsymbol{c}$. Then*

$$Q(\boldsymbol{A}, \boldsymbol{b}, \lambda, p) = \sum_{i=1}^{q} c_i Q(\boldsymbol{A}_{i,\xi}, \boldsymbol{b}, \lambda, p). \tag{3}$$

This recursion is an immediate consequence of the more general recursion given by Huffer (1988). See Huffer (1988), Lin (1993), and Huffer and Lin (1999, 2001) for applications of this general recursion.

**Result 2** *If a matrix $\boldsymbol{A} = (a_{ij})$ and vector $\boldsymbol{b} = (b_j)$ satisfy (for some $k \geq 1$) the following three conditions*

*(R1) $a_{1j} = 0$ for $j > k$,*

*(R2) $a_{ij} = a_{i1}$ for $j \leq k$, (i.e., the first $k$ columns of $\boldsymbol{A}$ are identical),*

*(R3) $a_{11} > 0$ and $b_1 > 0$,*

*then*

$$Q(\boldsymbol{A}, \boldsymbol{b}, \lambda, p) = \sum_{i=0}^{k-1} \frac{\delta^i}{i!} \times Q(\boldsymbol{A}^*_{(-i)}, \boldsymbol{b}^* - \delta \boldsymbol{a}^*, \lambda + \delta, p + i), \tag{4}$$

*where $\delta = b_1/a_{11}$, $\boldsymbol{A}^*$ is a matrix obtained by deleting the first row of $\boldsymbol{A}$, $\boldsymbol{A}^*_{(-i)}$ is a matrix obtained by deleting the first $i$ columns of $\boldsymbol{A}^*$, $\boldsymbol{b}^*$ is a vector obtained by deleting the first entry of $\boldsymbol{b}$, and $\boldsymbol{a}^*$ is a vector obtained by taking the first column of $\boldsymbol{A}$ and deleting the first entry.*

Each recursion is used to re-express a probability as in (1) by decomposing it into a sum of similar, but simpler components. The same recursions are then applied to each of these components and so on. The process is continued until we obtain components which are simple and easily expressed in closed form.

# 3 Exact Interval Estimation Based On Doubly Type-II Censored Samples

Let $X_{(r+1)} \le X_{(r+2)} \le \cdots \le X_{(n-s)}, 1 \le r+1 \le n-s \le n$, be a doubly Type-II censored sample from the one-parameter exponential distribution with probability density function (pdf)

$$f(x;\sigma) = \frac{1}{\sigma}e^{-x/\sigma}, x \ge 0, \sigma > 0. \qquad (5)$$

Denote $S_i = (n-i+1)(X_{(i)} - X_{(i-1)})$. Then the best linear unbiased estimator of $\sigma$ based on this sample is given by [see Balakrishnan and Cohen (1991)] $\sigma^* = \frac{1}{K}\sum_{i=r+1}^{n-s} a_i X_{(i)} = \sum_{i=1}^{n-s} c_i S_i$, where $K = (n-r-s-1) + \{\sum_{\ell=n-r}^{n} \frac{1}{\ell}\}^2 / \{\sum_{\ell=n-r}^{n} \frac{1}{\ell^2}\}$,

$$a_i = \begin{cases} \frac{\sum_{\ell=n-r}^{n} 1/\ell}{\sum_{\ell=n-r}^{n} 1/\ell^2} - (n-r-1) \\ \text{for } i = r+1, \\ 1 \\ \text{for } r+2 \le i \le n-s-1, \\ s+1 \\ \text{for } i = n-s, \end{cases}$$

and

$$c_i = \begin{cases} \frac{1}{K(n-i+1)} \sum_{j=r+1}^{n-s} a_j \\ \text{for } i = 1, 2, \ldots, r+1, \\ \frac{1}{K(n-i+1)} \sum_{j=i}^{n-s} a_j \\ \text{for } i = r+2, \ldots, n-s. \end{cases}$$

From the property of normalized spacings and the algorithm of Huffer and Lin (2001), we can find the exact values of $t$ satisfying $P(\sigma^*/\sigma \ge t) = P(\sum_{i=1}^{n-s} c_i Z_i > t) = \alpha$ for specified values of $\alpha$.

