International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 2, 2003, pp 1–10
© *2003 Dept. of Statistics, Univ. of Rajshahi, Bangladesh*

An Extension of the Analysis of Marginal Count Failure Data

Md. Rezaul Karim
Department of Statistics,
University of Rajshahi, Bangladesh
E-mail: mrkarim@librabd.net

Kazuyuki Suzuki
Department of Systems Engineering,
The University of Electro-Communications
Tokyo 182-8585, Japan.
E-mail: suzuki@se.uec.ac.jp

[Received September 1, 2003; Revised December 1, 2003; Accepted December 15, 2003]

Abstract

Karim, Yamamoto and Suzuki (2001a) defined the *marginal count failure data* for warranted products in which data arise from two separate independent databases. This data is incomplete as it does not give the exact number of failures of the specific products that were sold in a particular month and thus it can not distinguish the failures of different ages. The properties of the estimators proposed by Karim, et al. (2001a) based on nonhomogeneous Poisson process for repairable products and Poisson approximation model for nonrepairable products, are investigated more formally in this article.

Keywords and Phrases: Field failure data, Warranty data, Consistency, Bootstrap, MLE, EM algorithm.

1 Introduction

The level of customer satisfaction is one key indicator used by manufacturers to understand how well their products perform in the field. Field reliability data for manufactured products helps us to achieve and improve customer satisfaction. A prime source of field reliability data is a database collected automatically from warranty claims. By collecting and analyzing field reliability data from warranty claims, manufacturers can: predict future claims; determine whether a recall, halt in production, or modification is necessary; ascertain whether product reliability is affected by the manufacturing process or usage environment; and compare failure rates among similar or competing products. Warranty data also relate to costs, causes of failure, and variations in claims based on time and place of manufacturing spot, in addition to usage. However, due to cost constraints and often diffused organizations of service departments or repair service networks, many industrial companies construct warranty databases by gathering data from different sources for a particular time period (Karim, et al., 2001a). For example, Table 1 illustrates a general structure of the data, constructed by combining the monthly sales information, obtained by the sales department, and the number of claims registered for a given month, provided by the service department.

Table 1: General data structure of warranty claims counted monthly

Month of sale, y	No. of products sold in y, N_y	Warranty claims in a calendar time (month, j)							
		1	2	3	$\cdots$	Y	$Y+1$	$\cdots$	T
1	N_1	r_{10}	r_{11}	r_{12}	$\cdots$	$r_{1,Y-1}$	r_{1Y}	$\cdots$	$r_{1,T-1}$
2	N_2		r_{20}	r_{21}	$\cdots$	$r_{2,Y-2}$	$r_{2,Y-1}$	$\cdots$	$r_{2,T-2}$
3	N_3			r_{30}	$\cdots$	$r_{3,Y-3}$	$r_{3,Y-2}$	$\cdots$	$r_{3,T-3}$
$\vdots$	$\vdots$					$\cdots$	$\cdots$	$\cdots$	$\cdots$
Y	N_Y					r_{Y0}	r_{Y1}	$\cdots$	$r_{Y,T-Y}$
No. of claims in month j, r_j		r_1	r_2	r_3	$\cdots$	r_Y	r_{Y+1}	$\cdots$	r_T

Note: N_y and $\{r_j\}$ (marginal count) are observed data; $\{r_{yt}\}$ (complete data) can not be observed.

In Table 1, N_y is the number of products sold in the yth month for $y = 1, 2, \cdots, Y$; $\{r_{yt}\}$ is the number of failures at age t $(t = 0, 1, \cdots, T - y)$ for products sold in month y, where T $(T \geq Y)$ is the number of observed months; and r_j is the count of failures occurring in the jth month, $r_j = \sum_{y=1}^{\min(j,Y)} r_{y,j-y}$, $j = 1, 2, \cdots, T$. Karim, et al. (2001a) defined $\{r_j\}$ as the *marginal count failure data*, which can be observed, and $\{r_{yt}\}$ is the complete data, which can not be observed. The marginal count failure data is incomplete, as it does not give the exact number of failures of the specific products that were sold in a particular month, and therefore it can not distinguish failures of different ages.

Karim, et al. (2001a) modeled the marginal count failure data and derived the maximum likelihood estimates of the models parameters. In this article we study some properties of those estimators and show that the MLEs of the parameters are consistent. Also we define the bootstrap estimates of sample variances and show that bootstrap estimates yield better precision than the asymptotic variances.

