SSC-N-509

SUFFICIENT STATISTIC FOR ABSTRACTING THE PROBABILITY DISTRIBUTION FOR EXPECTED MAGNET FAILURE RATE FROM ACCELERATED LIFE TEST

Ely Shrauner
SSC Central Design Group* and Washington University

April 1988

*Operated by Universities Research Association for the U. S. Department of Energy

ABSTRACT

We extend our program of Bayseian analysis of inferences about reliability and availability of the SSC magnet system from the results of accelerated life tests (ALT). We show that under assumption of the exponential failure function hypothesis the joint statistic (n,T) comprising the number of failures n and the total time on test T is necessary and sufficient for determining the best estimate for the probability density for the expected magnet failure rate λ. There is therefore no finesse by which the ALT protocol may be manipulated to yield higher probabilistic confidence that the expected failure rate is less than that corresponding to the very severe reliability requirements specified for the SSC with smaller test samples and testing times than the large ones established in our previous analyses. We show this in general for the exponential model and with two explicit examples of common ALT sampling protocols.

INTRODUCTION

We have previously described the analysis to extract the probabilistic confidence in a failure rate corresponding to specified availability for the SSC magnet system that could be inferred from results of accelerated life tests (ALT) on samples of the magnets[1,2] This analysis was essentially Bayesian and (thereby) involved prior probabilities that had to be presumed. Two stages were illustrated of this Bayseian analysis.

First, the prior probability distribution for magnet failure rate λ was presumed to be of the microcanonical form: uniform in some finite domain of λ , $\lambda_c \leqslant \lambda \leqslant \lambda_0$, and zero outside of this domain[3] With this presumed prior distribution on the magnet failure rate λ we computed the probability that the true λ_T of our magnets was less than the value λ_S corresponding to specified availability requirements could be inferred from evidence that n magnets out of a sample of N fail to survive y equivalent operating years of ALT cycling[2] (One week of ALT cycling is about equivalent to one year of expected operational cycling according to the ALT protocol currently considered.)

Second, the probabilistic confidence that $\lambda_T \leqslant \lambda_S$ was computed with the canonical distribution that results from a maximum information-entropy analysis[4] In this more refined theoretical approach the probability distribution which variationally maximizes the information entropy (i.e., incorporates the minimum non-substantiated information subject to the constraints of the data in expectational form —i.e., incorporates all the information of the test data as expectations with respect to this distribution) is of canonical form. Computations with this mathematically unique and minimally subjective canonical distribution for the expected rate λ recapitulate the probabilistic results of our ALT analysis that were indicated with less theoretically substantial and accurate earlier treatments that presumed a microcanonical distribution. Empirical rules for interpolations derived with the microcanonical approximation hold also in the maximum-information-entropy, canonical case.

In the present paper we extend our Bayseian analysis of the magnet ALT. We use a slightly different and more expedient approach than our previous treatments. The difference stems from employment of the failure rate probability distribution densities, rather than the probability distributions, and is associated with a slightly different, but similar and simpler maximization process. The advantages of working with probability densities and introducing the concept of the likelihood function and its maximization were suggested to me by Prof. R. E. Barlow[5] This method has some limitations, which we acknowledge before employing it to illustrate the concept of sufficient statistic as it applies in our analysis and how as a consequence the maximum information derivable from our ALT analysis for the magnet failure rate is essentially independent of variations in our ALT protocol[6] There is no finesse by manipulation of the ALT protocol by which smaller test sample sizes and shorter testing periods than the large ones indicated in our previous analyses can yield higher probabilistic confidence that the expected failure rate is less than that corresponding to the SSC availability specifications.

The large ALT sample sizes and testing periods required to meet the demanding specifications for the SSC indicated in our previous analyses might make one wish to manipulate the ALT protocol to see if some faster route can be discovered. For example; is it more profitable to test a sample of N magnets a) for a predetermined time interval, or b) until a predetermined number of failures occur? We illustrate with these examples that there is no such device because the total time on test plus the number of failures constitute the necessary and sufficient statistic for determining the maximum likehood estimator for the failure rate λ, and its posterior density on evidence of the outcome of the ALT subject to the presumed failure function model.