**Example** : This example comes from Lawless (1982, pp. 138) with $n = 12, r = 2, s = 1$, and the doubly Type-II censored sample is as follows:

$$--, --, 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, 98.1, 138.6, --$$

In this case, we have $\sigma^* = 71.1385$ and $\boldsymbol{A}$ in (2) as

$$\boldsymbol{A} = \Big(\frac{9955}{120609}\ \frac{3620}{40203}\ \frac{3982}{40203}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\Big).$$

The approximate 95% confidence interval for $\sigma$ obtained by Balakrishnan and Gupta (1998) was $[22\sigma^*/36.78071, 22\sigma^*/10.98232] = [42.5508, 142.5060]$, which is quite close to our exact 95% confidence interval for $\sigma$ given by $[22\sigma^*/36.79546, 22\sigma^*/10.97781] = [42.5337, 142.5646]$.

# References


Balakrishnan, N. and Aggarwala, R. (2000). *Progressive Censoring: Theory, Methods, and Applications.* Boston: Birkhäuser.

Balakrishnan, N. and Cohen, A. C. (1991). *Order Statistics and Inference: Estimation Methods.* San Diego: Academic Press.

Balakrishnan, N. and Gupta, S. S. (1998). Higher-order moments of order statistics from exponential and right-truncated exponential distributions and applications to life-testing problems. In N. Balakrishnan and C.R. Rao (Eds.), *Handbook of Statistics-17: Order Statistics - Applications*, pp. 25–59. Amsterdam: North-Holland.

Balakrishnan, N. and Sandhu, R. A. (1996). Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type-II censored samples. *Sankhyā, Series B 58*, 1–9.

Huffer, F. W. (1988). Divided differences and the joint distribution of linear combi-

nations of spacings. *Journal of Applied Probability 25*, 346–354.
Huffer, F. W. and Lin, C. T. (1999). An approach to computations involving spacings with applications to the scan statistic. In J. Glaz and N. Balakrishnan (Eds.), *Scan Statistics and Applications*, pp. 141–163. Boston: Birkhäuser.
Huffer, F. W. and Lin, C. T. (2001). Computing the joint distribution of general linear combinations of spacings or exponential variates. *Statistica Sinica 11*(4), 1141–1157.
Lawless, J. F. (1971). A prediction problem concerning samples from the exponential distribution, with application in life testing. *Technometrics 13*, 725–730.
Lawless, J. F. (1982). *Statistical Models & Methods for Lifetime Data.* New York: Wiley.
Lin, C. T. (1993). The computation of probabilities which involve spacings, with applications to the scan statistic. Ph.D. Thesis, Florida State University, Tallahassee.
Lingappaiah, G. S. (1973). Prediction in exponential life testing. *Canadian Journal of Statistics 1*, 113–118.
Nelson, W. (1982). *Applied Life Data Analysis.* New York: Wiley.
Rènyi, A. (1953). On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungaricae 4*, 191–231.
Sukhatme, P. V. (1937). Tests of significance for samples of the $\chi^2$ population with two degrees of freedom. *Annals of Eugenics 8*, 52–56.
Viveros, R. and Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics 36*, 84–91.

**Theorem 5.2.** *If $g$ denotes $Q = Q(\Delta; t, z)$, $Q^{(k)} = Q^{(k)}(\Delta; t, z)$, $Q^{(tr)} = Q^{(tr)}(\Delta; t, z)$ or $e(t, z)$ then the distribution of the random variable $\sqrt{n}(\hat{g} - g)$ tends to the normal law with the mean 0 and the variance $V(\hat{g})$. The asymptotic variances are:*