2 Modeling and Estimation

It can be assumed that the distributions of the numbers of failures for different monthly sales at a certain age are the same for the products which are produced according to the same design and specifications. In real situation, the proportions of the observed number of failures at a certain age for each monthly sales are not always the same, because of the effects of usage environments or the effects of unfavorable seasons. And these effects are strong, especially for the products which are used outside of houses, buildings, etc. However, for some products, such as television, personal computer, cellular phone, etc., these effects are negligible, and it is reasonable to assume that the distributions of the numbers of failures for different monthly sales at a certain age are the same. That is, in Table 1, the distributions of $\{r_{yt}\}$ and $\{r_{zt}\}$ are the same for all y and z at each t. This implies the assumption that the distributions of the numbers of failures for different monthly sales at a certain age are the same in the sense that the samples have come from the same population, and the expected number of failures per product depends on the age of the product and is independent of other factors. This assumption is also used in Karim, et al. (2001a) to model the numbers of failures.

To model the failure counts for repairable products, we define the parameter, q_t, as the expected number of failures per product at age t, $t = 0, 1, \cdots, T-1$, and assume that, for each sales month y, the $\{r_{yt}\}$, $t = 0, 1, \cdots, T-y$, in Table 1, are independently distributed as Poisson distributions with mean $\{N_y q_t\}$. This is a discretization to monthly intervals of an age-based nonhomogeneous Poisson process. Karim, et al. (2001a & 2001b) and Suzuki, Karim and Wang (2001) also proposed this model to analyze warranty claims data.

If all the complete data, r_{yt}'s, are available, the complete data log likelihood function can be written as

$$\log L_c(q_t; r_{yt}) \propto \sum_{y=1}^{Y} \left\{ -\sum_{t=0}^{T-y} N_y q_t + \sum_{t=0}^{T-y} r_{yt} \log(N_y q_t) \right\}. \tag{1}$$

However, we consider situation where only monthly marginal counts r_j's are available, where $r_j = \sum_{y=1}^{\min(j,Y)} r_{y,j-y},\ j = 1, 2, \cdots, T$, are the sum of the complete data $r_{y,j-y}$. Under the above model (1), r_j, $j = 1, 2, \cdots, T$, are the sum of independent Poisson variables which are again distributed as Poisson with means equal to the sum of the means of the variables $r_{y,j-y}$, $j = 1, 2, \cdots, T$; $y = 1, 2, \cdots, \min(j, Y)$. That is, r_j's are independently distributed as Poisson distributions with means $\sum_{y=1}^{\min(Y,j)} N_y q_{j-y}$, $j = 1, 2, \cdots, T$, and the observed data log likelihood function becomes

$$\log L(q_t; r_j) \propto \sum_{j=1}^{T} \left\{ -\sum_{y=1}^{\min(j,Y)} N_y q_{j-y} \right\} + \sum_{j=1}^{T} \left\{ r_j \log \left(\sum_{y=1}^{\min(j,Y)} N_y q_{j-y} \right) \right\}. \tag{2}$$

Karim, et al. (2001a & 2001b) and Suzuki, et al. (2001) also considered this model.

Karim et al. (2001a) proposed a multinomial model for nonrepairable products and applied the EM algorithm to maximize the likelihood function and to estimate the probability of failures at age t, denoted by q_t. Furthermore, they suggested a simple Poisson approximation under the multinomial model to reduce the computation time necessary in calculating the conditional expectations in the E-step of the EM algorithm for the multinomial model. The validity of the Poisson approximation is investigated numerically by several simulations. They showed that the observed data likelihood under Poisson approximation model is the same as of the likelihood (2), where q_t is used to represent the probability of failure at age t.