Throughout we will be concerned with the analysis of the outcome of an ALT experiment which we denote by $(N, y_1, \ldots\ldots, y_n, y_0)$ in which a sample of N identical, representative magnets are independently tested for a period of y_0 equivalent years of operational cycling (one week of ALT testing is equivalent to about one year of expected operational cycling) with n magnets withdrawn due to failure at times $y_1, \ldots, y_n$.

DEMONSTRATION

Bayes' Theorem

The basis for our whole analysis is Bayes' theorem. We employ it here in the form:

$$P(\lambda|c,N,y_1,\ldots y_n,y_0) = \frac{P(N,y_1,\ldots y_n,y_0|\lambda)\,P_c(\lambda)}{\int d\lambda \quad \text{''} \quad \text{''}}$$

in which,

$P(N,y_1,\ldots y_n,y_0|\lambda)$ = the joint probability density of the ALT outcome $(N,y_1,\ldots y_n,y_0)$ conditional on failure rate λ ;

$P_c(\lambda) \equiv P(\lambda|c)$ = prior marginal probability density for λ conditional on all prior information represented by c, which might be from prior testing or from our subjective information or lack of information.

$P(\lambda|c,N,y_1,\ldots y_n,y_0)$ = Bayseian posterior probability density for λ on evidence of the outcome $(N,y_1,\ldots y_n,y_0)$ of the ALT experiment augmenting all the prior information c, whose relevance here is contained in $P(c|\lambda) = P_c(\lambda)$;

This analysis is complete and correct within the context of Bayes' theorem and axiomatic probability theory. The posterior density $P(\lambda|c,N,y_1,\ldots y_n,y_0)$ contains all knowable statistical information about λ on evidence of the prior information contained in $P_c(\lambda)$ augmented by the statistical information obtained from the ALT results $(N,y_1,\ldots y_n,y_0)$.

Likelihood Function for

The likelihood function for λ quantifies information on λ from evidence of the ALT results as the joint probability density of the ALT outcome conditional on λ ; i.e.,

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

Strictly speaking L is not a probability density in the variable λ.

Maximum Likelihood Estimator of λ

The maximum likelihood estimator (MLE) of λ is determined from

$$0 = \partial L/\partial\lambda \quad .$$

The MLE describes where L, and thus, through Bayes' theorem, where $P(\lambda|c,N,y_1,\ldots y_n,y_0)$ concentrate.

Model–Exponential Failure Function Hypothesis

In order to proceed we must assume a failure function model which allows us to transcribe the outcome of the ALT into the likelihood function for λ, $L(\lambda|N,y_1,\ldots y_n,y_0)$. We continue with our previously assumed exponential failure function hypothesis, for which the failure function is

$$F(y|\lambda) = 1 - e^{-\lambda y}.$$

F is, of course, the probability of failure in the time [0,y] of a single sample magnet conditional on failure rate λ. The corresponding failure density is

$$f(y|\lambda) = \lambda e^{-\lambda y} \quad .$$

With this model the likelihood function for λ is

$$\begin{aligned} L(\lambda|N,y_1,\ldots y_n,y_0) &= P(N,y_1,\ldots y_n,y_0|\lambda) \\ &= \prod_{j=1}^{n} [\lambda e^{-\lambda y_j}][e^{-\lambda y_0}]^{(N-n)} \\ &= \lambda^n e^{-\lambda T(N,y_1,\cdots y_n,y_0)} \end{aligned} \quad ,$$

In this expression the quantity

$$\begin{aligned} T(N,y_1,\ldots y_n,y_0) &= \sum_{j=1}^{n} y_j + (N-n)y_0 \\ &= Ny_1 + \sum_{j=1}^{n-1}(N-j)(y_{j+1} - y_j) + (N-n)(y_0 - y_n) \end{aligned}$$

is the total time on test: the sum of (N - j) magnets on test between the j-th and (j + 1)-th failures plus (N - n) magnets on test from y_n to y_0. Clearly the likelihood function is specific to the failure function model chosen to conform to prior information about the nature of the (magnet) system under consideration. But within this model L is independent of the prior probability density $P(\lambda|c)$ for λ ; except that it might be said that the information required to choose the failure function model is similar to that required to chose $P(\lambda|c)$. Within the context of this failure model, the likelihood function for λ , $L(\lambda|N,y_1,\ldots y_n,y_0)$ is necessarily and sufficiently parametrized completely by the joint statistic $(n,T(N,y_1,\ldots y_n,y_0))$.