$$V(\hat{Q}^{(k)}) = \int_z^{z_0 \wedge z_1} \left\{ h'(u,a) \int_u^{z_0 \wedge z_1} h'(y,a) \exp\left\{ -\int_z^y h'(v,a) d\Lambda^{(\cdot)}(v) \right\} d\Lambda_n^{(k)}(y) \right\}^2 d\sigma_{-k}^2(u)$$
$$+ \int_z^{z_0 \wedge z_1} (h'(u,a))^2 \left\{ -\int_u^{z_0 \wedge z_1} h'(y,a) \exp\left\{ -\int_z^y h'(v,a) d\Lambda^{(\cdot)}(v) \right\} d\Lambda_n^{(k)}(y) \right.$$
$$\left. + \exp\left\{ -\int_z^u h'(v,a) d\Lambda^{(\cdot)}(v) \right\} \right\}^2 d\sigma_k^2(u)),$$
$$V(\hat{Q}^{(tr)}) = (1 - Q^{(tr)})^2 \int_z^{z_0 \wedge z_1} (h'(u,a))^2 d\sigma^2(u),$$
$$V(\hat{e}(t,z)) = \int_z^{z_0} (h'(y,a))^2 \left\{ \int_y^{z_0} (h(u,a) - h(y,a)) h'(u,a) \exp\left\{ -\int_z^u h'(v,a) d\Lambda^{(\cdot)}(v) \right\} d\Lambda^{(\cdot)}(u) \right.$$
$$\left. + (h(y,a) - h(z_0,a)) \exp\left\{ -\int_z^{z_0} h'(v,a) d\Lambda^{(\cdot)}(v) \right\} \right\}^2 d\sigma^2(y).$$

### 2.4 Estimation of the ideal reliability characteristics

It may seem that the ideal characteristics could be estimated using the model with a smaller number of failure modes. This is incorrect because the ideal characteristics are estimated using data about failures of all modes, not only of modes with indexes from the set $\omega$.

Suppose that $g$ is any unconditional reliability characteristic and $g_{0\omega}$ is the corresponding ideal characteristic. The formulas for the estimator $\hat{g}_{0\omega}$ and its asymptotic variance are obtained from corresponding formulas for the estimator $\hat{g}$ replacing $\Lambda^{(\cdot)}$ by $\Lambda^{(\omega)}$ and $\Lambda^{(-k)}$ by $\Lambda^{(\omega\setminus\{k\})}$, where $\hat{\Lambda}^{(\omega)}(z) = \sum_{k\in\omega} \hat{\Lambda}^{(k)}(z)$, $\hat{\Lambda}^{(\omega\setminus\{k\})}(z) = \sum_{l\in\omega\setminus\{k\}} \hat{\Lambda}^{(l)}(z)$.

## 3 Semi-parametric and parametric estimation

Suppose that the function $\lambda^{(k)}(z)$ is from the class $\lambda^{(k)}(z, \gamma_k)$, where $\gamma_k$ is a (possibly multi-dimensional) parameter. The maximum likelihood estimators $\hat{\gamma}_k$ (in the case of the linear degradation model) verify the following equations:

$$\sum_{i=1}^n \frac{\partial}{\partial \gamma_k} \ln\{\lambda^{(k)}(Z_i; \hat{\gamma}_k)\} 1_{\{V_i = k\}} - \sum_{i=1}^n A_i \frac{\partial}{\partial \gamma_k} \Lambda^{(k)}(Z_i; \hat{\gamma}_k) = 0.$$

The functions $\Lambda^{(k)}(z) = \Lambda^{(k)}(z, \gamma_k)$ are estimated by $\hat{\Lambda}^{(k)}(z) = \Lambda^{(k)}(z, \hat{\gamma}_k)$. Semi-parametric estimators of the main reliability characteristics are obtained replacing all functions $\Lambda^{(k)}(t)$ by their estimators $\Lambda^{(k)}(t; \hat{\gamma}_k)$ and the distribution function $\pi(a)$ by its estimator $\hat{\pi}$ in the expressions of these characteristics.

In the case of parametric estimation the distribution function is taken from a specified family of distributions $\pi(a) = \pi(a, \eta)$, and the estimators $\hat{\eta}$ of the unknown parameters $\eta$ of this distribution are estimated by the method of maximum likelihood using the complete data $A_1, \ldots, A_n$.

## 4 Analysis of real tire wear data

We analysed failure time and wear data of 101 bus tires 01-73 B manufactured at the Omsk tire plant and used in the first quarter of 2000 at the Tashkent bus park N7 on buses DEU-BS-106 made in South Korea.

## References


Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*. New York: Wiley.




# Exact Inference based on exponential distribution for censored and progressively censored samples


**N. Balakrishnan**
Department of Mathematics and Statistics
McMaster University
Hamilton, Ontario
Canada
*bala@mcmail.cis.mcmaster.ca*

**Chien-Tai Lin**
Department of Mathematics
Tamkang University
Tamsui, Taiwn
R.O.C.
*chien@math.tku.edu.tw*



### Abstract

In this paper, we make use of an algorithm of Huffer and Lin (2001) in order to develop the exact interval estimation for the exponential scale parameter based on doubly Type-II and general progressively Type-II censored samples. This approach enables us to determine the exact percentage points of the pivotal quantity based on the best linear unbiased estimator, and construct the exact prediction intervals for the $\ell$-th order statistic in a sample of size $n$ based on doubly Type-II censored samples.