2.1 Unconstrained MLE of the Parameter

Here the MLE of q_t is derived by directly maximizing the log likelihood (2). The score equations are

$$\frac{\partial \log L}{\partial q_t} = \begin{cases} \displaystyle\sum_{j=t+1}^{Y} \left(-N_{j-t} + \frac{r_j N_{j-t}}{\sum_{i=0}^{j-1} N_{j-i}\hat{q}_i} \right) = 0, & \text{if } t = 0, \\ \displaystyle\sum_{j=t+1}^{Y} \left(-N_{j-t} + \frac{r_j N_{j-t}}{\sum_{i=0}^{j-1} N_{j-i}\hat{q}_i} \right) \\ \qquad + \displaystyle\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)} \left(-N_{j-t} + \frac{r_j N_{j-t}}{\sum_{i=j-Y}^{j-1} N_{j-i}\hat{q}_i} \right) = 0, & \text{if } 0 < t < Y, \\ \displaystyle\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)} \left(-N_{j-t} + \frac{r_j N_{j-t}}{\sum_{i=j-Y}^{j-1} N_{j-i}\hat{q}_i} \right) = 0, & \text{if } t \geq Y. \end{cases} \tag{3}$$

These equations gives the MLE $\hat{q}_t$ as

$$\hat{q}_t = \begin{cases} r_1/N_1, & \text{if } t = 0, \\ \left(r_{t+1} - \sum_{y=1}^{\min(Y-1,t)} N_{y+1}\hat{q}_{t-y} \right) /N_1, & \text{if } t = 1, 2, \cdots, T-1. \end{cases} \tag{4}$$

A deficiency of this estimator is that it is liable to produce negative estimates for some probabilities. Particularly, for any given data set of N_y and r_j, if the inequalities

$$r_{t+1} \geq \sum_{y=1}^{\min(Y-1,t)} N_{y+1}\hat{q}_{t-y}$$

do not hold, equation (4) produces negative estimates for $\hat{q}_t$ for $t = 1, 2, \cdots, T-1$. To avoid the problem of negative estimates, in the next section a constrained estimator is developed based on the Expectation-Maximization (EM) algorithm (Dempster, Laird and Rubin, 1977; McLachlan and Krishnan, 1997).

2.2 Constrained MLE of the Parameter via the EM Algorithm

The E-step and M-step of the EM algorithm for model (2) are as follows.

E-step

For this model at the $(k+1)$th iteration in the E-step we have to compute

$$\mathrm{E}_{q_t^{(k)}}[r_{y,j-y}|r_j] = \frac{r_j N_y q_{j-y}^{(k)}}{\sum_{i=1}^{\min(Y,j)} N_i q_{j-i}^{(k)}},\ y = 1,\ 2, \cdots, \min(Y,j), \tag{5}$$

for $j = 1, 2, \cdots, T$.

M-step

At the $(k+1)$th iteration, the M-step finds

$$\hat{q}_t^{(k+1)} = \frac{\sum_{y=1}^{\min(Y,T-t)} \mathrm{E}_{\hat{q}_t}^{(k)}[r_{yt}|r_{y+t}]}{\sum_{y=1}^{\min(Y,T-t)} N_y},\ t = 0, 1, \cdots, T-1. \tag{6}$$

Iterating between (5) and (6) until it meets a convergence criterion, the EM algorithm finds the MLE of q_t, $t = 0, 1, \cdots, T-1$. An attractive feature of the estimators obtained by the EM algorithm is that the positivity constraints on the estimates of q_t are automatically satisfied, providing that the initial estimates $q_t^{(0)}$ are all positive for all t, $t = 0, 1, \cdots, T-1$. Although it gives positive estimates, however sometimes it belongs on the lower boundary of the parameter space $\Omega = (0, 1)$. More details on the above estimation techniques are also discussed in Karim, et al. (2001a).

3 Properties of the Estimators

3.1 The Consistency of the MLE of the Parameters

Using the marginal data score function from equation (3) under the condition $0 < t < Y$, and replacing the marginal data, $\{r_j\}$, with the sum of complete data, $\{r_{yt}\}$, we have

