BAYESIAN TECHNIQUES TO REDUCE THE SAMPLE SIZE IN AUTOMOTIVE ELECTRONICS ATTRIBUTE TESTING

by

Andre Kleyner
Delco Electronics Corp.

Shrikar Bhagath
Delco Electronics Corp.

Mauro Gasparini
Department of Statistics
Purdue University

Jeffrey Robinson
Mathematics Department
General Motors Res. Laboratories

and

Mark Bender
Delco Electronics Corp.

Technical Report # 96-26

Department of Statistics
Purdue University
West Lafayette, IN USA

June 1996

*

Bayesian Techniques to Reduce the Sample Size in Automotive Electronics Attribute Testing

Andre Kleyner[1], Shrikar Bhagath[1], Mauro Gasparini[2],
Jeffrey Robinson[3], and Mark Bender[1]

[1] Delco Electronics Corporation, Kokomo, IN 46904-9005

[2] Department of Statistics, Purdue University, West Lafayette, IN 47907-1399

[3] Mathematics Department, General Motors Research Laboratories,
Warren, MI 48090-9055

Key Words:- Bayesian reliability, Attribute test, Reliability demonstration, Sample size reduction, Beta priors, Mixture priors.

SUMMARY

The paper discusses the application of Bayesian techniques to the determination of sample sizes required for an attribute test of a product in order to demonstrate a target reliability with a specified confidence. The method is based on analyzing statistical data on similar products and incorporating them into a Bayesian prior distribution for the unknown reliability. A mixture prior obtained by combining a beta prior with the uniform rectangular prior (representing the unknown content of the new product design) is discussed. The suggested method can significantly lower sample sizes for attribute tests and thus reduce cost, time, and resources currently being spent on reliability demonstration testing. A numerical example at the end of the paper illustrates the method.

INTRODUCTION

In the pursuit of high quality and high reliability in a mass production environment, the automotive manufacturers require their suppliers to prove a target reliability with an assigned confidence level on a supplied product. This is usually done through a reliability demonstration test by running a certain number of samples under conditions simulating the mission life, an experiment which is sometimes called test to a bogey. Most of the time the sample size is determined only by the required reliability and the confidence level. Most of the methods currently used in the industry presume no prior information about the product or its predecessors, though very often this information is available. With the ever increasing reliability requirements the sample size to be tested is growing out of proportion and out of economical sense, requiring larger and larger amounts of human resources and capital equipment. Based on the fact that many new automotive products are evolutionary and not revolutionary, Bayes method can be one of the approaches to incorporate prior knowledge about the product, thus reducing the number of test samples and the amount of resources dedicated to the test programs.

EXISTING TECHNIQUES FOR SAMPLE SIZE DETERMINATION.

Statistical experiments are generally performed to learn more about unknown parameters characterizing our material of interest. In an automotive setup, the unknown parameter is the product reliability R, that is, the probability of surviving a specified mission life under standard condition: an attribute reliability experiment is performed to learn more about it. The experiment consists of observing N successes out of N reliability test trials. A peculiar feature is that most often no less than a 100% success rate is required - failing which corrective actions are to be taken- whereas in the usual reliability trials the success rate, albeit usually high, is random.

Techniques commonly utilized to calculate sample sizes for reliability demonstration of a product when a 100% success rate is required are generally referred to as Success Run Formulae [1,2]. The likelihood function, that is the probability of observing all successes given a certain value of the unknown product reliability R, is

$$L(data|R) = R^N \tag{1}$$

Based on this equation, the expression

$$C = 1 - {R_L}^N \tag{2}$$

is known in reliability circles as the "Success Run Formula" (see [2]) and can be obtained from the classical Clopper and Pearson [3] approach to the determination of confidence intervals for a binomial proportion. In equation (2), R_L is the lower bound of a one-sided $C\times 100\%$ confidence interval for the unknown reliability R. R_L is referred to from now on as the *demonstrated reliability*. In the automotive industry C and R_L are usually stipulated by the customer; the Success Run formula is then used for the determination of the required sample size N.