## 1 Introduction

It is well known that the normalized spacings from an exponential distribution are independent and identically distributed as exponential; see, Sukhatme (1937) and Rènyi (1953). Similarly, this property can be extended to the general progressively Type-II censored samples; see Viveros and Balakrishnan (1994). The property of normalized spacings allows the development of exact chi-square confidence intervals for the scale parameter of an exponential distribution based on Type-II right censored samples. It has also been utilized by many authors to construct exact prediction intervals for failure times of items having observed the first $n - s$ failures from a sample of $n$ items placed on a life test; see, for example, Lawless (1971, 1977), Likeš (1974), and Lingappaiah (1973). However, when the available sample from the life-testing experiment is doubly Type-II censored and general progressively Type-II censored, no exact distributional result is available for the appropriate pivotal quantity based on the best linear unbiased estimator (BLUE) of the scale parameter that is useful for the interval estimation in the one-parameter exponential distribution.

In this paper, we develop exact confidence intervals for the exponential scale parameter based on doubly Type-II censored and general progressively Type-II censored samples by utilizing an algorithm of Huffer and Lin (2001). We also construct exact prediction intervals for the $\ell$-th order statistic in a sample of size $n$ based on doubly Type-II censored samples. Our results obtained in the case of the Type-II right censored samples (case $r = 0$) and the progressively Type-II right censored samples (case $r = 0$) all agree with those of Lawless (1971), Likeš (1974), Balakrishnan and Aggarwala (2000), and Viveros and Balakrishnan (1994). Moreover, they can also serve as a criterion to evaluate the accuracy of a chi-square approximation proposed by Balakrishnan and Gupta (1998) for the distribution of the BLUE. Finally, we present two examples to illustrate the exact methods of inference developed here.

## 2 Algorithm

Our work relies heavily on the algorithm proposed by Huffer and Lin (2001). For completeness, we now introduce the basic idea of this algorithm. Suppose that $Z_1, Z_2, \ldots, Z_{n+1}$ are i.i.d. exponentials with mean 1. We define $Z^{(n)} = (Z_1, Z_2, \ldots, Z_{n+1})'$. Huffer and Lin (2001) recently developed an algorithm for evaluating the probabilities involving linear combinations of i.i.d. exponentials with arbitrary rational coefficients of the form

$$P(AZ^{(n)} \leq tb), \tag{1}$$

where $A$ is any matrix of rational values, $b$ is any vector of rational values, and $t > 0$ is a real-valued scalar. This algorithm depends on the repeated, systematic use of the two recursions [given in Eqs. (3) and (4) below] which are stated in terms of a function $Q$ defined by

$$Q(A,b,\lambda,p) = p!\,R(p,\lambda)\,P\left((1-\lambda t)AZ^{(n-p)} > tb\right), \tag{2}$$

where $R(p,\lambda) = \frac{t^p}{p!}e^{-\lambda t}$, for integers $p \geq 0$ and real values $\lambda \geq 0$. In particular, $Q(\emptyset,\emptyset,\lambda,p) = p!\,R(p,\lambda)$.

**Recursion 1** *Let $A$ be an arbitrary matrix. Let $r$ and $q$ be the number of rows and columns of $A$. For any $r \times 1$ vector $x$, we define $A_{i,x}$ to be the matrix obtained by replacing the $i^{th}$ column of $A$ by $x$. Let $c = (c_1,\ldots,c_q)'$ be any $q \times 1$ vector satisfying $\sum_{i=1}^{q} c_i = 1$. Define $\xi = Ac$. Then*

$$Q(A,b,\lambda,p) = \sum_{i=1}^{q} c_i Q(A_{i,\xi},b,\lambda,p). \tag{3}$$

This recursion is an immediate consequence of the more general recursion given by Huffer (1988). See Huffer (1988), Lin (1993), and Huffer and Lin (1999, 2001) for applications of this general recursion.

