S2-38
N91-27002
P-16

COMPOUND ESTIMATION PROCEDURES IN RELIABILITY

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1990

Johnson Space Center

H2086788

Prepared By: Ron Barnes , Ph.D.

Academic Rank: Associate Professor

University & Department: University of Houston-Downtown, Department of Applied Mathematical Sciences, One Main Street, Houston, Texas 77002

NASA/JSC

Directorate: Safety, Reliability and Quality Assurance

Division: Reliability and Maintainability

JSC Colleague: Richard Heydorn, Ph.D.

Date Submitted: August 16, 1990

Contract Number: NGT-44-005-803

ABSTRACT

At NASA, components and subsystems of components in the Shuttle and Space Station generally go through a number of redesign stages. While data on failures for various design stages are sometimes available, the classical procedures for evaluating reliability only utilize the failure data on the present design stage of the component or subsystem. Often, few or no failures have been recorded on the present design stage.

Last Summer, in the NASA Faculty Fellow Program , Bayesian estimators for the reliability of a single component, conditioned on the failure data for the present design, were developed. These new estimators permit NASA to evaluate the reliability, even when few or indeed no failures have been recorded. Point estimates for the later evaluation were not possible with the "classical" procedures.

Since different design stages of a component (or subsystem) generally have a good deal in common, the development of new statistical procedures for evaluating the reliability, which consider the entire failure record for all design stages, has great intuitive appeal.

A typical subsystem consists of a number of different components and each component has evolved through a number of redesign stages.

The investigations this Summer considered compound estimation procedures and related models. Such models permit the statistical consideration of all design stages of each component and thus incorporate all the available failure data to obtain estimates for the reliability of the present version of the component (or subsystem).

A number of models were considered to estimate the reliability of a component conditioned on its **total** failure history from **two** design stages.

It was determined that reliability estimators for the present design stage , conditioned on the complete failure history for two design stages have **lower risk** than the corresponding estimators conditioned only on the most recent design failure data.

Several models were explored and preliminary models involving the bivariate Poisson distribution and the Consael Process (a bivariate Poisson process) have been developed. Possible shortcomings of the models are noted. An example is given to illustrate the procedures.

These investigations are ongoing with the aim of developing estimators that will extend to components (and subsystems) with three or more design stages.

INTRODUCTION

Components in the NASA Shuttle and indeed in many other complex systems often go through a number of redesign stages. Classical reliability estimators rely only on the failure data for the present design. Since previous design stages often have a good deal in common with the present design, statistical procedures for estimating reliability may be improved by also taking into account the failure data on the earlier design versions as well.

Last year in the NASA Summer Faculty Fellow program (SFF), Bayesian estimators for the expected value of the reliability of a single component, conditioned on its failure history (for the present design stage) were developed for the cases of (1) a constant failure rate - the exponential model and (2) a variable failure rate - Weibull model [1],[2].

For the constant failure rate model it was shown that:

$$E[\,R(t) \mid N(T) = n\,] = 1/(1 + t/T)^{n+1} \qquad (1)$$

where

R(t) = the reliability or probability that the component will successfully function up to time t in the future,

N(T) = the number of failures up to time T (in the past failure history), and

n = the number of failures recorded up to time T for this component.

These new estimators enable NASA to evaluate reliabilities when few or even no failures have been recorded. Evaluation in the latter case was not possible with the previous "classical " estimators.

In March 1990, these Bayesian estimators were employed along with the classical estimators in a NASA report on the reliability of Orbiter APU hydraulic hoses [3]. Both sets of estimators predicted a very high reliability for success of the hoses on the next mission (.999..).

This report will focus on investigations to extend the constant failure rate model to utilize the total failure history on a component with 2 or more design stages. The investigations considered compound estimation procedures in order to utilize this failure data history. The report incorporates the Bayesian estimators developed last year and noted earlier in this report.

Motivation for the Study

Since redesign versions of a component would appear to have some commonality, the idea of reliability estimators which incorporate the total failure history of all design stages of a component seems worth considering. When data occurs, for example, on a number of identical valves which have been through 2 design stages with the number of failures on the j-th valve with the first design denoted by N_{1j} and the number on the i-th valve with the new design denoted by N_{2i}, one can plot N_1 versus N_2 The data often suggest some correlation between the two failure counts. In general, the number of failures N_1 and N_2 are not independent. In the discussion to follow, $N_1(T_1)$ and $N_2(T_2)$ the number of failures up to times T_1 and T_2 recorded for the old and new designs respectively are each assumed to have a Poisson distribution with possibly different failure rates λ_1 and λ_2 which are unknown. Since $N_1(T_1)$ and $N_2(T_2)$ are not assumed to be independent, the problem of obtaining Bayesian reliability estimators conditioned on the failure data N_1 and N_2 requires some form of a joint or compound probability distribution for N_1 and N_2 .