In a Bayesian approach instead, we use prior distributions on the unknown parameters of a statistical experiment to exploit useful pre-experimental information, for example the data from previous test results or similar product usage. For Success Run experiments, the likelihood (1) has to be combined with the prior distribution on R to obtain a posterior distribution on R. Such a posterior distribution summarizes all available information about the unknown product reliability R.

The Bayesian version of the classical Success Run Formula uses a *Uniform Prior,* also called a Rectangular Prior, which presumes an equal likelihood for the reliability value to fall anywhere between 0 and 1 and expresses the idea of "vague" prior information. In other words, since this prior assigns the same weight to every value of R, we expect it to produce results similar to the classical Success Run formula.

Combining the uniform prior and the likelihood using Bayes theorem we obtain the Bayesian version of the Success Run formula from the posterior probability

$$C = p(R_L < R < 1) = \frac{\int_{R_L}^{1} R^N dR}{\int_{0}^{1} R^N dR} = 1 - R_L^{N+1} \tag{3}$$

where C is the coefficient of the *credible* interval $[R_L, 1]$. R_L, the *(1-C)* quantile of the posterior distribution of R, is still to be referred to as the (Bayesian) *demonstrated reliability.* An interpretation of (3) is that, after the successful completion of a Success

Run experiment with N units, the unknown reliability R lies in the interval $[R_L,1]$, with $C\times 100\%$ probability.

The sample size calculated using equation (2) is one sample more than what we would get using equation (3). Some common reliability demonstration requirements and the sample sizes for Success Run of these demonstrations are given in Table 1. Equation (3) has been used in the calculation of these sample sizes.

Table 1 : Some Common Reliability Demonstration Requirements

Reliability to be Demonstrated	Confidence Level	Sample Size (Success Run formula)
0.95	0.9	45
0.97	0.7	40
0.99	0.5	69
0.99	0.9	229

FROM BETA PRIORS TO MIXTURES OF BETA PRIORS FOR PRODUCT RELIABILITY.

A generalization of the Success Run formula (3) can be obtained from priors other than the uniform. In Bayesian statistics, it is well known that for a binomial likelihood such as (1), a beta prior distribution on R, with density

$$\pi^*(R) = \frac{R^{A-1}(1-R)^{B-1}}{\beta(A,B)} \qquad if \quad 0 \le R \le 1 \tag{4}$$

Where $\beta(A,B)=\dfrac{\Gamma(A)\Gamma(B)}{\Gamma(A+B)}$

is particularly convenient; the constants A and B (sometimes called hyperparameters) have a nice interpretation - A being thought, sometimes, as the number of successes out of $A+B$ trials in a similar pre-experiment, real or imaginary. More importantly, the beta prior distribution is conjugate to binomial sampling, that is, the posterior is a beta distribution as well. This allows for a continuous updating of the posterior within the same general class of distributions. The uniform prior is a special case of (4) for $A=B=1$.

The posterior density on R obtained by combining (1) and (4) through Bayes theorem is

$$\pi^*(R|data) \quad = \quad \frac{\left(\dfrac{R^N R^{A-1}(1-R)^{B-1}}{\beta(A,B)}\right)}{\left(\dfrac{\int S^N S^{A-1}(1-S)^{B-1}dS}{\beta(A,B)}\right)} \quad = \quad \frac{R^{A+N-1}(1-R)^{B-1}}{\beta(A+N,B)} \tag{5}$$

that is, a beta density with parameters $(A+N)$ and B. The use of beta priors for binomial sampling has a long history, starting somewhere in the prehistory of modern Bayesian statistics. For an account of the uses of beta distributions in attribute reliability trials, see for example Martz and Waller [4,5]. As in the case of the standard Success Run formula (3), the immediate use of posterior (5) is to establish a reliability level R_L above which there is a high Bayesian credibility C that the reliability R will be met. For this purpose, we use equation

$$C = \int_{R_L}^{1} \frac{R^{A+N-1}(1-R)^{B-1}}{\beta(A+N,B)} dR \tag{6}$$

which tells us that there is a C posterior probability that R will be greater than R_L.

