1994

A comparative Study of Four Estimators for Analyzing the Random Event Rate of the Poisson Process

Sohn, So Young

Taylor & Francis

Sohn, So Young. "A comparative study of four estimators for analyzing the random event rate of the Poisson process." Journal of Statistical Computation and Simulation 49.1-2 (1994): 1-10.
http://hdl.handle.net/10945/60784

J. Statist. Comput. Simul., Vol. 49, pp. 1–10
Reprints available directly from the publisher
Photocopying permitted by license only

A COMPARATIVE STUDY OF FOUR ESTIMATORS FOR ANALYZING THE RANDOM EVENT RATE OF THE POISSON PROCESS

SO YOUNG SOHN

Department of Operations Research, Naval Postgraduate School, Monterey, CA 93943

(Received June 1, 1993; in final form September 5, 1993)

In this paper, a random effect Poisson regression model is considered for the prediction of the failure rate which would follow a lognormal distribution. A two-stage procedure is used to obtain the regression estimator of the failure rate as well as the shrinkage estimator. These estimators are compared to both the raw estimator which entirely depends on the historical failure records and a shrinkage estimator in which a gamma distribution is used mistakenly in place of the lognormal prior distribution. Results of Monte-Carlo simulation indicate the following in terms of the MSE: (1) overall, the shrinkage estimator based on the lognormal prior distribution performs best; (2) with the moderate variability in the failure rates (0–2.5), the performance of the shrinkage estimator based on the gamma distribution is not significantly different from that of the shrinkage estimator based on the lognormal distribution; (3) when there exists considerable variability in the failure rates (0–10), the raw estimator appears to replace shrinkage estimators. In terms of the Bias, the raw estimator performs better than the others.

KEY WORDS: Poisson regression; Lognormal distribution; Shrinkage estimator; Two-Stage Estimation; Maximum likelihood estimation.

1. INTRODUCTION

Consider a collection of N items that independently experience n_i failures in accordance with Poisson processes during time t_i, each with rate parameter λ_i for item i, $i = 1, \ldots, N$. Possibly also available are concomitant observations on other variables x_i that may in part influence the values of the rates λ_i. Often the estimated failure rates of λ_i, n_i/t_i, vary and exhibit extraneous variance, that is variance greater than expected under a Poisson sampling theory. Taking into account this extra-Poisson variability, one can form a regression model for λ_i as a function of random error as well as concomitant variables.

To predict the future failure rates of individual units, a probability distribution should be specified for the random error. This distribution is called a superpopulation for the failure rate, λ_i. Both the lognormal and the gamma distribution have been utilized as the superpopulation for the event rate of the Poisson process in the literature (Hinde, 1982), (Gaver, 1985), (Lawless, 1987) and (Gaver *et al.*, 1990). Depending on the kind of superpopulation (gamma or lognormal), different (Poisson-gamma and Poisson-lognormal) regression models for the failure rates

This paper was prepared in conjunction with research funded by the Naval Postgraduate School.

can be formulated. The features of each regression model are as follows: (1) The Poisson-gamma regression model: The gamma is a natural conjugate prior associated with the Poisson process and hence yields pleasant analytical simplicity. However it may not necessarily represent the actual extra-Poisson variability; (2) The Poisson-lognormal regression model: The lognormal can be a good candidate superpopulation when event rates have potential outliers while the prediction of the future event rates requires numerical integration which can be computationally burdensome. For both gamma and lognormal superpopulations, the shrinkage estimators, $\hat{\lambda}_i^{SG}$ and $\hat{\lambda}_i^{SL}$, which are the composite estimators of the regression estimator, $\hat{\lambda}_i^R$, and the raw estimator, $\hat{\lambda}_i = n_i/t_i$, are derived in (Gaver, 1985), respectively. The raw estimator, $\hat{\lambda}_i$, considers only its own past experience to predict the future failure rate of the item i while the regression estimator, $\hat{\lambda}_i^R$, is obtained in relation to concomitant variables. The shrinkage estimator is the weighted mean of the regression estimator and the raw estimator.