The posterior density is completely specified from the prior density $P(\lambda|c)$ and the joint statistic obtained from the ALT experiment comprising the number of failures observed n and the total time on test $T(N,y_1,\ldots y_n,y_0)$. That is, with this failure function model the only observables necessary and sufficient for estimating the posterior density $P(\lambda|c,N,y_1,\ldots y_n,y_0)$ containing all knowable information about λ on evidence of the prior information contained in $P(\lambda|c)$ augmented by the information obtained in the ALT results $(N,y_1,\ldots y_n,y_0)$ are n and $T(N,y_1,\ldots y_n,y_0)$. The prior density $P(\lambda|c)$ is, of course, presumed for determining this posterior probability density with Bayes' theorem. But the number of failures and total time on test, or their equivalents, were by the same arguement necessary and sufficient statistics for determining the prior density $P(\lambda|c)$ from prior ALT or equivalent input. Thus, all prior densities for λ with this failure function model can be described with parameters equivalent to some n_c and T_c from some prior stage of information acquistion. Bayes' theorem updates information so that yesterday's posterior is today's prior.

The maximum likelihood estimator from the ALT statistics with this failure function model is obtained from

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

or

$$\lambda_{max} = n/T = n/T(N,y_1,\ldots y_n, y_0) \quad .$$

The MLE might not exist for some more complicated model's, but it is obviously of this form for failure models not greatly dissimilar from the exponential model.

Microcanonical Prior Distribution On λ

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 for λ :

$$P(\lambda|c) = P_c(\lambda) = 1/(\lambda_o - \lambda_c) \;, \quad \lambda_c \leqslant \lambda \leqslant \lambda_o,$$

$$= 0 \qquad\qquad , \text{ otherwise.}$$

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

$$P(\lambda|c,N,y_1,...y_n,y_o) = \frac{L(\lambda|N,y_1,...y_n,y_o)/(\lambda_o - \lambda_c)}{\int_{\lambda_c}^{\lambda_o} d\lambda \quad '' \quad ''} \quad .$$

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

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

This condition usually connotes large sample size N . All information about λ obtained from prior experiments and contained in the prior probability density on λ , $P(\lambda|c)$ plus all information additional on evidence of the ALT outcome $(N,y_1,...y_n,y_o)$ is contained in the posterior density $P(\lambda|c,N,y_1,...y_n,y_o)$. The likelihood function L is of course specific to the failure function model, but because of its sharp maximum under condition $\lambda_c << n/T << \lambda_o$, the posterior density is much more sharply concentrated than the microcanonical prior density, and is completely specified by the joint statistic $(n,T(N,y_1,...y_n,y_o))$.

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

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

the same MLE as for the case of this microcanonical prior unaugmented.

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 and prediction.

TESTING TIME TO CONFIRM SSC AVAILIBILITY FROM MAXIMUM LIKELIHOOD ESTIMATOR

How much testing time is required to establish the rate λ corresponding to the 96% availibility specified for the SSC magnet system? The general relation among the stationary availibility A_∞ and the failure rate λ of each magnet of a system of 9600 in serial fault configuration with mean time to repair τ is[2]:

$$.96 = A_\infty = \frac{1}{1+9600\lambda\tau} \Longleftrightarrow \lambda\tau = (1/.96 - 1) = 4\cdot 10^{-6} .$$

The value of $\tau \leqslant 5$ day $= .014$ year to replace a magnet ;
or $\tau \geqslant 1$ hour $= .00011$ year for a small quench.

So the limits on the failure rate λ in year^{-1} are :

$$4\ 10^{-6}/.014 = 3\cdot 10^{-4} \leqslant \lambda \leqslant 4\cdot 10^{-6}/.00011 = 3.6\cdot 10^{-2}\ \text{year}^{-1} .$$

The MLE on evidence of n failures in testing N magnets for y_o equivalent operating years with total time on test T when the prior information was equivalent to n_c , T_c is

$$\hat{\lambda} = (n + n_c)/(T + T_c) \Longleftrightarrow T = (N - n)y_o + \Sigma y_j = (n + n_c)/\hat{\lambda} - T_c$$

in which must be $\hat{\lambda} \geqslant \lambda_{min} \cong 2\cdot 10^{-4}$ year^{-1} from above. Thus, the test time required is

$$y_o > \left[\frac{(n + n_c)/\lambda_{min} - T_c - \Sigma y_i}{(N - n)}\right] = \frac{(n + n_c)/\lambda_{min} - T_c - n\bar{y}}{(N - n)} ,$$

in which $\bar{y} = \Sigma y_j / n$ is the average y_j .