If, before the experiment, we require a certain Bayesian credibility C based on the contractual specifications, for given A and B the only unknown in expression (6) is the sample size N. For a given prior (4) we have to solve (6) numerically for N, in order to know how large a sample size we have to observe, with 100% success rate, to satisfy the required C and R_L.

The choice of the parameters of the prior A and B is a crucial one. It seems reasonable, in automotive reliability applications, to base such a choice on failure data, which are easily available and contain a lot of relevant information on past models or similar products. In the presence of information on the success rate of n previous life tests, a possible way to obtain A and B is based on an empirical Bayes approach discussed in [6]. See, for example, Martz and Waller [4] where empirical Bayes estimates of A and B are derived (see Appendix A).

Beta priors of the form (4) have a long history and are mathematically convenient, but for our purposes they are too restrictive. The best way to understand this is observing that an industrial product is in continuous evolution and, although a lot of similarity exists between old and new models, we always have a margin of novelty which should be accounted for. On the other hand, we do want to use prior information on similar products in our research; this is the reason why we want to use Bayesian methods in the first place. The right compromise between these conflicting goals seems to be generalizing the class of beta priors to the larger class of finite mixtures of beta priors. The plan is then to put together a prior distribution derived from failure data, and a margin of uncertainty intrinsic to the new model. The latter margin of uncertainty can be expressed as a uniform prior on the reliability.

Our proposal is therefore the use of a two-component mixture of beta distributions, with density

$$\pi(R) = \rho \frac{R^{A-1} \times (1-R)^{B-1}}{\beta(A,B)} + (1-\rho) \qquad if \quad 0 \le R \le 1 \tag{7}$$

The first component of the mixture is a beta prior with parameters A and B to be derived from failure data. The second component of the mixture is a uniform prior (a special case of the beta) representing uncertainty about the new product reliability. The two components are combined according to weights ρ and (1-ρ), where ρ is a "*knowledge factor*" representing how similar the new product is to the old one, and (1-ρ) is an "*innovation factor*", reflecting the proportion of new content in the new product. Notice that the use of a uniform prior alone would lead to the Bayesian version of the Success Run formula; the use of mixtures represents therefore a reasonable compromise between Bayesian and classical methods.

The idea of using mixture priors in the context of product reliability could be generalized to the case of heterogeneous prior information, in particular to the case where failure data is available for different past products, some more similar than others to the new product. In that case, the analysis could be generalized to the consideration of prior densities of the form

$$\pi(R) = \sum_i \left(\rho_i \frac{R^{A_i - 1} \times (1-R)^{B_i - 1}}{\beta(A_i, B_i)} \right) + (1-\rho) \tag{8}$$

where $\rho = \sum_i \rho_i$

and the different knowledge coefficients ρ_ι reflect different degrees of similarity between the new and the old products. Another reference to the use of mixture priors in Bayesian reliability is the article by Savchuk and Martz [7].

Only mixtures with two component are considered here. Combining (1) and (7) using Bayes theorem we obtain the posterior density,

$$\pi(R|data) = \frac{(1-\rho)R^N + \rho \dfrac{R^{A+N-1} \times (1-R)^{B-1}}{\beta(A,B)}}{\dfrac{(1-\rho)}{N+1} + \rho \dfrac{\beta(A+N,B)}{\beta(A,B)}} \tag{9}$$

and the corresponding expression

$$C = \int_{R_L}^{1} \pi(R|data)dR \tag{10}$$

where a required demonstrated reliability R_L and credibility C can be achieved. The solution of equation (10) has to be found, in general, by numerical methods.

AN EXAMPLE TO DEMONSTRATE APPLICATION OF THE TECHNIQUE

The sample size determination technique described in previous sections of this paper has been applied to a real life example to demonstrate a significant reduction in sample size. Table 2 shows failure data for an electronic vehicle control product (slightly modified from actual data for security reasons) in terms of IPTV (Incidents Per Thousand Vehicles). Table 2 shows breakdown by model years and body styles, totally constituting 12 test sets ($n = 12$). The observed failure rates R_j, are calculated from the IPTV data using:

$$R_j = 1 - \frac{IPTV}{1000} \tag{11}$$

Using equations from Appendix A, the values of A and B for the data in Table 2 are found to be 769.34 and 2.53 respectively. The cumulative distribution functions (CDFs) of the uniform, beta, and mixture distributions are shown in Figure 1 for the crucial range of $0.98 \leq R \leq 1$