Therefore, in order to obtain shrinkage estimators, both the regression estimator, $\hat{\lambda}_i^R$, and the weights should be estimated. Estimating $\hat{\lambda}_i^R$ and the weights for $\hat{\lambda}_i^{SL}$, unlike $\hat{\lambda}_i^{SG}$, requires numerical integration. It may render blind use of the gamma superpopulation where the random λ_i would actually be associated with the lognormal distribution.

The objective of this paper is two-fold: (1) to use a simplified procedure in order to obtain $\hat{\lambda}_i^{SL}$ and compare the performance of $\hat{\lambda}_i^{SL}$ to that of $\hat{\lambda}_i$ and $\hat{\lambda}_i^R$; and (2) to analyze the impact of using the gamma superpopulation assumption on predicting λ_i which would be associated with a lognormal distribution. First, a two-stage procedure is used to obtain $\hat{\lambda}_i^R$ based on the Poisson-lognormal model and the weight for the shrinkage estimator, $\hat{\lambda}_i^{SL}$. This two-stage estimation procedure is relatively easy to use compared to the procedure used in (Gaver, 1985 and Gaver *et al.*, 1990) which requires the numerical approximation. Next, the three resulting estimators (raw, regression, and shrinkage) are then compared in terms of the mean squared error and squared bias by way of a Monte-Carlo simulation. The same simulation design is used to evaluate the degree of inaccuracy on prediction due to mis-specification of superpopulation (e.g., gamma superpopulation in place of lognormal superpopulation).

In section 2, the Poisson regression model is illustrated based on the lognormal superpopulation. A two-stage method is described to obtain the regression estimator, $\hat{\lambda}_i^R$. In addition, the shrinkage estimators, $\hat{\lambda}_i^{SL}$ and $\hat{\lambda}_i^{SG}$ are given for lognormal and gamma superpopulations, respectively. Section 3 deals with a simulation study. The results are discussed in Section 4.

2. POISSON REGRESSION MODEL

Assume that for an individual i, the number of failures n_i, experienced in a stated period t_i has a Poisson distribution with parameter $\lambda_i t_i$ that constitutes the following within-individual model:

For $i = 1, \ldots, N$

$$n_i \text{ follows Poisson}(\lambda_i t_i) \tag{1}$$

where λ_i is referred to as the failure rate that is a function of concomitant variables $(x_{i1},x_{i2}, \ldots, x_{ip})$, and random error ϵ_i. In order to ensure non-negativity of failure rate, the following between-individual model is assumed:

$$\lambda_i = \exp(\beta_1 + \beta_2 x_{i2} \ldots + \beta_p x_{ip} + \epsilon_i) \quad \text{or}$$
$$\log(\lambda_i) = \beta_1 + \beta_2 x_{i2} \ldots + \beta_p x_{ip} + \epsilon_i \tag{2}$$

where the random error ϵ_i follows the $N(0,\sigma^2)$ or $\exp(\epsilon_i)$ follows lognormal distribution.

According to the conventional approach, the likelihood contribution of individual unit i is, up to irrelevant constants,

$$L_i(\sigma^2,\beta';t,n,x) = \int_{-\infty}^{\infty} e^{-\lambda_i t_i}(\lambda_i t_i)^{n_i} \exp(-\epsilon_i^2/2\sigma^2)\, d\epsilon_i, \tag{3}$$

where β' is a $1xp$ row vector of regression coefficients, $(\beta_1,\beta_2, \ldots, \beta_p)$.

So the total likelihood is

$$L(\sigma^2,\beta';t,n,x) = \prod_{i=1}^{N} L_i(\sigma^2,\beta';t,n,x). \tag{4}$$

The approximate conditional means and second moments are cases of

$$E(\epsilon_i^k|t_i,n_i,x_i) = \frac{1}{L_i(\hat{\sigma}^2,\hat{\beta}';t,n,x)} \int_{-\infty}^{\infty} \epsilon_i^k e^{-\lambda_i t_i}(\lambda_i t_i)^{n_i} \exp(-\epsilon_i^2/2\sigma^2)\, d\epsilon_i, \tag{5}$$

where $\hat{\sigma}^2$ and $\hat{\beta}$ are maximum likelihood (ML) estimates.