**Recursion 2** *If a matrix $A = (a_{ij})$ and vector $b = (b_j)$ satisfy (for some $k \geq 1$) the following three conditions*

*(R1) $a_{1j} = 0$ for $j > k$,*
*(R2) $a_{ij} = a_{i1}$ for $j \leq k$, (i.e., the first $k$ columns of $A$ are identical),*
*(R3) $a_{11} > 0$ and $b_1 > 0$,*

*then*

$$Q(A,b,\lambda,p) = \sum_{i=0}^{k-1} \frac{\delta^i}{i!} Q(A^*_{(-i)}, b^* - \delta a^*, \lambda+\delta, p+i), \tag{4}$$

*where $\delta = b_1/a_{11}$, $A^*$ is a matrix obtained by deleting the first row of $A$, $A^*_{(-i)}$ is a matrix obtained by deleting the first $i$ columns of $A^*$, $b^*$ is a vector obtained by deleting the first entry of $b$, and $a^*$ is a vector obtained by taking the first column of $A$ and deleting the first entry.*

Each recursion is used to re-express a probability as in (1) by decomposing it into a sum of similar, but simpler components. The same recursions are then applied to each of these components and so on. The process is continued until we obtain components which are simple and easily expressed in closed form.

## 3 Exact Interval Estimation Based On Doubly Type-II Censored Samples

Let $X_{(r+1)} \leq X_{(r+2)} \leq \cdots \leq X_{(n-s)}, 1 \leq r+1 \leq n-s \leq n$, be a doubly Type-II censored sample from the one-parameter exponential distribution with probability density function (pdf)

$$f(x;\sigma) = \frac{1}{\sigma}e^{-x/\sigma}, x \geq 0, \sigma > 0. \tag{5}$$

Denote $S_i = (n-i+1)(X_{(i)} - X_{(i-1)})$. Then the best linear unbiased estimator of $\sigma$ based on this sample is given by [see Balakrishnan and Cohen (1991)] $\sigma^* = \frac{1}{K}\sum_{i=r+1}^{n-s} a_i X_{(i)} = \sum_{i=1}^{n-s} c_i S_i$, where $K = (n-r-s-1) + \{\sum_{\ell=n-r}^{n} \frac{1}{\ell}\}^2/\{\sum_{\ell=n-r}^{n} \frac{1}{\ell^2}\}$,

$$a_i = \begin{cases} \left[\sum_{\ell=n-r}^{n} 1/\ell\right]/\left[\sum_{\ell=n-r}^{n} 1/\ell^2\right] - (n-r-1) & \text{for } i = r+1, \\ 1 & \text{for } r+2 \leq i \leq n-s-1, \\ s+1 & \text{for } i = n-s, \end{cases}$$



$$c_i = \begin{cases} \frac{1}{K(n-i+1)}\sum_{j=r+1}^{n-s} a_j & \text{for } i = 1,2,\ldots,r+1, \\ \frac{1}{K(n-i+1)}\sum_{j=i}^{n-s} a_j & \text{for } i = r+2,\ldots,n-s. \end{cases}$$

From the property of normalized spacings and the algorithm of Huffer and Lin (2001), we can find the exact values of $t$ satisfying $P(\sigma^*/\sigma \geq t) = P(\sum_{i=1}^{n-s} c_i Z_i > t) = \alpha$ for specified values of $\alpha$. In a similar fashion, we can construct exact prediction intervals for the order statistic $X_{(\ell)}$, $n-s < \ell \leq n$, by finding the exact values of $t$ such that $P([X_{(\ell)} - X_{(n-s)}]/\sigma^* > t) = P(\sum_{i=1}^{\ell} d_i Z_i > 0) = \alpha$ with

$$d_i = \begin{cases} -\frac{t}{K(n-i+1)}\sum_{j=r+1}^{n-s} a_j & \text{for } i = 1,2,\ldots,r+1, \\ -\frac{t}{K(n-i+1)}\sum_{j=i}^{n-s} a_j & \text{for } i = r+2,\ldots,n-s, \\ \frac{1}{n-i+1} & \text{for } i = n-s+1,\ldots,\ell. \end{cases}$$

