SSC-N-518 , 6/88

STABLE DENSITY FORM FOR BAYESIAN ANALYSIS OF ACCELERATED LIFE TESTS

J. ELY SHRAUNER

SSC/CDG and WASHINGTON UNIVERSITY

ABSTRACT

We extend our program of Bayesian analysis of accelerated life tests with the exhibition of a stable form of the marginal probability density for the objective variable that evolves under repeated application of Bayes' theorem. We relate these general concepts within the context of analysis of accelerated life tests for determination of the probability density for the probable failure rate λ of individual representative test magnets. The concept of a stable form of the probability density for λ containing all the information consistently relevant for its determination supplements and connotes the previously demonstrated concept of a stable minimal set of statistics sufficient for its determination. The stable form for the case of the exponential failure function is the Poisson distribution, and its behavior under the accumulation of augmentary information is described quantitatively.

I INTRODUCTION

We extend our program of Bayesian analysis with the exhibition of a stable form of the marginal probability density for the objective variable that evolves under repeated application of Bayes' theorem. We relate these general concepts within the context of analysis of accelerated life tests for determination of the probability density for the probable failure rate λ of individual representative magnets. The concept of a stable probability density for λ, which contains all the probabilistic information that is consistently relevant for λ supplements and connotes the concept of a stable minimal set of statistics sufficient for the determination of the λ density. The latter was demonstrated previously.[1.]

Identification of a stable minimal set of statistics sufficient for determination of the probability density for λ makes possible the most efficient extraction of relevant results from (ALT) experiments and dictates the most efficient parameterization of the probability densities. Both the stable minimal set of sufficient statistics and the stable form of the probability densities *reflect the particular model employed for expressing the fundamental failure process.* In our case, as in our previous treatments, demonstrations are simplest with the exponential failure-function model with constant failure rate λ. The concepts are much more general; for example, extensions to two-parameter Weibull-type models are immediate.

The formulation that focuses on the probability densities which was adopted for the demonstration of ref. 1. is also convenient for our present purposes in preference to the formulation that focuses on the probabilities themselves which was used in our earlier treatments.[2,3.] Both formulations reflect the failure-function model employed and both will accommodate the maximum-information-entropy refinements exhibited in ref. 3.

The *density-emphasizing formulation introduces the readily obtained and explicitly model* dependent likelihood function L, which, although not universally a strict conditional probability density for the objective variable λ, incorporates the maximal probabilistic information relevant to the determination of the density for λ on evidence of the experimental outcome used to augment the prior probabilistic information on λ. The details of the augmentation, and therefore usually of the density for λ posterior to Bayesian augmentation are revealed by the likelihood function and the maximum likelihood estimator, MLE. The MLE is the most likely value of where the likelihood function L has its maximum, $\partial L/\partial\lambda = 0$. The likelihood function L and/or MLE may not exist for some complexly parameterized models, or ill-posed sampling protocols, or some outcomes of otherwise satisfactory models and experimental protocols. Furthermore, there are approximations endemic to this formulation; but they in no way detract from its convenience for demonstrations, and quite generally become more negligible as ALT sample size and/or total time on test become large. In some of our applications with protocols involving small samples these approximations begin to compromise the numerical accuracy of this formulation. Never-the-less, its convenience for our demonstration still recommends it.[4.]

Our demonstration of the evolution of a stable form of the probability density for the objective variable λ starts from an assumed microcanonical form of the prior probability density for and augments it with information obtained from experiment by employment of Bayes' theorem. The resulting posterior density for λ is of the form of a Poisson distribution parameterized with the components of the minimal set of statistics sufficient for determination of the λ-density with *this model as obtained from the experiment. This minimal set of statistics sufficient for determi*nation of the probability density with this model was previously demonstrated by the same method to be of stable form.[5] Further augmentation by means of Bayes' theorem to incorporate further experimental (or other) information using the previously-determined Poisson distribution as the new prior density results in a posterior density for λ that is again of the form of a Poisson distribution; again parameterized with the minimal set of statistics sufficient for λ with this model. The form of the probability density for the objective variable λ has evolved from the initial microcanonical form to the stable form for this model, the Poisson distribution. The form of the minimal set of statistics sufficient for determination of the probability density for λ in this model is also stable. When the prior probability density is of the form of the Poisson distribution augmentation with Bayes' theorem results in a posterior density that is also of the stable form of the Poisson distribution and it is parameterized by the stable minimal set of statistics sufficient for determination of the λ-density in this model comprised of components which are each the cummulative statistic in which the statistic parameterizing the prior density has been directly augmented with the corresponding statistic from the new experiment. Further Bayesian augmentation in this formulation, once it has been expressed in terms of these stable forms, merely increments the necessary parameters.[6]