In order to obtain the ML estimtes of regression parameters, Gaver *et al.* (1990) used a variant of the Gauss-Hermite technique that involves an initial correction of the function in equation (3) by Laplace's method. This technique requires iterative procedures of various steps to adjust β's and σ^2 based not only on the initial correction of the likelihood function but also on the regression analysis. This approach appears computationally burdensome and difficult to implement. In an effort to improve on these problems, we consider a two-stage estimation (Vonesh and Carter, 1987) where within-individual parameters are separately estimated from the between-individual regression model.

First, based on the information obtained in the within-individual model (1), the raw estimator $\hat{\lambda}_i$ $(= n_i/t_i)$ is used to estimate λ_i.

Next, $\hat{\lambda}_i$, replaces unobservable λ_i in the between-individual model (2). As a result of replacement, estimation error δ_i that is independent of ϵ_i is added to model (2):

$$\log(\hat{\lambda}_i) = \beta_1 + \beta_2 x_{i2} \ldots + \beta_p x_{ip} + \epsilon_i + \delta_i. \tag{6}$$

Since a delta method approximation to $\text{var}(\log(n_i/t_i)|\lambda_i)$ is $1/\lambda_i t_i \simeq 1/n_i$, the $\log(\hat{\lambda}_i)$ may be regarded as having approximate normal distributions with means $x_i\beta$ and

variances $\sigma^2 + 1/n_i$ where x_i is a $1xp$ row vector of $(1, x_{i2}, \ldots, x_{ip})$. Often σ^2 is not known and is to be estimated using the available data. The maximum likelihood (ML) estimators of σ^2 and β can be obtained from the following log likelihood function after the irrelevant constant term is discarded:

$$\prod_{i=1}^{N} L_i(\sigma^2,\beta';\hat{\lambda},x,n) = -0.5N \log\left(\sigma^2 + \frac{1}{n_i}\right) - \sum_{i=1}^{N} \frac{(\log\hat{\lambda}_i - x_i\beta)^2}{2(\sigma^2 + 1/n_i)}. \tag{7}$$

To solve the normal equations of (7), the following ordinary least square estimator of β and the resulting $\hat{\sigma}_o^2$ can be used as initial values of the ML estimates, $\hat{\beta}$ and $\hat{\sigma}^2$, respectively:

$$\hat{\beta}^o = (X'X)^{-1}(X' \log\hat{\lambda}) \text{ and } \hat{\sigma}_o^2 = \sum_{i=1}^{N} (\log\hat{\lambda}_i - x_i\hat{\beta}^o)^2/N$$

where x is an NxP matrix of x_i's and $\log\hat{\lambda}$ is an $Nx1$ vector of $\log\hat{\lambda}_i$. A closed form expression for the ML estimate, $\hat{\beta}$, is $(X'\hat{W}X)^{-1}(X'\hat{W}\log\hat{\lambda})$ where an $N \times N$ diagonal matrix $\hat{W}$ consists of diagonal elements, $(\hat{\sigma}^2 + 1/n_i)$. This is essentially the same expression as what Graver *et al.* (1990) derived as the ML estimates of β except for the way $\hat{\sigma}^2$ is derived. We use equation (7) to estimate σ^2 while they use numerical approximation of equation (4) for the same purpose.

Once the ML estimates, $\hat{\beta}$ and $\hat{\sigma}^2$ are obtained from (7), the regression estimator, $\hat{\lambda}_i^R$, can be derived by placing β in (6) with ML estimators, $\hat{\beta}$:

$$\log(\hat{\lambda}_i^R) = x_i\hat{\beta} \tag{8}$$

The regression estimator, $\hat{\lambda}_i^R$, along with $\hat{\lambda}_i$ as well as $\hat{\sigma}^2$ is, in turn, used as the input for the shrinkage estimator, $\hat{\lambda}_i^{SL}$, derived in (Gaver, 1985):

$$\hat{\lambda}_i^{SL} = \exp[(n_i \log(n_i/t_i) + x_i\hat{\beta}/\hat{\sigma}^2)/(n_i + 1/\hat{\sigma}^2)] \tag{9}$$

In order to evaluate the impact of mis-specified superpopulation (gamma in place of lognormal) the following shrinkage estimator, $\hat{\lambda}_i^{SG}$, (Gaver, 1985) is also considered.