$$\begin{aligned}
&\frac{1}{\prod_{y=1}^{Y} N_y}\frac{\partial \log L}{\partial q_t}\\
&\quad= \frac{1}{\prod_{y=1}^{Y} N_y}\sum_{j=t+1}^{Y}\left\{-N_{j-t}+\frac{N_{j-t}\sum_{i=0}^{j-1} r_{j-i,i}}{\sum_{i=0}^{j-1} N_{j-i}q_i}\right\}\\
&\quad+\frac{1}{\prod_{y=1}^{Y} N_y}\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)}\left\{-N_{j-t}+\frac{N_{j-t}\sum_{i=j-Y}^{j-1} r_{j-i,i}}{\sum_{i=j-Y}^{j-1} N_{j-i}q_i}\right\}\\
&\quad= \sum_{j=t+1}^{Y}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=0}^{j-1} N_{j-i}q_i}\sum_{i=0}^{j-1}\left(\frac{r_{j-i,i}/N_{j-i}}{\prod_{y=1}^{j-i-1} N_y\prod_{y=j-i+1}^{Y} N_y}\right)\right\}\\
&\quad+\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=j-Y}^{j-1} N_{j-i}q_i}\sum_{i=j-Y}^{j-1}\left(\frac{r_{j-i,i}/N_{j-i}}{\prod_{y=1}^{j-i-1} N_y\prod_{y=j-i+1}^{Y} N_y}\right)\right\}.
\end{aligned}\tag{7}$$

Here we note that

$$\operatorname*{p\lim}_{N_y\to\infty,\forall y}\left(\frac{r_{yt}}{N_y}\right)=q_t. \tag{8}$$

Taking the probability limit in (7) and using (8), we obtain

$$\begin{aligned}
&\operatorname*{p\lim}_{N_y\to\infty,\forall y}\frac{1}{\prod_{y=1}^{Y} N_y}\frac{\partial \log L}{\partial q_t}\\
&\quad= \sum_{j=t+1}^{Y}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=0}^{j-1} N_{j-i}q_i}\sum_{i=0}^{j-1}\left(\frac{q_i}{\prod_{y=1}^{j-i-1} N_y\prod_{y=j-i+1}^{Y} N_y}\right)\right\}\\
&\quad+\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=j-Y}^{j-1} N_{j-i}q_i}\sum_{i=j-Y}^{j-1}\left(\frac{q_i}{\prod_{y=1}^{j-i-1} N_y\prod_{y=j-i+1}^{Y} N_y}\right)\right\}\\
&\quad= \sum_{j=t+1}^{Y}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=0}^{j-1} N_{j-i}q_i}\sum_{i=0}^{j-1}\left(\frac{N_{j-i}q_i}{\prod_{y=1}^{Y} N_y}\right)\right\}\\
&\quad+\sum_{j=\max(t+1,Y+1)}^{\min(T,Y+t)}\left\{-\frac{N_{j-t}}{\prod_{y=1}^{Y} N_y}+\frac{N_{j-t}}{\sum_{i=j-Y}^{j-1} N_{j-i}q_i}\sum_{i=j-Y}^{j-1}\left(\frac{N_{j-i}q_i}{\prod_{y=1}^{Y} N_y}\right)\right\}=0.
\end{aligned}\tag{9}$$

If $t=0$ or if $t \geq Y$, the proofs follow immediately from (9). Hence the MLE, $\hat{q}_t$ obtained from model (2), is a consistent estimator of q_t.

3.2 Bootstrap Estimation of Sample Variance

As mentioned in the previous section, the constrained MLE sometimes belongs on the lower boundary of the parameter space, $\Omega = (0, 1)$. In this situation, the marginal asymptotic variance estimation by using Louis (1982) formula, given in Karim, et al. (2001a), is not appropriate, especially if the N_y or the observed number of failures, r_j is not sufficiently large. Therefore, this section discusses the bootstrap method, introduced by Efron (1979), to estimate the sample variances of the estimators, and compares these variances with that of the asymptotic variances by simulations. To compare the bootstrap estimates of sample variances with asymptotic variances, here we use the following three sets of parameters given in Table 2, which express constant, increasing and decreasing mean number of failures per product. Monthly sales amounts, $N_y = 5,000$ are considered for $y = 1, 2, 3$. Karim, et al. (2001a) also considered the same parameters settings for Monte Carlo simulations.

Table 2: Parameter sets for simulation studies

Set No.	q_t		
	q_0	q_1	q_2
1	0.005	0.005	0.005
2	0.003	0.005	0.007
3	0.007	0.005	0.003

The parametric bootstrap approach (e.g., see, Efron and Tibshirani, 1993; Meeker and Escobar, 1998) is applied with the following steps:

Step 1. Generate r_1, r_2, r_3, from the model (2), with $N_y = 5,000$, $y = 1, 2, 3$, and the true values of q_t, $t = 0, 1, 2$, given in Table 2, and put them as $\tilde{r}_j$, $j = 1, 2, 3$.

