Bayes Estimation of Component-Reliability from Masked System-Life Data

Dennis K.J. Lin
Pennsylvania State University, University Park
John S. Usher, Member IEEE
University of Louisville, Louisville
Frank M. Guess, Associate Member IEEE
University of Tennessee, Knoxville

Key Words — **Masked data, Competing risk, Bayes estimation**

Summary & Conclusions — **This paper estimates component reliability from masked *series*-system life data, *viz*, data where the exact component causing system failure might be unknown. It focuses on a Bayes approach which considers prior information on the component reliabilities. In most practical settings, prior engineering knowledge on component reliabilities is extensive. Engineers routinely use prior knowledge and judgment in a variety of ways. The Bayes methodology proposed here provides a formal, realistic means of incorporating such subjective knowledge into the estimation process. In the event that little prior knowledge is available, conservative or even non-informative priors, can be selected.**

The model is illustrated for a 2-component *series* system of exponential components. In particular it uses discrete-step priors because of their ease of development & interpretation. By taking advantage of the prior information, the Bayes point-estimates consistently perform well, *ie*, are close to the MLE. While the approach is computationally intensive, the calculations can be easily computerized.

1. INTRODUCTION

Acronyms[1]

MLE maximum likelihood estimate
MTTF mean time to failure.

Reliability analysts are often interested in estimating the reliability of each component in a system through the analysis of system life data. Unlike individual component life testing, this type of analysis yields estimates that reflect component reliability after their assembly into an operational system. As such, the estimates account for the many degrading effects introduced by the system manufacturing, assembly, distribution, and installation. The resulting estimates can then be used to predict performance of new systems better.

Because of these advantages, companies are beginning to implement this type of estimation methodology; [9] describes one such implementation at IBM that has been successfully used to predict the reliability of newly developed computer hardware.

Component-reliability is often estimated from system-life data by using a *series*[2] system assumption and applying a competing-risks model. The observable quantities of interest are the system-life (failure or censoring time) and the exact component causing failure. Finding MLE for component-life distribution parameters has been widely addressed in the literature. However, in practice, this approach is often confounded by *masking* (the exact cause of system failure is unknown). *Masking* occurs frequently when exact diagnosis of the failure cause is too resource-consuming to conduct on every failed system. For example, in a complex system like a computer, it is often more cost effective to isolate the failure-cause to only a few circuit cards which can be quickly replaced. The analyst is then left with the time to system failure, but only partial knowledge of the failure-causing component.

Estimating component reliability from *masked* system-life data has received attention in the literature, but mostly from a classical statistics perspective. For example, Miyakawa [5] considers a 2-component *series* system of "exponential" components and derives closed-form expressions for the MLE. Under the same exponential assumption, [8] extend the Miyakawa results to a 3-component system; in all but a few special cases, closed-form MLE are intractable, and a simple iterative solution was proposed. Ref [3] further developed a procedure for finding the exact MLE in the 3-component case.

Ref [1] extends & clarifies the derivation of the general likelihood in the *masked* data case and examines the effect of *masking* on the *s*-bias and mean square-error of the MLE for a special-case, 3-component system of "exponential" components. Ref [1] also points out that these results are based upon *s*-independent (of failure cause) *masking*. Ref [2] extends [1] by investigating the effects of degrees of proportional-dependent *masking* on the MLE for a 2-component system; [7] provides a pseudo-graphical approach for estimating Weibull component reliability from *masked* data.

This paper presents a Bayes methodology for estimating component reliabilities from *masked* system-life data. This type of approach allows the analyst to quantify directly the prior engineering judgment in the development of the component reliability estimates. The prior function represents the degree-of-belief in each component and is incorporated into the reliability estimates. Our focus is on the use of step-wise functions to represent the component priors under the assumption that each component has exponentially distributed life. Section 2 illustrates the development of the Bayes model. Section 3 applies the model to a 2-component system. Section 4 illustrates its use with a numerical example.