$$\hat{\lambda}_i^{SG} = \frac{(\exp(x_i\beta^*)t_i/\hat{E}(u_i))(n_i/t_i) + (\hat{E}(u_i)\exp(x_i\beta^*))/\hat{V}(u_i)}{\exp(x_i\beta^*)t_i/\hat{E}(u_i) + 1/\hat{V}(u_i)} \tag{10}$$

where $u_i = \exp(\epsilon_i)$ which follows the gamma distribution with mean $E(u_i)$ and variance $V(u_i)$. We assume one parameter (α) gamma distribution on u_i where the probability density function of u is,

$$g(u;\alpha) = \exp(-u/\alpha)(u/\alpha)^{\alpha^{-1}-1}/\Gamma(\alpha^{-1})\alpha.$$

In such a case, $E(u_i) = 1$ and $V(u_i) = \alpha$. The maximum likelihood estimates, β^* and $\hat{V}(u_i)$ can be obtained from the following likelihood function:

$$L(\alpha,\beta:n,t) = \prod_{i=1}^{N} \frac{\Gamma(n_i + 1/\alpha)}{n_i!\Gamma(1/\alpha)} \left(\frac{\alpha \exp(x_i\beta)t_i}{\alpha \exp(x_i\beta)\, t_i + 1}\right)^{n_i} \left(\frac{1}{\alpha \exp(x_i\beta)t_i + 1}\right)^{1/\alpha} \tag{11}$$

Notice that the estimated empirical Bayes posterior mean and variance of the λ_i are as follows:

$$\hat{\lambda}_i^{SG} = \frac{\exp(x_i\beta^*)(n_i + 1/\hat{\alpha})}{\exp(x_i\beta^*)t_i + 1/\hat{\alpha}}, \tag{12}$$

$$\hat{var}(\lambda_i|n_i,x_i) = \frac{(\exp(x_i\beta^*))^2(n_i + 1/\hat{\alpha})}{(\exp(x_i\beta^*)t_i + 1/\hat{\alpha})^2} \tag{13}$$

3. SIMULATION

The objective of the following simulation is to compare the accuracy of the raw, regression and shrinkage estimators ($\hat{\lambda}_i, \hat{\lambda}_i^R, \hat{\lambda}_i^{SL}$, and $\hat{\lambda}_i^{SG}$) of the random event rate which would be associated with the lognormal distribution.

Simulation parameters are set following the design used in (Gaver *et al.*, 1990) except for the distribution of ϵ_i. The basic simulation design (SD1) is as follows:

1. Set $N = 20$ and $p = 2$.
2. For $i = 1, \ldots, 20$, generate ϵ_i from i.i.d. $N(0,1)$.
3. Use $x_{i2} = 1$ for $i = 1, \ldots, 10$ and $x_{i2} = -1$ for $i = 11, \ldots, 20$.
4. Fix $\beta_1 = 0.5$, $\beta_2 = 0.1$, while $t_i = 2$ for $i = 1, \ldots, 20$.
5. Obtain $\lambda_i = \exp(\beta_1 + \beta_2 x_{i2} + \epsilon_i)$ for $i = 1, \ldots, 20$.
6. Generate $(n_i)^r$ from $Poisson(\lambda_i t_i)$ for each i, where $r = 1, \ldots, 100$.
7. Calculate the raw estimates, $(\hat{\lambda}_i)^r = (n_i)^r/t_i$ for each i where $r = 1, \ldots, 100$.
8. For each r, use N observations of x, and $(\hat{\lambda}_i)^r$ to obtain initial estimates, $(\hat{\beta}^o)^r = (X'X)^{-1}(X' \log(\hat{\lambda})^r)$; and $(\hat{\sigma}_o^2)^r = \Sigma_{i=1}^N (\log(\hat{\lambda}_i^r) - x_i(\hat{\beta}_o)^r)^2/N$.
9. Obtain ML estimates of $(\hat{\beta})^r$ and $(\hat{\sigma}^2)^r$ using equation (7) for $r = 1, \ldots, 100$.
10. Calculate, $(\hat{\lambda}_i^R)^r = x_i(\hat{\beta})^r$ for $r = 1, \ldots, 100$.
11. By replacing n_i, $\hat{\beta}$, $\hat{\sigma}^2$, β^* and $\hat{\alpha}$ in (9) and (12) with the corresponding n_i^r, $(\hat{\beta})^r$, $(\hat{\sigma}^2)^r$, $(\beta^*)^r$ and $(\hat{\alpha})^r$ obtain $(\hat{\lambda}_i^{SL})^r$ and $(\hat{\lambda}_i^{SG})^r$, respectively, for $r = 1, \ldots, 100$.
12. For each estimate, $\hat{\lambda}_i^*$, evaluate $\text{MSE}(\hat{\lambda}_i^*) = \Sigma_{r=1}^{100}((\hat{\lambda}_i^*)^r - \lambda_i)^2/100$ and the squared bias, $\text{BIAS}^2(\hat{\lambda}_i^*) = (E(\hat{\lambda}_i^*) - \lambda_i)^2$ where $\hat{\lambda}_i^*$ represents $\hat{\lambda}_i$, $\hat{\lambda}_i^R$, $\hat{\lambda}_i^{SL}$, and $\hat{\lambda}_i^{SG}$, respectively, and the $E(\hat{\lambda}_i^*) = \Sigma_{r=1}^{100}(\hat{\lambda}_i^*)^r/100$.