II DEMONSTRATION

Likelihood Function In Terms of Sufficient Statistic

In the context of present interest we consider the outcome of an ALT experiment in which n representative magnets fail at times y_j during a test of a total sample of N magnets spanning a time y_0 . We denote this experimental outcome by $(N, y_1, \ldots y_n, y_0)$ and the probability of this outcome predicated on the failure rate λ by $P(N, y_1, \ldots y_n, y_0 \mid \lambda)$. The likelihood function for the objective variable λ is then

$$(1) \quad L(\lambda \mid N, y_1, \ldots y_n, y_0) = P(N, y_1, \ldots y_n, y_0 \mid \lambda) \; .$$

Since a sufficient statistic for determination of the probability density for λ within the context of the present exponential failure function model has been shown to be comprised of the total time on test T and the number of failures n , the individual failure times y_j are superfluous informational details for this limited purpose. We can thus write the likelihood function for on evidence of the experimental outcome $(N, y_1, \ldots y_n, y_0)$ as

$$(2) \quad L(\lambda \mid N, T, n) = L(\lambda \mid N, y_1, \ldots y_n, y_0) \; ,$$

where T is understood to be the total time on test for this outcome,

$$(3) \quad T(N, y_1, \ldots y_n, y_0)$$
$$= N\, y_1 + (N-1)(y_2 - y_1) + (N-2)(y_3 - y_2) + \ldots (N-n)(y_0 - y_n)$$
$$= \sum_{j=1}^{n} y_j + N\, y_0 \quad .$$

We have argued from our previously established general concepts to arrive at this functional form. We shall see that it is the correct result of a simple brute force approach.

Explicit Model Dependence

With the exponential failure function model the probability of the failure of a single representative magnet in the time interval $[0,y]$ conditional on its failure rate being λ is given by the failure function

$$(4) \quad F(y \mid \lambda) = 1 - e^{-\lambda y} \quad .$$

The corresponding marginal probability density for failure is

$$(5) \quad f(y \mid \lambda) = \lambda\, e^{-\lambda y} \quad .$$

With this model the joint probability density of the ALT outcome $(N, y_1, \ldots y_n, y_0)$ is

$$(6) \quad P(N, y_1, \ldots y_n, y_0 \mid \lambda) = \prod_{j=1}^{n} [\lambda e^{-\lambda y_j}][e^{-\lambda y_0}]^{N-n} \; C(N,n)$$
$$= \lambda^n e^{-\lambda T(N, y_1, \cdots y_n, y_0)} C(N,n) \; ,$$

where $C(N,n)$ is a λ -independent, protocol-dependent combinatoric normalization factor that will proove to be totally innocuous in our Bayesian analysis. Identifying this expression with the likelihood function as in Eq.(1) confirms our assertion that L is indeed of the form predicted on general principal; as shown in Eqs.(2,3). It is dependent only on T and n, and not on the individual y_j 's . The explicit model dependence is manifest here as well as the dependence on the stable minimal set of statistics sufficient for determining the probability for λ in this model. This is further apparent when the above expression is incorporated as the likelihood function in Bayes' theorem.

Bayes' Theorem

The essence of our analysis is Bayes' theorem in the form

$$(7) \qquad P(\lambda|c,N,y_1,...y_n,y_o) = \frac{P(N,y_1,...y_n,y_o|\lambda)P(\lambda|c)}{\int d\lambda \quad '' \qquad ''} \quad ,$$

in which

(8) $P(\lambda|c)$ = the prior marginal probability density for λ conditional on all prior relevant information represented by c , which might be from prior testing or from other, perhaps subjective, information or lack of information ;

and

$$(9) \qquad P(\lambda|c,N,y_1,...y_n,y_o) = P(\lambda|c,N,T,n)$$

= the Bayesian posterior probability density for λ containing all knowable probabilistic information about λ on evidence of all prior relevant information contained in $P(\lambda|c)$ augmented through Bayes' theorem with the relevant joint statistic $(n,T(N,y_1,...y_n,y_o))$ obtained from the outcome $(N,y_1,...y_n,y_o)$ of the ALT experiment.