Using equation (10) and solving numerically for the sample size, N, for a demonstrated reliability of $R_L = 0.99$ with $C = 90\%$, the sample sizes for various knowledge factors are as shown in Table 3. Using the classical Success Run formula (no prior knowledge about the product or knowledge factor $\rho = 0$), 229 samples of the product A will have to be tested with no failures to demonstrate a 0.99 reliability with 90% confidence. From Table 3 it is seen that with only a 10 % prior knowledge of the product (knowledge factor $\rho = 0.1$), the sample size reduces to 54 and as the knowledge factor increases, the sample size decreases.

CONCLUSION:

The method presented in this paper has great potential for cost reduction in reliability demonstration testing in a mass production environment like an automotive electronics industry. The failure data on similar products used to build a "prior" can significantly decrease the number of test items to a bogey. Even in cases with a low knowledge factor such as 0.2 or 0.3 (20-30% prior knowledge about the product), the method may present significant sample size reductions.

In cases with a favorable prior, the sample size may sometimes go down to zero or even become negative. The zero or negative sample sizes would mean that the required reliability has already been demonstrated during the previous stages of product development and no further testing is needed.

In instances with an unfavorable prior the number of samples to be tested may actually exceed the number computed using the classical method. This means that the product's prior has already shown that the product's reliability is most likely less than the desired outcome and no further testing should be performed without appropriate design corrections.

ACKNOWLEDGMENTS

We would like to gratefully acknowledge the help and support we received from Joe Boyle, Thomas Torri, Jay Rosen, and Ted DeGarmo at Delco Electronics; and Joe Wolkan at the General Motors Proving Grounds.

Table 2 : Calculation of Coefficients A and B from IPTV / Reliability Data

	Product	Model	Body Style	IPTV	Volume Sold l_j	Reliability R_j
1	Product A	19XX	Type I	2.94	170	0.9971
2			Type II	3.57	121	0.9964
3			Type III	2.45	206	0.9976
4			Type IV	5.32	35	0.9947
5			Type V	2.38	52	0.9976
6			Type VI	8.68	38	0.9913
7		19YY	Type I	1.75	306	0.9983
8			Type II	1.12	113	0.9989
9			Type III	4.06	87	0.9959
10			Type IV	1.61	27	0.9984
11			Type V	1.12	173	0.9989
12			Type VI	4.41	156	0.9956
n = 12					**$K = \Sigma l_j^{-1} = 1.66E\text{-}04$**	**$\Sigma R_j = 11.96$**

A = 769. 34

B = 2. 53

Table 3. Sample sizes for various knowledge factors at R = 0.99 and C = 90%

Knowledge Factor (ρ)	Sample Size N
1	0
0.9	1
0.8	2
0.7	4
0.6	6
0.5	9
0.4	13
0.3	19
0.2	30
0.1	54
0	229