Table 1 Simulation design

SD	β_2	σ^2
1	0.1	1
2	0.1	0.1
3	0.1	0.01
4	0.3	1
5	0.3	0.1
6	0.3	0.01

13. Calculate the average $\text{MSE}(\hat{\lambda}^*) = \Sigma_{i=1}^{20} \text{MSE}(\hat{\lambda}_i^*)/20$ and the average $\text{BIAS}^2(\hat{\lambda}^*) = \Sigma_{i=1}^{20} \text{BIAS}^2(\hat{\lambda}_i^*)/20$.
14. Repeat above procedures (1–13) 100 times by changing the seed used in step 2.
15. Obtain grand average MSE's and grand average BIAS^2's of λ, λ^R, $\hat{\lambda}^{SL}$ and $\hat{\lambda}_i^{SG}$, respectively, over 100 replicated simulation of procedures (1–13):
Grand average $\text{MSE}(\hat{\lambda}^*) = \Sigma_{l=1}^{100}$ (average $\text{MSE}(\hat{\lambda}^*))^l/100$;
Grand average $\text{BIAS}^2(\hat{\lambda}^*) = \Sigma_{l=1}^{100}$ (average $\text{BIAS}^2(\hat{\lambda}^*))^l/100$.

Random numbers are generated using the available routines in SAS (1989), and the optimization package GAMS (1988) is used for the evaluation of maximum likelihood estimates. Several variations of the basic simulation design, SD1, are considered to observe how the performance of each estimator changes as the regression parameters change. Since the λ_i vary depending on the β and σ^2 with fixed x_i, we consider a factorial design which consists of six combinations of the values of β_2 (0.1 and 0.3) and σ^2 (1, 0.1, and 0.01). The six simulation designs (SD1, SD2, SD3, SD4, SD5, and SD6) which correspond to these combinations are given in Table 1.

For each design, simulation is replicated 100 times and the resulting grand average MSE and the grand average squared BIAS of each estimation method are summarized in Tables 2 and 3.

4. DISCUSSION

Accuracy of the various estimation methods is evaluated in terms of the grand average MSE and squared bias using a Monte Carlo simulation. Results of the factorial experiment indicate that there is a significant interaction effect between estimators and σ^2 on both MSE and BIAS at significance level 5%, while the interaction effect between estimators and β_2 turns out to be insignificant under the experimental design considered in the simulation.