[1] The singular & plural of an acronym are always spelled the same.

[2] The terms, *series* & *parallel* are used in their logic-diagram sense, irrespective of the schematic-diagram or physical-layout.

The Bayes analysis uses probability as a measure of degree-of-belief, not of relative frequency.

Notation

i	system index, $i=1,\ldots,n$ unless otherwise stated
T_i	random life of system-i
$T_{i,j}$	random life of component-j in system-i
$f_j(t)$, $R_j(t)$	[pdf, Sf] of life of component-j
θ_j	scale parameter of life distribution (also MTTF)
λ_j	$1/\theta_j$: failure rate of component-j
S_i	set of components known to contain the true cause of failure
$f(\boldsymbol{t}\|\boldsymbol{\theta})$	likelihood function of the sample data
$f(\theta_j)$	prior distribution for θ_j
$g(\boldsymbol{\theta}\|\boldsymbol{t})$	posterior distribution for $\boldsymbol{\theta}$
$\hat{}$, $\tilde{}$	implies: [MLE, Bayes point-estimate]
n_1, n_2, $n_{1,2}$	number of observations (in the sample) where S_i = [{1}, {2}, {1,2}]
$\sum_i$, $\prod_i$	[sum, product] over all i
$\mathcal{I}(\cdot)$	indicator function: $\mathcal{I}$(True)=1, $\mathcal{I}$(False)=0.

Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

2. BAYES MODEL

2.1 Assumptions

1. The $T_{i,j}$ are s-independent r.v., and are i.i.d r.v. for each j.

2a. Component-j life is exponentially distributed, with mean θ_j and failure rate λ_j.

2b. The θ_j are s-independent.

3. Systems are observed until failure (no censoring).
4. All system-components are in *series*.
5. *Masking* is s-independent of the failure-cause.

2.2 Model Derivation

A sample of n J-component *series* systems is life tested. *Series* implies: $T_i = \min_j\{T_{i,j}\}$ for all $i=1,2,\ldots,n$. We obtain estimates of θ_j, based upon our prior knowledge and our sample data. Under the Bayes framework, this is done by organizing one's prior degree-of-belief into prior distributions on all θ_j. These priors then represent the degree-of-belief about the value of θ_j prior to taking sample data. When combined with the sample results (the *masked* system-life data), the priors are then suitably transformed into one's posterior degree-of-belief, *viz*, a posterior distribution, through the use of Bayes theorem. Bayes estimates of the θ_j are then found as the mean of posterior distribution.

For each system, the observed quantities are t_i, and $S_i \subset \{1,2,\ldots,J\}$. If S_i contains a single element j, then the cause of system failure is known to be component j. If S_i contains all possible elements, $1,2,\ldots,J$, then the cause of failure is completely unknown. This subset approach considers the full range of possible information on system-failure causes.

Under assumption #5, a reduced (partial) likelihood is [1]:

$$f(\boldsymbol{t}|\boldsymbol{\theta}) = \prod_i \left[\sum_{j\in S_i} (f_j(t_i) \cdot \prod_s R_s(t_i))\right], \tag{1}$$

$\prod_s$ implies: product over s from 1 to J, excluding $s=j$.

For exponentially-distributed component-lives,

$$f(\boldsymbol{t}|\boldsymbol{\theta}) = \exp\left(-\left(\sum_i t_i\right)\cdot \sum_{s=1}^{J} (\lambda_s)\right)\cdot \prod_i \left[\sum_{j\in S_i} \lambda_j\right]. \tag{2}$$

We use a simple step-function prior of the form:

$$f(\theta_j) = \begin{cases} \alpha_{j,k}, \text{ for } \theta_j \in [a_{k-1}, a_k) \\ 0, \text{ for } \theta_j < 0; \end{cases}$$

$$\bigcup_{k=1}^{\infty} [a_{k-1}, a_k] = [0,\infty),$$

$$0 = a_0 < a_1 < a_2 < \ldots,$$

$$\sum_{k=1}^{\infty} \alpha_{j,k}\cdot(a_k - a_{k-1}) = 1.$$

There is always a partition of the time axis $[0,\infty)$ where this equation holds for all priors.