The impact of such estimators that utilize the total failure history from 2 or more design stages of a component is indicated by the following result.

Proposition

$$E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 \leqq \quad E[R(t)-E(R(t)|\ N_2(T_2))]^2$$

i.e., the estimator conditioned on two stages of failure history has a **lower risk** than the corresponding risk for the estimator conditioned only on the most recent design failure history. Hence as additional failure data on earlier design stages are included, the corresponding risk decreases.

Proof

$$E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 \ =$$

$$E[R(t)-E(R(t)|N_2(T_2))-\ (E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))]^2 \ =$$

$$\begin{aligned}&E[R(t)-E(R(t)|N_2(T_2))]^2-\ 2E[(R(t)-E(R(t)|N_2(T_2)))\ *\\&\qquad (E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))]\ + \qquad\qquad (2)\\&\qquad E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2 \ =\end{aligned}$$

But since

$$\begin{aligned}&E[(R(t)-E(R(t)|N_2(T_2)))(E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2)))] \ =\\&E\{E[(R(t)-E(R(t)|N_2(T_2)))(E(R(t)|N_1(T_1),N_2(T_2))\ -\\&\qquad E(R(t)|N_2(T_2)))|N_1(T_1),N_2(T_2)]\ \}\ =\\&E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2\end{aligned}$$

then expression (2) becomes

$$\begin{aligned}&E[R(t)-E(R(t)|N_1(T_1),\ N_2(T_2))]^2 \ =\\&E[R(t)-E(R(t)|N_2(T_2))]^2\ -\ E[E(R(t)|N_1(T_1),N_2(T_2))-E(R(t)|N_2(T_2))]^2\end{aligned}$$

Since the last term is non-negative the result is established. Note that this result implies that, if the last term is positive, a strict inequality holds.

DISCUSSION

The present investigations have considered two preliminary models for the distribution of N_1 and N_2 which lead to reliability estimators.

Model A : Bivariate Poisson Model

The first of these employs the bivariate Poisson distribution.. The joint probability function for a 2 dimensional Poisson process $(N_1(t), N_2(t))$ with $\mathrm{cov}(N_1(t), N_2(t)) = \upsilon t$ is given in [4] by:

$$P[\,N_1(t) = n_1,\, N_2(t) = n_2] = f(n_1, n_2, t) \qquad (3)$$

where

$$f(n_1, n_2, t) = \exp(-\lambda_1 t - \lambda_2 t + \upsilon t) \sum_{j=0}^{\min(n_1, n_2)} ((\lambda_1 - \upsilon)^{n_1 - j}(\lambda_2 - \upsilon)^{n_2 - j} *$$

$$t^{n_1 + n_2 - 2j} (\upsilon t)^j) \,/\, [(n_1 - j)!\,(n_2 - j)!\,j!] \qquad (4).$$

Note that in the discussion to follow, in place of t in equation (3) one could use $T = T_1 + T_2$ which is the total elapsed time for failures for both designs. Utilizing expression (4) and Baye's formula,where $X2(t) = 1$ denotes that the component in its second design stage is still operating up to time t, one can show that

$$E[R(t) | N_1(T)=n_1, N_2(T)=n_2] = P[X2(t)=1 \,| N_1(T)=n_1,\, N_2(T))=n_2] =$$

$$P[N_1(T)=n_1, N_2(T)=n_2 \,| X2(t)=1]\; P[X2(t)=1] / P[\,N_1(T) = n_1,\, N_2(T) = n_2] \qquad (5)$$

Assuming a joint Bayesian prior distribution for λ_1, λ_2, this expression becomes:

$$\frac{\int\int P[N_1(T)=n_1, N_2(T)=n_2 \mid X2(t)=1, \lambda_1, \lambda_2]\ P[X2(t)=1 \mid \lambda_1, \lambda_2]\ g(\lambda_1, \lambda_2)\ d\lambda_1, d\lambda_2}{\int\int P[N_1(T)=n_1, N_2(T)=n_2 \mid \lambda_1, \lambda_2]\ g(\lambda_1, \lambda_2)\ d\lambda_1, d\lambda_2} \tag{6}$$