All λ-independent factors cancel out of Bayes' theorem in the form of Eq.(7) , which itself is of the form of a normalization condition; confirming the innocuous nature of these factors mentioned above.

Maximum Likelihood Estimator (MLE) For The Objective Variable

The maximum likelihood estimator for the objective variable λ for this experimental outcome with this model is given by

$$(10) \qquad 0 = \frac{1}{L}\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - T \quad ,$$

or

$$(11) \qquad \text{MLE} = n/T = n/T(N,y_1,...y_n,y_o) \; .$$

Microcanonical Prior Distribution For λ —Gives Poisson Posterior Distribution

As described for our earlier analyses, if our prior knowledge of the expected true value of λ is very vague we might choose to start with the microcanonical prior probability density :

$$(12) \quad P(\lambda|c) = 1/(\lambda_0 - \lambda_c) \ , \quad \lambda_c \leqslant \lambda \leqslant \lambda_0,$$
$$= 0 \qquad , \text{ otherwise.}$$

The posterior density, which includes information from the prior $P(\lambda|c)$ augmented by the experimental information in $L(N,y_1, ...y_n, y_0)$, is computed with Bayes' theorem as:

$$(13) \quad P(\lambda|c,N,y_1, ...y_n, y_0) = L(\lambda|N,y_1, ...y_n, y_0)/(\lambda_0-\lambda_c)$$

The *normalization factor* $(\lambda_0-\lambda_c)$ *cancels out of this expression of* Bayes' *theorem*, so the microcanonical interval could become very large. If this interval becomes large so that $\lambda_c << n/T << \lambda_0$ then the denominator becomes after the cancellation

$$(14) \quad \int_0^\infty \lambda^n e^{-\lambda T} d\lambda = T^{-n-1} \ n! \quad .$$

This condition usually connotes large sample size N . All relevant information about λ obtained from prior experiments and contained in the prior probability density for λ , $P(\lambda|c)$, plus all information additional on evidence of the ALT outcome $(N,y_1, ...y_n, y_0)$ is contained in the posterior density $P(\lambda|c,N,y_1, ...y_n, y_0)$. The likelihood function L is of course specific to the failure function model as well as the experimental outcome, but because of its sharp maximum under condition $\lambda_c << n/T << \lambda_0$, the posterior density is much more sharply concentrated than the microcanonical prior density, and is completely specified in this case by the joint statistic $(n,T(N,y_1, ...y_n, y_0))$ of the likelihood function, since no such information was contributed from the prior density in this case. This is the result hoped for in any contemplated experiment, i.e. information on the subjective variable has been considerably increased. We will elaborate further on this later.

If we compute the maximizer of the posterior density $P(\lambda|c,N,y_1, ...y_n, y_0)$ we obtain with

$$(15) \quad 0 = \frac{1}{P}\frac{\partial P}{\partial \lambda} = \frac{n}{\lambda} - T \quad ,$$

the same as the MLE for the likelihood function alone with no contribution from this "white-noise", λ -independent microcanonical prior.

The maximum information entropy variationally minimizes the unsubstantiated information contained in a probability distribution subject to the data relavent to its determination expressed in the form of expectations with respect to this probability distribution. The maximum likelihood estimator is a wholely different thing. It does not contribute to the variational determination of a probability distribution. Rather, in our usage the MLE merely tells where the experimental outcome, through its likelihood function, is trying to shift the concentration of probability in the augmented posterior distribution. Indeed, in a given problem the prior may be determined by the *maximum information entropy variational method* and also the maximum likelihood estimator be useful in assessing the effect of the augmentation. In fact, the augmented posterior density in our present treatment behaves very similarly to the maximum entropy prior of the previous treatment; giving almost the same empirical rules for interpolation.