## 4 Exact Interval Estimation Based On General Progressively Type-II Censored Samples

Suppose $X^{(R_{r+1},\ldots,R_m)}_{r+1:m:n} \leq X^{(R_{r+1},\ldots,R_m)}_{r+2:m:n} \leq \ldots \leq X^{(R_{r+1},\ldots,R_m)}_{m:m:n}$ is a general progressively Type-II censored sample from a distribution with pdf in (5). With $X^{(R_{r+1},\ldots,R_m)}_{0:m:n} = 0$, let $D_i = (n-i+1)(X^{(R_{r+1},\ldots,R_m)}_{i:m:n} - X^{(R_{r+1},\ldots,R_m)}_{i-1:m:n})$, for $i = 1,2,\ldots,r+1$, and $W_j = (n - \sum_{i=r+1}^{j-1} R_i - j + 1)(X^{(R_{r+1},\ldots,R_m)}_{j:m:n} - X^{(R_{r+1},\ldots,R_m)}_{j-1:m:n})$, for $j = r+2,\ldots,m$. Then, the best linear unbiased estimator of the exponential scale parameter $\sigma$ based on this general progressively Type-II right censored sample [see Balakrishnan and Sandhu (1996)] can be expressed as $\sigma^*_g = \frac{1}{m-r-1+\frac{\alpha^2_{r+1}}{\beta_{r+1}}}[\sum_{i=r+2}^{m} W_i + (\frac{\alpha_{r+1}}{\beta_{r+1}})\sum_{i=1}^{r+1} \frac{D_i}{n-i+1}]$, where $\alpha_{r+1} = \sum_{i=1}^{r+1} \frac{1}{n-i+1}$ and $\beta_{r+1} = \sum_{i=1}^{r+1} \frac{1}{(n-i+1)^2}$. Define

$$h_i = \begin{cases} \frac{\alpha_{r+1}}{[\beta_{r+1}(m-r-1)+\alpha^2_{r+1}](n-i+1)} & \text{for } i = 1,2,\ldots,r+1, \\ \frac{\beta_{r+1}}{\beta_{r+1}(m-r-1)+\alpha^2_{r+1}} & \text{for } i = r+2,\ldots,m. \end{cases}$$

Then, from the property of normalized spacings and the algorithm of Huffer and Lin (2001) again, we can easily find the exact values of $t$ satisfying $P(\sigma^*_g/\sigma > t) = P(\sum_{i=1}^{m} h_i Z_i > t) = \alpha$ for specified values of $\alpha$.

## 5 Illustrative Examples

**Example 1:** This example comes from Lawless (1982, pp. 138) with $n = 12, r = 2, s = 1$, and the doubly Type-II censored sample is as follows:

$$--, --, 24.4, 28.6, 43.2, 46.9, 70.7, 75.3, 95.5, 98.1, 138.6, --$$

In this case, we have $\sigma^* = 71.1385$ and $A$ in (2) as

$$A = \left(\frac{9955}{120609}\ \frac{3620}{40203}\ \frac{3982}{40203}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\ \frac{10981}{120609}\right).$$

The approximate 95% confidence interval for $\sigma$ obtained by Balakrishnan and Gupta (1998) was $[22\sigma^*/36.78071, 22\sigma^*/10.98232] = [42.5508, 142.5060]$, which is quite close to our exact 95% confidence interval for $\sigma$ given by $[22\sigma^*/36.79546, 22\sigma^*/10.97781] = [42.5337, 142.5646]$. Let $\ell = 12$. Since $P([X_{(\ell)} - X_{(n-s)}]/\sigma^* > 4.38371) = 0.025$ and $P([X_{(\ell)} - X_{(n-s)}]/\sigma^* > 0.02535) = 0.975$, an exact 95% prediction interval for the last failure time $X_{(12)}$ is obtained as

$$[138.6 + 0.02535 \times 71.1385,\ 138.6 + 4.38371 \times 71.1385] = [140.403,\ 450.451].$$

**Example 2:** Consider Nelson's data (1982, p. 228, Table 6.1) which gives data on times to breakdown of an insulating fluid in an accelerated test conducted at various test voltages. For the purpose of illustrating the methods of inference presented here, we generated the following general progressive Type-II censored sample from the $n = 19$ observations recorded at 34 KV in Nelson's Table 6.1 (with one smallest observation censored and three stages of progressive censoring).