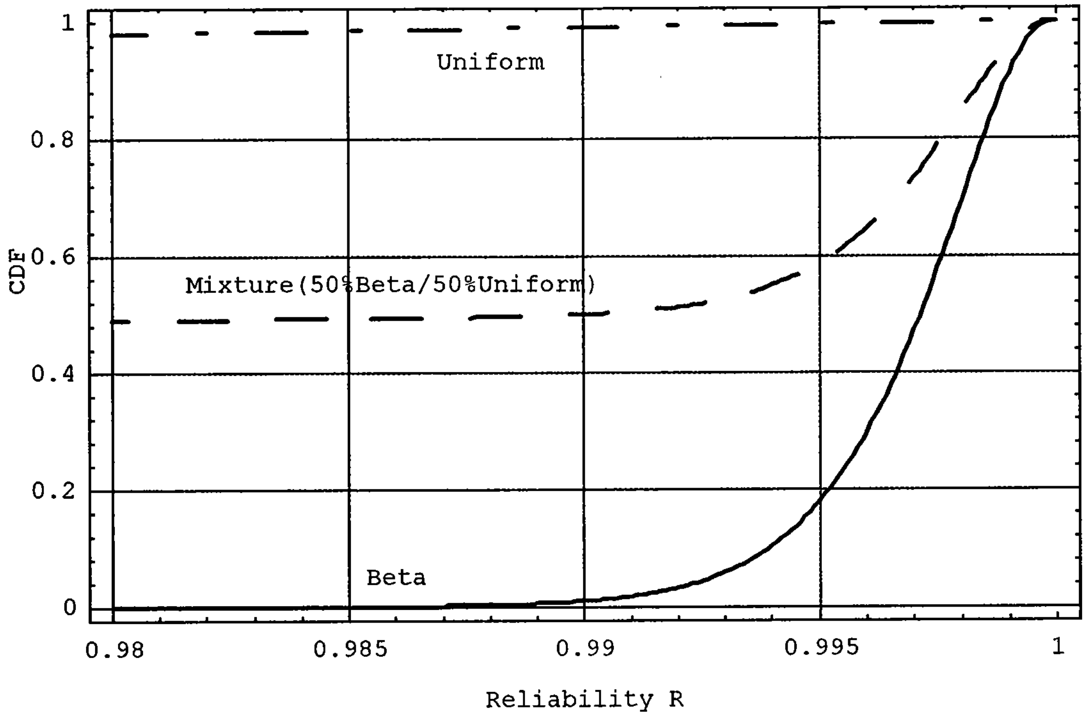

Figure 1. CDFs for Beta, Mixture and Uniform Distributions with A = 769.34, B = 2.53

REFERENCES

1. Johnson, L. G., GMR Rel. Manual, GMR-302, General Motors Research Laboratories, Aug. 1960.

2. Benedict. A. G., "Reliability-Confidence Combination for Small Sample Tests of Aerospace Ordnance Items.", NASA Technical Report 32-1165, JPL, California Institute of Technology, Pasadena, Calif, 1967.

3. Clopper, C.J. and Pearson, E.S., "The Use of Confidence or Fiducial Limits Illustrated in Case of the Binomial," Biometrika, 26, pp404-413, 1934.

4. Martz, H.F. and Waller, R.A., "The Basics of Bayesian Reliability Estimation from Attribute Test Data". Los Alamos Scientific Laboratory, Report UC-79p (February 1976).

5. Martz, H. F. and Waller, R. A. Bayesian Reliability Analysis. John Wiley & Sons, 1982.

6. Copas J. B., "Empirical Bayes Methods and Repeated Use of a Standard", Biometrika 59 p.p.349-360,1972.

7. Savchuk, V. and Martz, H., "Bayes Reliability Estimation Using Multiple Sources of Prior Information: Binomial Sampling.", IEEE Transactions on Reliability, Vol. 43, No. 1, March,1994.

APPENDIX A

Martz and Waller [4] derive empirical Bayes estimates of A and B as follows:

$$A+B=\frac{n^2(\sum_{j=1}^{n}R_j-\sum_{j=1}^{n}R_j^2)}{n(n\sum_{j=1}^{n}R_j^2-K\sum_{j=1}^{n}R_j)-(n-K)(\sum_{j=1}^{n}R_j)^2} \tag{i}$$

and

$$A=(A+B)\overline{R}$$

Where n is a number of life tests

l_j is a number of units in j^{th} test.

R_j is the j^{th} observed failure rate $= \frac{Number\ of\ failures\ in\ the\ j^{th}\ test}{l_j}$

$$K=\sum_{j=1}^{n}l_j^{-1}$$

$$\overline{R}=\frac{\sum_{j=1}^{n}R_j}{n}$$

When n is small, sampling error may cause equation (i) to yield negative estimates. If this occurs, Martz and Waller [4] suggest using the following form of this equation

$$A + B = \left(\frac{n-1}{n}\right)\left(\frac{n\sum R_j - (\sum R_j)^2}{n\sum R_j^2 - (\sum R_j)^2}\right) - 1 \qquad \text{(ii)}$$

These equations can be applied to processing real life data, where n would be the number of test sets for the similar products and R_j would be the reliability data from each set.