Poisson Prior Distribution For λ ---Gives Poisson Posterior Distribution, Establishing Stability Of This Form

It follows from the above analysis that as soon as anything is known relevant to the objective variable of interest that information can be included in its probability density in the stable form of sufficient statistics for this variable (in the particular analysing model) parameterizing the stable form of its probability density. In our illustrative case employing the exponential failure function model for analysing ALT results for the failure rate λ the stable form for its probability density is the Poisson distribution. We have shown that this form evolves immediately from the microcanonical density representing minimal prior information relevant to our objective variable λ by a single Bayesian augmentation with experimental information. The posterior probability density resulting from augmenting the minimal information in the microcanonical prior with new experimental information in the stable form of the sufficient statistics for λ in this model (n_c ,T_c) was shown to be the Poisson distribution parameterized by these statistics. It follows immediately that if this density is then taken as the prior density for the next Bayesian augmentation with new experimental information in the form of the statistics (n , T) then the density for λ posterior to this augmentation is of the stable form of the Poisson distribution parameterized by the combined statistics ($n + n_c$,$T + T_c$).

To see this explicitly consider the implementation of this augmentation by Bayes' theorem. The prior density for λ is the normalized Poisson distribution

$$(16) \quad P(\lambda|c,N,T_c,n_c) = P(\lambda|N,T_c,n_c) = T_c^{n_c+1}\lambda^{n_c}e^{-\lambda T_c}/\,n_c! \quad ,$$

and the likelihood function for λ from the new experimental information is

$$(17) \quad L(\lambda|N,T,n) = \lambda^n\, e^{-\lambda T}\, C(N,n) \quad .$$

The Bayesian augmentation gives for the updated posterior density for λ is

$$(18) \quad P(\lambda|N,T_c,n_c;N,T,n) = \frac{L(\lambda|N,T,n)P(\lambda|N,T_c,n_c)\,C(N,n)}{\int_0^\infty d\lambda \quad '' \qquad '' \qquad ''} \,.$$

The normalizing denominator integral (after cancelling the C 's) is

$$(19) \quad (T_c^{n_c+1}/\,n_c!)\int_0^\infty d\lambda\, \lambda^{n_c}e^{-\lambda T_c}\lambda^n e^{-\lambda T}$$

$$= (T_c^{n_c+1}/\,n_c!)[(n+n_c)!\,/(T+T_c)] \quad ,$$

so the posterior density (18) is just

$$(20) \quad P(\lambda|N,T_c,n_c;N,T,n) = P(\lambda|N,T_c+T,n_c+n) = (T+T_c)^{n+n_c+1}\lambda^{n+n_c}\;e^{-\lambda(T+T_c)}/\,(n+n_c)! \quad ,$$

which is the normalized Poisson distribution with the augmented statistics $[(n + n_c), (T + T_c)]$ in place of the prior statistics (n_c, T_c). This demonstrates the stability of the Poisson density form for this model, as described above. The normalized density for λ for this model at any stage of testing is of this form parameterized with n = the sum of all testing failures , and T = the sum of all time on test.

III NARROWING OF DISTRIBUTION WITH INCREASING INFORMATION

A Measure Of The Informational Narrowing

The probability distribution for λ is very broad and flat when very little information relevant to it is known, as described above with the microcanonical distribution. Conversely, when a large amount of information relevant to its probabilistic description is known its probability density reflects this with a concentrated narrow peak. In the limit of exact knowledge the density would approach a delta function. It is useful for some purposes to be able to quantify this reflection in the probability density of the amount of information it contains relevant to the objective variable—always within the context of the model chosen for its analysis. A useful figure of merit for this purpose is the ratio of the standard deviation of the distribution to its mean, sometimes called the coefficient of variation of the distribution.

The Mean Of The Poisson Distribution

The mean of our Poisson distribution is

$$(21) \quad \bar{\lambda} = \frac{\int_0^\infty e^{-\lambda T} \lambda^n (T^{n+1}/n!)\, \lambda\, d\lambda}{\int_0^\infty e^{-\lambda T} \lambda^n (T^{n+1}/n!)\, d\lambda} = (n+1)/T \ ,$$

which is a close approximation to $1/\bar{y}$, in which $\bar{y}$ is the mean value of the n times to failure y_j in a test to some fixed time y_0 of our sample. We again call attention to the fact that only n and T are involved here, and not the more detailed information contained the y_j 's nor the sample size N.