Step 2. Estimate q_t, $t = 0, 1, 2$, from $\tilde{r}_j$, $j = 1, 2, 3$, using the algorithm proposed in Section 2.2, and put these estimates as $\hat{\mathbf{q}} = (\hat{q}_0, \hat{q}_1, \hat{q}_2)'$.

Step 3. Generate B replications of r_1, r_2, r_3, from the model (2), with N_y, $y = 1, 2, 3$, and the estimates $\hat{q}_t$, $t = 0, 1, 2$, and put them as r^*_{jb}, $j = 1, 2, 3$; $b = 1, 2, \cdots, B$. These are called the bootstrap sample.

Step 4. Again estimate q_t, $t = 0, 1, 2$, for each set $\{r^*_{1b}, r^*_{2b}, r^*_{3b}\}$, $b = 1, 2, \cdots, B$, using the same algorithm as in Step 2. And obtain B sets of $\hat{\mathbf{q}}^*_b = (\hat{q}^*_{0b}, \hat{q}^*_{1b}, \hat{q}^*_{2b})'$. These are called the bootstrap replications of estimates of $\mathbf{q}$.

Step 5. Calculate the sample covariance matrix of $\hat{\mathbf{q}}^*_b$ as

$$\hat{\mathbf{V}} = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\mathbf{q}}^*_b - \bar{\hat{\mathbf{q}}}^*)(\hat{\mathbf{q}}^*_b - \bar{\hat{\mathbf{q}}}^*)', \tag{10}$$

where $\bar{\hat{\mathbf{q}}}^* = \sum_{b=1}^{B} \hat{\mathbf{q}}_b^*/B$.

The usual parametric bootstrap contains only Steps 2 – 5. However, as Step 1 uses the known parameters, to evaluate the precision of this approach, Steps 1 – 5 are repeated several times.

Table 3: Comparison of variance estimates for the constrained MLEs

Set No.		NHPP parameters		
		q_0	q_1	q_2
1	Sample variance†	0.0102	0.0296	0.0484
	Sample mean of mar. avar†	0.0100	0.0299	0.0499
	Bootstrap estimates: average of $\widehat{sv}_B$	0.0098	0.0281	0.0454
2	Sample variance†	0.0060	0.0220	0.0451
	Sample mean of mar. avar†	0.0060	0.0220	0.0460
	Bootstrap estimates: average of $\widehat{sv}_B$	0.0060	0.0215	0.0452
3	Sample variance†	0.0140	0.0354	0.0444
	Sample mean of mar. avar†	0.0139	0.0377	0.0536
	Bootstrap estimates: average of $\widehat{sv}_B$	0.0134	0.0328	0.0408

Note: † copied from Table 6 of Karim, et al. (2001a); all variances should be multiplied by 10^{-4}; "mar. avar" means "marginal asymptotic variance".

Based on the true values of Table 2, the bootstrap results obtained from 1000 simulations with $B = 500$ are summarized in Table 3. This table compares the marginal asymptotic variances and the bootstrap variances. For example, for set 3 the bootstrap estimate of sample variance of q_2 is 0.0408×10^{-4} whereas the sample mean of marginal asymptotic variance is 0.0536×10^{-4}. The bootstrap estimates of sample variances are smaller compared with the marginal asymptotic variances, particularly for Set 3 where the constrained MLE gives the estimates on the boundary of the parameter space with highest frequency out of the three sets (Karim, et al., 2001a). Therefore, Table 3 suggests that when the maximum likelihood estimators belong on the lower boundary of the parameter space, the bootstrap estimate of sampling variance yields better precision than the asymptotic variance.

4 An Example

In this example, the same approach as discussed in Section 3.2, is applied to the data set of warranty claims of an electronic equipment given in Karim, et al. (2001a) in which $(N_1, N_2, N_3) = (4496, 9296, 7235)$ and $(r_1, r_2, r_3) = (1, 15, 121)$ to compare the bootstrap estimates of the variances using the unconstrained and constrained MLEs.

For the example data set, the Step 1 and Step 2 (see Section 3.2) are replaced by a single step in which both the unconstrained (Section 2.1) and constrained (Section 2.2) MLEs of q_t, $t = 0, 1, 2$, are computed based on the observed data. The resulting estimates of (q_0, q_1, q_2) are (0.000222, 0.002876, 0.020608) for unconstrained MLEs and (0.000222, 0.002876, 0.020607) for constrained MLEs. In Step 3, the bootstrap samples are generated from model (2). Both the unconstrained and constrained methods are also used in Steps 3–5, and the procedure is stopped after the Step 5.