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $X_{i:m:n}$ | – | 0.78 | 0.96 | 1.31 | 2.78 | 4.85 | 6.50 | 7.35 |
| $R_i$ | – | 0 | 3 | 0 | 3 | 0 | 0 | 5 |

If we assume a one-parameter exponential distribution for the data at hand, we get $\sigma_g^* = 9.110$ and $Var(\sigma_g^*) = \sigma^2/7.99937$ from Balakrishnan and Sandhu (1996). Using the method described here, we have $P(\sigma_g^*/\sigma > 0.431705) = 0.975$ and $P(\sigma_g^*/\sigma > 1.802934) = 0.025$. Thus, an exact 95% confidence interval for $\sigma$ is obtained as $[\sigma_g^*/1.820934, \sigma_g^*/0.431705] = [5.002927, 21.102373]$.

## References


Balakrishnan, N. and Aggarwala, R. (2000). *Progressive Censoring: Theory, Methods, and Applications.* Boston: Birkhäuser.

Balakrishnan, N. and Cohen, A. C. (1991). *Order Statistics and Inference: Estimation Methods.* San Diego: Academic Press.

Balakrishnan, N. and Gupta, S. S. (1998). Higher-order moments of order statistics from exponential and right-truncated exponential distributions and applications to life-testing problems. In N. Balakrishnan and C.R. Rao (Eds.), *Handbook of Statistics-17: Order Statistics - Applications*, pp. 25–59. Amsterdam: North-Holland.

Balakrishnan, N. and Sandhu, R. A. (1996). Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type-II censored samples. *Sankhyā, Series B* *58*, 1–9.

Huffer, F. W. (1988). Divided differences and the joint distribution of linear combinations of spacings. *Journal of Applied Probability* *25*, 346–354.

Huffer, F. W. and Lin, C. T. (1999). An approach to computations involving spacings with applications to the scan statistic. In J. Glaz and N. Balakrishnan (Eds.), *Scan Statistics and Applications*, pp. 141–163. Boston: Birkhäuser.

Huffer, F. W. and Lin, C. T. (2001). Computing the joint distribution of general linear combinations of spacings or exponential variates. *Statistica Sinica* *11*(4), 1141–1157.

Lawless, J. F. (1971). A prediction problem concerning samples from the exponential distribution, with application in life testing. *Technometrics* *13*, 725–730.

Lawless, J. F. (1977). Prediction intervals for the two-parameter exponential distribution. *Technometrics* *17*, 255–261.

Lawless, J. F. (1982). *Statistical Models & Methods for Lifetime Data.* New York: Wiley.

Likeš, J. (1974). Prediction of sth ordered observation for the two-parameter exponential distribution. *Technometrics* *16*, 241–244.

Lin, C. T. (1993). The computation of probabilities which involve spacings, with applications to the scan statistic. Ph.D. Thesis, Florida State University, Tallahassee.

Lingappaiah, G. S. (1973). Prediction in exponential life testing. *Canadian Journal of Statistics* *1*, 113–118.

Nelson, W. (1982). *Applied Life Data Analysis.* New York: Wiley.

Viveros, R. and Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data. *Technometrics* *36*, 84–91.




# Reliability Sampling Plans Under Progressive Censoring.


**Uditha Balasooriya*[1] and N. Balakrishnan+**

*Nanyang Business School, Nanyang Technological University, Nanyang Avenue, Singapore 639798

+ Department of Mathematics & Statistics, MacMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1

[1] Presently on leave from the Department of Statistics and Applied Probability, National University of Singapore, Singapore 119260


## ABSTRACT


In evaluating the reliability of a product or system, an important quality variable is the lifetime of a product. Our focus is on the design of sampling plans for life testing on products and components to determine their ability to perform the intended purpose for the desired period of time. These sampling plans are also useful for many other experiments.

Variable-sampling plans such as MIL-STD-414 can be only applicable if all specimens in the sample are tested to failure. However, for many high reliability products that are designed to operate for long periods of time, testing under normal use conditions might be exceedingly time consuming and expensive. In Type I censoring experiments, all specimens are put on test simultaneously and the test is terminated at a prespecified time, while in Type II censoring experiments, the test is terminated when a prespecified number of failures is observed. In some experimentations, successive censoring also occur at many stages of the experiment due to various factors. For example, to study the process of deterioration of a product or a component, it might be necessary to destructively test specimens withdrawn at various stages of the experiment. The specimens may be withdrawn at the time of failure of a few units (Type II progressive censoring) at prespecified