The Standard Deviation Of The Poisson Distribution

The standard deviation about the mean of our Poisson distribution is the square root of the quantity

$$(22) \quad \sigma^2 = \frac{\int_0^\infty e^{-\lambda T} \lambda^n (T^{n+1}/n!)(\lambda - \bar{\lambda})^2 d\lambda}{\int_0^\infty e^{-\lambda T} \lambda^n (T^{n+1}/n!)\, d\lambda} = (n+1)/T^2 \ ,$$

so the standard deviation about the mean is

$$(23) \quad \sigma = (n+1)^{1/2}/T \ .$$

The Coefficient Of Variation Of The Poisson Distribution

The coefficient of variation of our Poisson distribution for is the ratio

$$(24) \quad \sigma/\bar{\lambda} = [(n+1)^{1/2}/T]/[(n+1)/T] = 1/(n+1)^{1/2} \ ,$$

which does not even depend on T; let alone N and the y_j 's . This narrowing at a rate that goes like one over the square of the total number of failures observed in all testing to date is not surprising--once it is shown. Again, this reflects the fact that the sufficient statistic for determining the density for λ is just (n,T) so the only statistical parameter to scale the λ -density is 1/T, and both the mean $\bar{\lambda}$ and the standard deviation σ of the distribution for λ must be of the form of the quantity 1/T times a function of n only. The normalization condition on the density implies that if its width is say $2(2/3)\sigma$ then its effective height must be something like one over this quantity .

To meet the severe requirements of the availability specifications for the SSC magnets we must have the least possible number of failures over the largest affordable total time on test T, which in this case is about proportional to the test sample size N. The formulation described here applies to the case in which $n = 0$.

IV CONCLUSION

We have demonstrated the existence of a stable form of the probabililty density for the objective variable, the individual-magnet failure rate, of our Bayesian analysis of accelerated life tests. This concept, in supplement with that of the minimal stable set of statistics sufficient for the determination of the probability density for the objective variable, are very important indicators for heuristic estimations of the value of testing decisions that are useful or even valid for cases in which the analyzing model is more complex than those considered here. They will be extremely important guiding concepts in our magnet testing program.

REFERENCES AND FOOTNOTES

1. Sufficient Statistic for Abstracting The Probability Distribution For Expected Magnet Failure Rate From Accelerated Life Test, SSC-N-509, 4/88, E. Shrauner.

2. Availability Analysis of Accelerated Life Tests For SSC Magnets, SSC-N-254, 10/86; and Preliminary Assessment of Magnet Accelerated Life Tests, SSC-N-215, 8/86, E. Shrauner.

3. Maximum Information Entropy-Bayesian Analysis For Availability Studies Of Accelerated Life Tests Of SSC Magnets, SSC-N-310, 12/86, 2/87, E. Shrauner.

4. The advantages of the formulation that emphasizes the densities and the employment of the MLE were suggested to me by Prof. R.E. Barlow, who has described these and other concepts treated in this paper in : R.E. Barlow and F. Proschan, in Theory of Reliability, Soc. Ital. di Fisica, Bologna, corso XCIV, 1986.

5. The concept of the stable minimal set of statistics for the determination of the probability density of the objective variable from experimentation analyzed in terms of a given model, i.e., the concept of "sufficient statistic" was suggested to me by Prof. E.T. Jaynes, who averred that it is more than a century old.

6. The stable form of the prior density for the exponential failure function model has been discussed by Barlow and Proschan in ref. 4., where they refer to it as the "natural conjugate prior" for this model. Other usages refer to it as the "stabilizer density" for a given model. The stable density form that propagates unchanged from the prior through the Bayesian augmentation to the posterior density may not exist for some models; but this may be associated with too complex of models to be directly illustrative for purposes such as ours.

ERRATA:

SSC-N-518, 6/88
STABLE DENSITY FORM FOR BAYESIAN ANALYSIS OF ACCELERATED LIFE TESTS
J. ELY SHRAUNER

1) On page 3 the last line of Equation (3) should read:

$$= \sum_{j=1}^{n} y_j \quad + (N - n)\, y_o \quad .$$

2) On page 8 the third line from the bottom should read:

.......distribution for λ is the ratio

3) On page 9 the top line should read:

like one over the square root of the number.....

ERRATA:

SSC-N-518, 6/88
STABLE DENSITY FORM FOR BAYESIAN ANALYSIS OF ACCELERATED LIFE TESTS
J. ELY SHRAUNER

1) On page 3 the last line of Equation (3) should read:

$$= \sum_{j=1}^{n} y_j \quad + (N - n)\, y_o \quad .$$

2) On page 8 the third line from the bottom should read:

.......distribution for λ is the ratio

3) On page 9 the top line should read:

like one over the square root of the number.....