Table 4: Comparison of bootstrap variance estimates using the unconstrained and constrained estimation methods

B	Bootstrap estimates of	Unconstrained method			Constrained method		
		q_0	q_1	q_2	q_0	q_1	q_2
200	sample mean, $\hat{\mathbf{q}}^*$	0.00023	0.00274	0.02088	0.00023	0.00274	0.02088
	sample variance, $\widehat{sv}_B$	0.00050	0.01071	0.10118	0.00047	0.01047	0.10037
500	sample mean, $\hat{\mathbf{q}}^*$	0.00023	0.00285	0.02057	0.00023	0.00285	0.02057
	sample variance, $\widehat{sv}_B$	0.00046	0.00997	0.09668	0.00045	0.00987	0.09634
1000	sample mean, $\hat{\mathbf{q}}^*$	0.00023	0.00284	0.02067	0.00023	0.00284	0.02067
	sample variance, $\widehat{sv}_B$	0.00047	0.01021	0.09866	0.00047	0.01016	0.09849
1500	sample mean, $\hat{\mathbf{q}}^*$	0.00023	0.00285	0.02061	0.00023	0.00285	0.02061
	sample variance, $\widehat{sv}_B$	0.00047	0.00963	0.09820	0.00046	0.00958	0.09805
2000	sample mean, $\hat{\mathbf{q}}^*$	0.00023	0.00285	0.02064	0.00023	0.00285	0.02064
	sample variance, $\widehat{sv}_B$	0.00048	0.00950	0.09777	0.00048	0.00946	0.09765

Note: all variances should be multiplied by 10^{-4}; B is bootstrap replications.

Table 4 provides a comparison of the unconstrained and constrained results for different choices of bootstrap replications, B equals 200, 500, 1,000, 1,500 and 2,000. The comparison shows that, the both methods are in close agreement for all B's. This is because for this example data, the unconstrained MLEs of q_t's are positive and the constrained MLEs are not close to zero, and the estimates obtained using the two methods are almost the same. Table 4 also indicates that the effect on the choice of B is not so significant for this data.

5 Concluding Remarks

Marginal counts of failures for a warranted product can be collected easily and cheaply, but manufacturers encounter difficulties in estimating the age-based (age-specific) failure rates of their products, since the data are incomplete and in a form inconvenient for analysis. Karim, et al. (2001a) proposed methods for analyzing marginal count failure data and discussed age-based analysis of warranty claims - the mean number of failures per product at age t for repairable products, and the probability of failures at age t for nonrepairable products.

In this article it is shown that the unconstrained maximum likelihood estimator proposed by Karim, et al. (2001a) is consistent. Also when the constrained MLE belongs to (or very close to) the lower boundary of the parameter space, the bootstrap sampling procedure yields better precision than the asymptotic variance estimation method.

Acknowledgements

The authors are grateful to anonymous referees for useful comments.

References

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society B*, **39** pp. 1-38.

Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Annals of Statistics*, **7**, No. 1, 1–26.

Efron, B., and Tibshirani, R. J. (1993). *An introduction to the bootstrap*, Monographs on Statistics and Applied Probability, **57**, Chapman and Hall, New York.

Karim, M. R., Yamamoto, W., and Suzuki, K. (2001a). Statistical analysis of marginal count failure data, *Lifetime Data Analysis*, Vol. **7**, No. **2**, 173–186.

Karim, M. R., Yamamoto, W., and Suzuki, K. (2001b). Change-point Detection from Marginal Count Failure Data, *Journal of the Japanese Society for Quality Control*, Vol. **31**, No. **2**, 318–338.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm, *Journal of the Royal Statistical Society B* **44**, 226–233.

McLachlan, G. J., and Krishnan, T. (1997). *The EM algorithm and extensions*, John Wiley & Sons, Inc., New York.

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

Suzuki, K., Karim, M. R., and Wang, L. (2001). Statistical Analysis of Reliability Warranty Data, *Handbook of Statistics: Advances in Reliability*, Vol. **20**, eds. N. Balakrishnan and C. R. Rao, Elsevier Science, 585–609.