If one assumes the joint prior distribution $g(\lambda_1, \lambda_2)$ to be uniform on finite rectangles with λ_1, λ_2 independent, i.e.

$$g(\lambda_1, \lambda_2) = \left| \begin{array}{ll} (1/\theta_1)\ (1/\theta_2) & \text{for } 0 < \lambda_1 < \theta_1 \text{ and} \\ & \quad 0 < \lambda_2 < \theta_2 \\ 0 & \text{elsewhere} \end{array} \right| \quad \text{then,}$$

one can show that the uniform priors, with the limits extended to the entire first quadrant, give

$$E[R(t)|N_1(T)=n_1, N_2(T)=n_2] =$$

$$\frac{e^{-\upsilon t}\sum_{j=0}^{\min(n_1,\, n_2)} [1/(1+ t/T)^{n_2-j+1}]\,(\upsilon t)^j/j!}{\sum_{j=0}^{\min(n_1,\, n_2)} (\upsilon t)^j/j!} \qquad (7)$$

Note that in the case when $n_1=0$ failures one has

$$E[R(t)|N_1(T)=0, N_2(T)=n_2] = e^{-\upsilon t}(1/(1+ t/T)^{n_2+1}) \qquad (8)$$

where T is the total elapsed time from the start of the first design. Notice the similarity of expression (8) to formula (1) developed previously for the reliability estimator given just the failure data for the second design.

Also observe that if N_1 and N_2 are completely independent then

$$E[R(t)|N_1(T)=n_1, N_2(T)=n_2] = E[R(t)|N_2(T)=n_2] = 1/(1+ t/T)^{n_2+1}.$$

Thus this model would, in some sense, appear to generalize the earlier model(1) which considered only one design stage.

Example

As an application of this reliability estimator consider the following example.

Arbous and Kerrich [5] recorded the number of accidents of 122 individuals (shunters) in two consecutive time periods. For each individual, the number of accidents in the first 6 year period was recorded and then after new insurance and safety procedures were implemented, the number of accidents for the same individual was recorded in the next 5 years. The authors estimated $\upsilon T \approx .257$ for the $T = 11$ years.

One can use expression (7) to evaluate the "reliability" for the next year of an individual,selected at random, given his/her total past accident (failure) history. Note that in this case the reliability is just the probability that a randomly selected individual with the given accident history will not have an accident in the next t years.

For a randomly selected individual with accident history given by

$$n_1=3\ ,\quad n_2 = 1\ ,\quad t=1\ ,\quad \upsilon T \approx .257\ ,\quad T_1 = 6,\quad T_2 = 5,\ T = T_1 + T_2=11$$

one finds by using expression (7) that

$$E[R(t)|N_1(11)=3\ ,N_2(11)=1\ ] = .6067624$$

Thus the reliability estimator obtained from this model suggests that for a randomly selected individual who experienced 3 accidents in the first 6 years and 1 in the next 5 years, the probability of no accidents in the next year is approximately 61 percent. Such information may be used in the setting of insurance premiums for the next year for various classes of individuals based on their past accident histories.

Note that in terms of NASA component reliability:

1. The individuals correspond to different copies of a single component with two design stages.

2. Each copy is located in a somewhat similar environment(e.g. the copies may be located on each of the three space shuttles).

3. Data on failures of each copy were recorded in the first design stage (i.e. the first 6 years of accident data).

4. A second improved design replaced the first and failures were recorded.

Then $E[R(t)|N_1(T_1)=n_1, N_2(T_2)=n_2]$ gives the probability that the second design stage with a given failure history in this location will not fail in the next t time units. Note that the NASA description assumes that when the component fails it is replaced or repaired so that it is equivalent to the original system before the failure.

It should be pointed out that in Model A some assumptions about the relationships between λ_1, λ_2, and υ were made. In particular, $\upsilon < \lambda_1, \lambda_2$ suggests that υ is related to the priors. With this in mind, a second preliminary model has been developed.

Model B : Consael Model

As was noted earlier, N_1 and N_2 are probably not independent in general. The Consael process [6] defines a bivariate compound Poisson process by

$$P[N_1(T)=n_1, N_2(T)=n_2] =$$

$$\int\int e^{-\lambda_1 T} [(\lambda_1 T)^{n_1}]/n_1! \; e^{-\lambda_2 T}[(\lambda_2 T)^{n_2}]/n_2! * g(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 \quad (9)$$

In the Consael process for fixed values of λ_1 and λ_2, N_1 and N_2 correspond to independent Poisson processes while λ_1 and λ_2, have a joint density function $g(\lambda_1,\lambda_2)$.