Figure 1 shows the overall pattern of the interaction effect between the estimators and the σ^2 in terms of the grand average MSE. First, performances of the three estimators are analyzed in terms of the grand average MSE. When σ^2 varies between 0.01 and 0.1 (SD3,6 and SD2,4), $\hat{\lambda}^{SL}$ performs best, while it is followed by the

Table 2 Comparison of four estimators: grand average MSE

SD	$\hat{\lambda}_i$	$\hat{\lambda}_i^R$	$\hat{\lambda}_i^{SL}$	$\hat{\lambda}_i^{SG}$
1	1.69253	1.98912	1.64129	1.85292
2	0.66134	1.61121	0.26291	0.48119
3	0.55222	1.63916	0.03039	0.05192
4	1.81101	1.93191	1.66298	2.05812
5	0.74891	1.68914	0.39255	0.60582
6	0.66536	1.62721	0.05683	0.16929

Table 3 Comparison of four estimators: grand average squared BIAS

SD	$\hat{\lambda}_i$	$\hat{\lambda}_i^R$	$\hat{\lambda}_i^{SL}$	$\hat{\lambda}_i^{SG}$
1	0.01121	0.08109	0.06311	0.08073
2	0.00720	0.07078	0.02411	0.02950
3	0.00882	0.06311	0.00366	0.01401
4	0.01371	0.09912	0.08052	0.09867
5	0.00880	0.08652	0.02947	0.03606
6	0.01078	0.07713	0.00448	0.01713