In the case of the microcanonical prior the MLE is $\hat{\lambda} = n/T$, i.e., $n_c = 0 = T_c$ in the above expression for y_o . Thus, for example:

$N = 10, n = 2 :\ y_o > 1.1\ 10$ year ,
$N = 20, n = 3 :\ y_o > 8.1\ 10$ year ,
$N = 20, n = 2 :\ y_o > 5.3\ 10$ year , etc.

These are about the same as earlier estimates; but now we have shown that they cannot be manipulated and are quickly computed.

TWO ALT PROTOCOLS FOR EXAMPLE

Protocol I: Test For Prescribed Time Interval

In this case our ALT protocol involves testing a sample of N magnets for a prescribed time y_o in which n failures are observed at $y_1,\dots y_n$. We have seen that $y_1,\dots y_n$ are unnecessary for inferring information about λ from ALT; only the total time on test $T(N,y_1,\dots y_n,y_o)$ and n are required. This is again seen in the likelihood function for λ on evidence of the ALT outcome $(N,y_1,\dots y_n,y_o)$ within context of the exponential failure function model:

$$L(\lambda\,|N,y_1,\dots y_n,y_o) = \binom{N}{1,\cdots 1,(N-n)} \prod_{j=1}^{n} [\lambda\, e^{-\lambda y_j}]\; e^{-\lambda(N-n)y_o}$$

$$= \binom{N}{1,\cdots 1,(N-n)} \lambda^n e^{-\lambda T(N,y_1\cdots y_n,y_o)},$$

with

$$\binom{N}{1,\cdots 1,(N-n)} = \frac{N!}{1!\cdots 1!(N-n)!} = \frac{N!}{(N-n)!}$$

the number of ways y_j , $j \leq n$ can each be chosen once and N - n not chosen out of the sample of N. As we have seen before, quite generally this normalization combinatoric cancels out of Bayes' theorem for determination of the posterior probability density $P(\lambda\,|c,N,y_1,\dots y_n,y_o)$ which contains all knowable information about λ on evidence of ALT outcome $(N,y_1,\dots y_n,y_o)$ and the prior information in $P(\lambda\,|c)$. If the prior information was equivalent to specifying n_c and T_c then the posterior density on λ is

$$P(\lambda\,|c,N,y_1,\dots y_n,y_o) = \lambda^{n+n_c}\,(T+T_c)^{n+n_c+1}\, e^{-\lambda(T+T_c)}/(n+n_c)!\ .$$

Its maximizer is

$$\lambda_{max} = (n+n_c)/(T+T_c)\ .$$

Protocol II: Test Until A Prescribed Number Fail

In this case the ALT protocol prescribes the number of failures n and tests until they are observed. The only difference from the previous case is that $y_o - y_n = 0$ in this case, where $y_o \geq y_n$ in the previous case, and in this case $y_n = y_o$ is a random variable with n prescribed; while in the previous case n was random with y_n prescribed. The likelihood function for λ on evidence of the ALT outcome $(N,y_1,\dots y_n)$ is

$$L(\lambda\,|N,y_1,\dots y_n) = \lambda^n\, e^{-\lambda T(N,y_1,\cdots y_n)}$$

with

$$T(N,y_1,\dots y_n) = \sum_{j=1}^{n} y_j + (N-n)y_n\ .$$

The posterior density for λ on evidence of this outcome and the information contained in the prior density $P(\lambda|c)$, again parameterized by (n_c,T_c) is

$$P(\lambda|c,N,y_1,...y_n) = \lambda^{n+n_c} (T + T_c)^{n+n_c+1} e^{-\lambda(T+T_c)}$$

with maximizer at

$$\lambda_{max} = (n + n_c)/(T + T_c) \quad .$$

We see that there is no finesse to be gained by manipulating the ALT protocol within the context of the exponential failure function hypothesis because the joint statistic (n,T) comprising the number of failures n and total time on test $T(N,y_1,...y_n,y_0)$ in the ALT outcome $(N,y_1,...y_n,y_0)$ constitutes the necessary and sufficient statistic for information on λ in this model.