Then,

$$f(\theta_j) = \sum_{k=1} \alpha_{j,k}\cdot \mathcal{I}(\theta_j \in [a_{k-1}, a_k)).$$

The use of such step-function (discrete) priors is advantageous due to the ease with which one can quantify degree-of-belief in each component's mean life. Martz & Waller [4] discuss various methods that can be useful to engineers faced with the task of developing such step-function distributions.

Under assumption #2b, the joint prior is:

$$f(\boldsymbol{\theta}) = \prod_{j=1}^{J} \sum_{k=1}^{\infty} \alpha_{j,k}\cdot \mathcal{I}(\theta_j \in [a_{k-1}, a_k)).$$

3. 2-COMPONENT SYSTEM

The system has 2 components in *series*. A sample of n such systems is placed on test. For each system we observe a life, and a) S_i = {1} or S_i = {2}, or b) S_i = {1,2}. Then,

$$f(\boldsymbol{t}|\theta_1,\theta_2) = \exp(-(\lambda_1+\lambda_2)\cdot \sum_i t_i)\cdot \lambda_1^{n_1}\cdot\lambda_2^{n_2}\cdot(\lambda_1+\lambda_2)^{n_{1,2}}. \tag{4}$$

From (3) & (4),

$$f(t,\theta_1,\theta_2) = \prod_{j=1}^{2} \sum_{k=1}^{\infty} \alpha_{j,k} \cdot f(t|\theta) \cdot \mathcal{I}(\theta_j \in [a_{k-1}, a_k));$$

$$f(t) = \sum_{k=1}^{\infty} \left[\int_{a_{k-1}}^{a_k} \int_{a_{k-1}}^{a_k} \prod_{j=1}^{2} \alpha_{j,k} \cdot f(t|\theta_1,\theta_2) \cdot \mathcal{I}(\theta_j \in [a_{k-1}, a_k))\, d\theta_1\, d\theta_2 \right];$$

$$g(\theta_1,\theta_2|t) = f(t,\theta_1,\theta_2)/f(t). \quad (5)$$

The $\tilde{\theta}_1$, $\tilde{\theta}_2$, are the mean of the marginal posterior distribution in (5).

4. NUMERICAL EXAMPLE[2]

As in section 3, consider a *series* system of 2 "exponential" components. The data in table 1 represent the life and the true cause of failure for a random sample of $n=30$ systems. (The true cause of failure was found by observing the minimum life of the 2 components.) The data were simulated under the exponential assumption with $\theta_1=12$, $\theta_2=15$. The total time on test is:

$$\sum_i t_i = 209.8018.$$

To simulate the effect of various levels of *masking*, we randomly *masked* 10%, 30%, 50%, 70% of the failure causes. *Masked* observations are denoted in table 1 (columns 4-7). The θ_1, θ_2 in table 2A are evaluated as:

$$\hat{\theta}_j = \left[\sum_{i=1}^{n} t_i \right] / (\eta \cdot n_j), \quad (6)$$

$$\eta \equiv 1 + n_{1,2}/(n_1+n_2),$$

as presented in [3, 5]. The engineer's degree-of-belief is expressed by the discrete $f(\theta_j)$:

$$f(\theta_1) = \quad (7)$$

0.6, for $11 \le \theta_1 < 12$

0.3, for $12 \le \theta_1 < 13$

0.1, for $13 \le \theta_1 < 14$

0, otherwise;

[3]The number of significant figures is not intended to imply any accuracy in the estimates, but to illustrate the arithmetic.