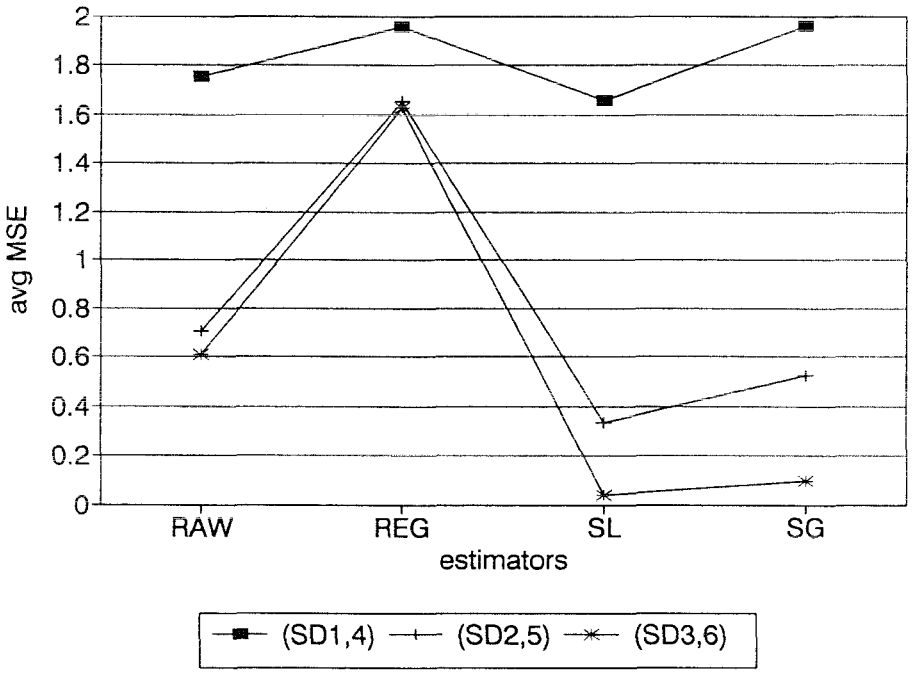

Figure 1 Grand average MSE

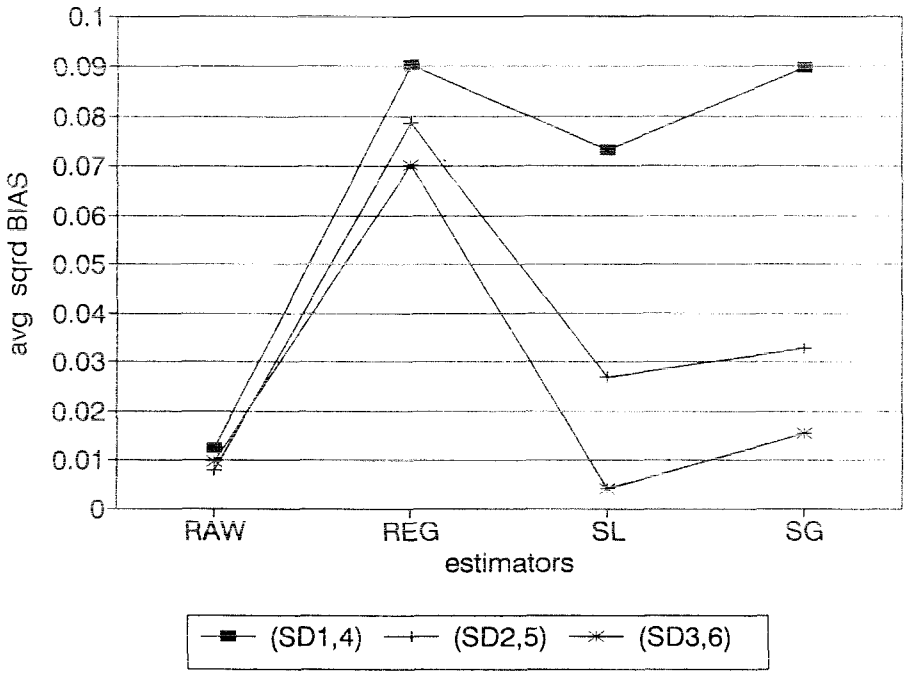

Figure 2 Grand average squared BIAS

performance of $\hat{\lambda}^{SG}$. It is interesting to note that the performance of $\hat{\lambda}^{SG}$ is not significantly different from that of $\hat{\lambda}^{SL}$ in (SD3,6 and SD2,4). It supports the expectation that the performance of these two estimators would be close as σ^2 approaches 0. The worst performance is given by the regression estimator when it interacts with a relatively small σ^2. With a relatively large σ^2 (SD1 and SD4), the grand average MSE of each estimator increases overall. In SD1 and SD4, $\hat{\lambda}^{SL}$ slightly outperforms $\hat{\lambda}$ while $\hat{\lambda}^{SG}$ and $\hat{\lambda}^{R}$ perform worst. In terms of the performance of $\hat{\lambda}^{SG}$, it is observed that the difference between the grand average MSE $\hat{\lambda}^{SG}$ and that of $\hat{\lambda}^{SL}$ in SD1 and SD4 significantly increases compared to differences made by the two estimators in (SD2,5) and (SD3,6), respectively.

Secondly, grand average values of the squared bias of the four estimators are compared in Figure 2. Except for $\hat{\lambda}$, patterns of grand average squared BIAS of the other three estimators are similar to those of grand average MSE. Apparently, $\hat{\lambda}$ provides the minimum bias while it is followed by the $\hat{\lambda}^{SL}$ and $\hat{\lambda}^{SG}$. Again $\hat{\lambda}^{R}$ performs worst.

In order to generalize the results obtained from this simulation study, the patterns of n_i, are considered for the selection of the estimator of the failure rate. Figure 3 shows the patterns of the ordered n_i generated by one replication ($l = 1$ and $r = 1$) of six simulation designs (SD1–SD6). In SD1 and SD4, 70% items experience from 0 to 5 failures while 30% n_i's are dispersed from 6 to 20. The four other designs generate 85% n_i's varying from 0 to 5 and the rest of them varying between 6 and 10. Since the n_i's are generated based on the time period, $t_i = 2$, the patterns