Other modifications of the sampling method for the ALT protocol have also been investigated. They serve to illustrate the concept of sufficient statistic with the present failure model, but therefore give essentially similar results.

REFERENCES AND FOOTNOTES

1. Accelerated life tests are necessary to assure that the SSC magnets can be expected with specified confidence levels to perform at their specified reliability, availability and lifetime. See: SSC Conceptual Design Report, SSC-SR-2020, Sec.5.2.1, p.287; and Report of Task Force on SSC Magnet System Test Site, SSC-SR-1001, Sec. III, A., p.9.

2. An accelerated life test protocol for the SSC magnets has been established in: V. Karpenko, correspondence with D. Brown, 7/1/86, and D. Brown, correspondence with V. Karpenko. The proposed test protocol is summarized in E. Shrauner, SSC-N-215, 8/86.

The proposed test protocol allows for sampling in about one week of testing the equivalent of about one year of ordinary operational cycling anticipated for the SSC. Availability requirements on the SSC magnet system are severe, because it is comprised of about 9500 individual magnets in a serial-fault configuration. The individual component magnets must therefore be extremely reliable. A general relation among the stationary availbility A_{∞} , the mean time to repair MTTR, and the mean time between failures MTBF is

$$A_{\infty} = 1/[\,1 + \mathrm{MTTR}/\mathrm{MTBF}\,] \quad .$$

For our purposes MTTR is specified as in Ref. 1. to be about 1 week, or about 1/50 year. This determines MTBF through the above relation for A_{∞} as specified in Ref. 1. Because of the serial fault configuration of the component magnets in the system, a mean effective failure rate λ for the individual component magnets (considered to be all the same) can be assigned by:

$$\lambda = 1/(9500\ \mathrm{MTBF}) = [(1/A_{\infty}) - 1]/[9500\ \mathrm{MTTR}] \quad .$$

The assignment of a constant effective failure rate implies that the individual magnets fail for a complex variety of causes. This so-called exponential-failure-function hypothesis is chosen mostly for the simplicity of computation that it allows for obtaining rough estimates. If another failure-function is required for reality it may be incorporated into the present analysis without alteration of the structure laid out here.

The SSC Conceptual Design Report, Ref. 1., specifies A_{∞} for the magnet system as required to be 98 percent. On the other hand, it also specifies the availability for the whole SSC as required to be 80 percent; and it might be thought that in the sense in which we are concerned the magnet system is in some way equivalent to the whole SSC. For this reason both these values are considered in Ref.3. as perhaps two extremes of a range of values for A_{∞} . The value $A_{\infty} = .98$ specifies $\lambda = 1/(4560$ year), and the value $A_{\infty} = .80$ specifes $\lambda = 1/(760$ year) , as determined through the above relation.

3. Analyses of the probabilistic confidence for specified availabiltylevels obtainable from ALTs on SSC magnets using microcanonical Bayesian prior probability distributions have been outlined before in: E. Shrauner, SSC-N-215, 8/86; and SSC-N-254, 10/86.

4. Analysis of the probabilistic confidence for specified availabilty levels obtainable from ALTs on SSC magnets using the canonical Bayesian prior probability distribution derivable from maximum information entropy Bayesian analysis has been outlined in: E. Shrauner, SSC-N-310, 12/86,2/87.

5. The advantages of defining and working with the densities and the maximum likelihood estimators were suggested to me by Prof. R.E. Barlow, who has described these and other methods and concepts treated in this paper in: R.E. Barlow and F. Proschan, in Theory of Reliability, Soc. Ital. di Fisica, Bologna, 1986, corso XCIV.

6. The concept of sufficient statistic was suggested to me by Prof. E.T. Jaynes, who averred that it is more than a century old.

7. The stability of this particular form of prior expressed in terms of parameters n_c, T_c and its facile incorporation of all prior information relavent to λ in the prior density for λ and the propagation of this information through to the augmented posterior density in the same form expressed in terms of the integrated statistics $n + n_c$ and $T + T_c$ is an important phenomenon to be the subject of a separate study.