If one assumes the triangular density function ($\lambda_2 \le \lambda_1$ since one assumes that the newer design is an improvement) one has:

$$g(\lambda_1,\lambda_2) = \left| \begin{array}{ll} 2/\theta_1^{\,2} & \text{for } 0 < \lambda_1 < \theta_1 \text{ and} \\ & \quad 0 < \lambda_2 < \lambda_1 \\ 0 & \quad \text{elsewhere} \end{array} \right| \qquad (10)$$

Utilizing expressions (9) and (10) ,and Baye's formula in expression (6) (and taking limits on the prior distribution) one can again obtain an estimate for the reliability:

$$E[R(t)|N_1(T)=n_1,N_2(T)=n_2]$$

At the present time investigations are continuing with this estimator which has a rather complex , highly combinatorial closed form. It is anticipated that further consideration of this model will indicate approaches to the development of the general model which can incorporate failure data from any number of redesign stages .

SUMMARY AND CONCLUSIONS

Various NASA failure data suggests support for the development of compound/mixture models to estimate the reliability of components that have failure data recorded on more than one design stage. Bayesian estimators that can utilize all "relevant" failure data, even from an earlier design of the component, were investigated.

A Bayesian estimator based on the bivariate Poisson distribution was developed and an example illustrating the technique applied to insurance data was given. The similarity of this example to NASA reliability problems was also noted.

An additional estimator based on the Consael process was developed. Investigations into this model are continuing. It is anticipated that these investigations will lead to a general model to utilize all relevant failure data on a component (or subsystem of components) that has experienced more that two design stages.

REFERENCES

1. Heydorn, Richard, unpublished paper "A Bayesian Approach to Reliability and Confidence," NASA/JSC Reliability and Maintainability Division, Spring 1989.

2. Barnes, Ron ,"A Bayesian Approach to Reliability and Confidence", NASA Summer Faculty Fellow Final Report, August 1989.

3. "A Reliability Assessment of the Orbiter APU Hydraulic Hose", prepared by NASA/JSC Reliability and Maintatinability Division, April 1990.

4. Haight, Frank A., Handbook of the Poisson Distribution, John Wiley and Sons Inc., New York, 1967.

5. Arbous, A. and Kerrich, J. ,"Accident Statistics and the Concept of Accident-Proneness," Biometrics, 7, 340-432, December 1951.

6. Everitt, B.S., and Hand, D.J., Finite Mixture Distributions, Chapman and Hall, London, 1981.

BIBLIOGRAPHY

1. Efron,B. and Morris, C. N., "Stein's Estimation Rule and Its Competitors - An Empirical Bayes Approach," J. Amer. Statist. Assoc. 68, 117-130, 1973.

2. ______________________, "Stein's Paradox in Statistics," Scientific American 236, (5) 119-127, 1977.

3. Haight, Frank A., "Index to the Distributiions of Mathematical Statistics," Journ. Res. of NBS -B.Math 65B, 1, Jan-Mar. 1961.

4. Heydorn, Richard, and Basu, Rehka, unpublished report "Estimating Parameters in a Finite Mixture of Probability Densities"NASA/JSC 1989.

5. Holgate, P. , "Estimates for the Bivariate Poisson Distribution," Biometrika 51, 1&2 , 741, 1964.

6. Martz, Harry F. and Waller, Ray A. , Bayesian Reliability Analysis, John Wiley and Sons Inc., New York, 1982.

7. Stein, Gillian and Juritz, June, "Bivariate Compound Poisson Distributions," Commun. Statist. - Theory Meth. 16(12), 3591-3607, 1987.

8 Stigler, Stephen , "The 1988 Neyman Memorial Lecture: A Galtonian Perspective on Shrinkage," Statistical Science, Vol.5, No. 1, 147-155, Feb. 1990.

9. Teicher, Henry "On the Multivariate Poisson Distribution," Skand. Aktuar. Tidskar. 37, 1-9.

ACKNOWLEDGEMENTS

The author would like to thank the members of the Reliability and Maintainability Division of Chief Duane Duston, especially Tom Edwards and Malcolm Himel. Special thanks also to Mary Gibson and the staff for their helpful support and assistance.

Finally, thanks to my NASA sponsor, Dr. Richard Heydorn for making my SFF experiences both stimulating and productive.