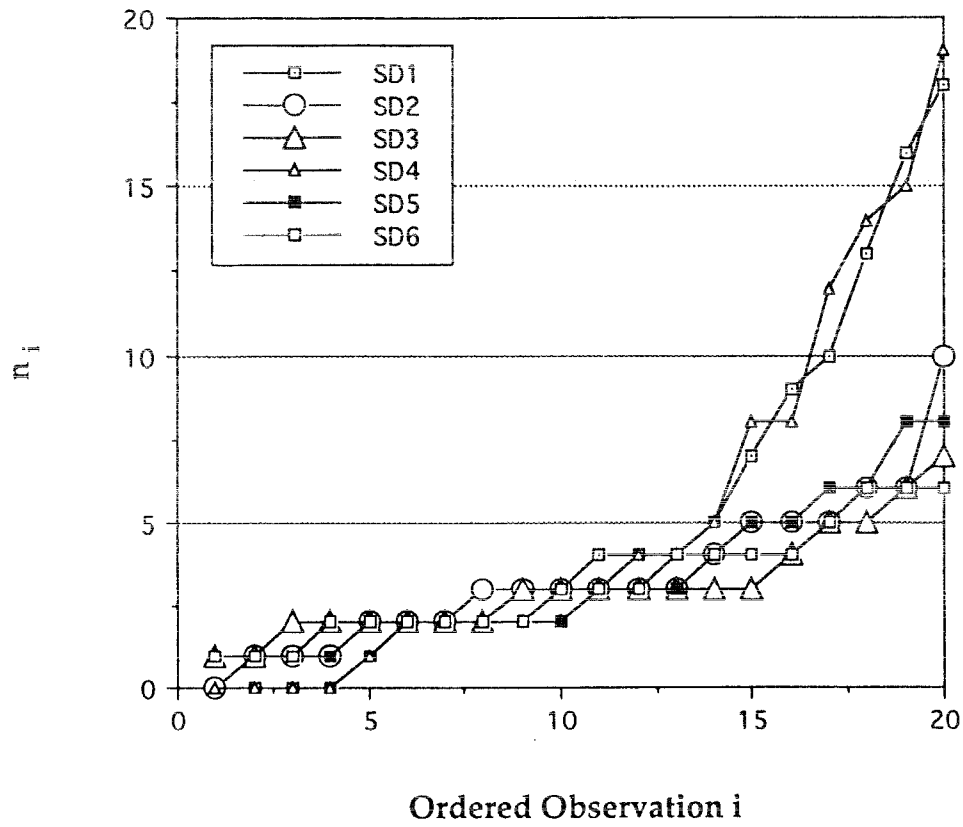

Figure 3 Patterns of n_i

can be represented in terms of $\hat{\lambda}_i$. In SD1 and SD4, 70% $\hat{\lambda}_i$'s are less than 2.5 while 30% $\hat{\lambda}_i$'s are dispersed from 3 to 10. The four other designs (SD2,3,5 and 6) result in 85% $\hat{\lambda}_i$'s varying from 0 to 2.5 and the rest of them varying between 3 and 5.

In sum, regardless of the range of $\hat{\lambda}_i$'s the shrinkage estimator, $\hat{\lambda}^{SL}$ turns out to perform best in terms of the MSE. However, when $\hat{\lambda}_i$'s vary from 0 to 2.5 with the degree of dispersion described in Figure 3, use of $\hat{\lambda}^{SG}$ does not appear to bring significantly different results from that of $\hat{\lambda}^{SL}$. When $\hat{\lambda}_i$'s vary from 0 to 10, the raw estimator can replace the role of $\hat{\lambda}^{SL}$ for the future prediction of failure rate. In addition, use of $\hat{\lambda}$ can enjoy the property of unbiased estimator. Overall, the regression estimator turns out to be the worst among others. It indicates that it might be better to observe the failure rates of new items for a certain time period and use the shrinkage estimates rather than directly employ the regression estimates alone in the prediction of the failure rates. Assessing the robustness of the lognormal superpopulation assumption to other distribution such as student t distribution is left as an area for future study.

References

GAMS (1988). *GAMS User's Guide*. The Scientific Press, Redwood City, CA.

Gaver, D. F. (1985). "Discrepancy-Tolerant Hierarchical Poisson Event-Rate Analysis," *Naval Postgraduate School Technical Report* NPS55-85-016.

Gaver, D. F., Jacobs, P. A. and O'Muricheartaigh, I. G. (1990). "Regression Analysis of Hierarchical Poisson-Like Event Rate Data: Superpopulation Model Effect on Predictors," *Communications in Statistics-Theory and Methodology*, Vol. **19,** No. 10, pp. 3779–3797.

Hinde, J. (1982). "Compound Poisson Regression Models," In GLIM 82: *Proceedings of International Conference in Generalised Linear Models* (R. Gilchrist, ed.), pp. 109–121, Berlin, Springer-Verlag.

Lawless, J. F. (1987). "Regression Methods for Poisson Process Data," *J. of American Statistical Association*, Vol. **82,** No. 399, pp. 808–815.

SAS (1989). *SAS User's Guide, Version 6 Edition*, SAS Institute, Cary, NC.

Vonesh, E. F. and Carter, R. L. (1987). "Efficient Inference for Random Coefficient Growth Curve Models with Unbalanced Data," *Biometrics*, Vol. **43,** pp. 617–628.