Table 1. Simulated System-Life Data

[for a 2-component ($j=1,2$) system with *masking*
i → system index
column headings in [] → masking-level (%)
column [0] gives true failure-cause: $j=1,2$
in other columns: 'blank' → non-masked, ? → masked]

		Masking-Level (%)				
#i	t_i	[0]	[10]	[30]	[50]	[70]
1	12.099	1	?			?
2	21.516	1		?		
3	9.436	1		?	?	
4	4.197	1			?	?
5	3.939	2		?		
6	1.492	1			?	?
7	3.249	2			?	?
8	1.535	1				?
9	11.534	2		?		
10	2.713	1			?	?
11	3.624	2		?		
12	11.756	1	?	?	?	
13	3.058	1				?
14	9.605	1			?	?
15	1.291	2	?			?
16	8.299	1				?
17	27.049	2		?	?	
18	5.998	2			?	?
19	8.492	1		?		
20	7.865	2			?	?
21	7.214	2				?
22	0.613	1			?	?
23	2.551	2		?	?	
24	7.487	1				?
25	4.651	2		?		
26	4.926	1				?
27	14.453	1			?	?
28	3.732	1	?		?	?
29	0.059	2				?
30	5.373	2				?

$$f(\theta_2) = \quad (8)$$

0.2, for $12 \le \theta_1 < 13$

0.5, for $13 \le \theta_1 < 14$

0.3, for $14 \le \theta_1 < 15$

0, otherwise.

Then, as shown in figure 1, $f(\theta_1, \theta_2) = \alpha_{i,j} =$ (9)

0.12, for $11 \le \theta_1 < 12$ and $12 \le \theta_2 < 13$

0.30, for $11 \le \theta_1 < 12$ and $13 \le \theta_2 < 14$

0.18, for $11 \le \theta_1 < 12$ and $14 \le \theta_2 < 15$

0.06, for $12 \le \theta_1 < 13$ and $12 \le \theta_2 < 13$

0.15, for $12 \le \theta_1 < 13$ and $13 \le \theta_2 < 14$

0.09, for $12 \le \theta_1 < 13$ and $14 \le \theta_2 < 15$

0.02, for $13 \le \theta_1 < 14$ and $12 \le \theta_2 < 13$

0.05, for $13 \le \theta_1 < 14$ and $13 \le \theta_2 < 14$

0.03, for $13 \le \theta_1 < 14$ and $14 \le \theta_2 < 15$

0, otherwise.

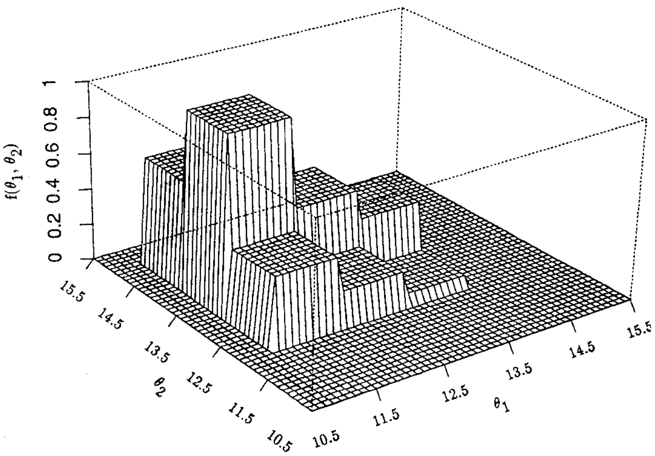

Figure 1. The Joint Prior: $f(\theta_1, \theta_2)$

While these degree-of-belief priors are generated arbitrarily for this illustration, discussions with engineers at IBM, Research Triangle Park reveal that they represent the manner by which subjective engineering judgment could be easily quantified. The $f(\boldsymbol{t},\theta)$, for

$$(n_1, n_2, n_{1,2}) = (14, 12, 4)$$

as an example, is shown in figure 2. The resulting degree-of-belief posterior, a function of n_1, n_2, $n_{1,2}$ as in (5), is shown in figure 3 and can be written as:

$$g(\theta|\boldsymbol{t}) = k^{-1}\cdot\alpha_{i,j}\cdot\exp\left(-(\lambda_1+\lambda_2)\cdot\sum_i t_i\right)$$

$$\cdot\lambda_1^{n_1}\cdot\lambda_2^{n2}\cdot(\lambda_1+\lambda_2)^{n_{1,2}},$$

$\alpha_{i,j}$ is given in (9),

$$k \equiv \int_\theta g(\theta|\boldsymbol{t})\, d\theta.$$

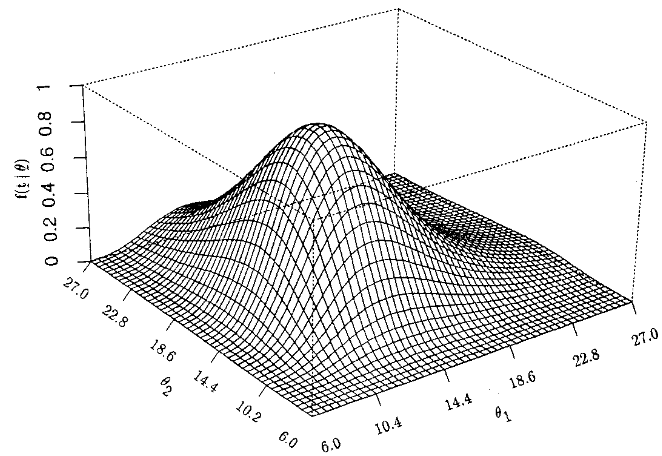

Figure 2. The Joint Likelihood Function: $f(\boldsymbol{t}|\theta)$
[for $(n_1, n_2, n_{1,2}) = (14, 12, 4)$]

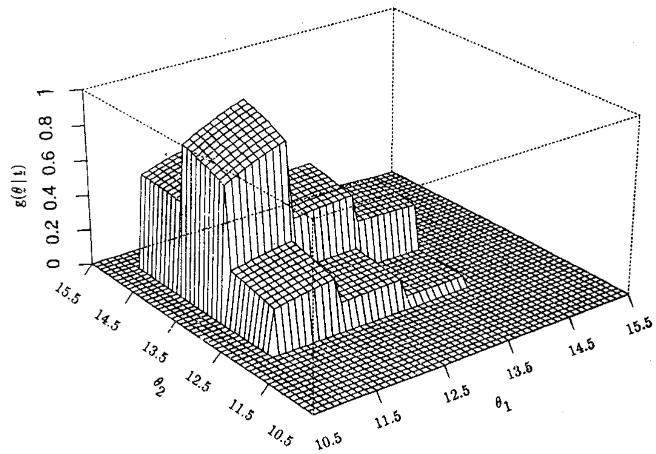

Figure 3. The Joint Posterior: $g(\theta,\boldsymbol{t})$
[for $(n_1, n_2, n_{1,2}) = (14, 12, 4)$]

Table 2. Component Parameter Estimates
[$\theta_1 = 12$, $\theta_2 = 15$]

	Masking-Level (%)				
	[0]	[10]	[30]	[50]	[70]
n_1	16	14	13	8	4
n_2	14	12	7	8	6
$n_{1,2}$	0	4	10	14	20
A. MLE					
$\hat{\theta}_1$	13.11	12.99	10.76	13.93	17.48
$\hat{\theta}_2$	14.99	15.15	19.98	13.93	11.66
B. Bayes Point-Estimate (Posterior Mean)					
$\tilde{\theta}_1$	11.52	11.52	11.58	11.54	11.55
$\tilde{\theta}_2$	12.99	13.00	13.09	12.97	12.94

Table 2B shows the Bayes point-estimates (posterior Mean) for each of the 4 levels of *masking*. The Bayes point-estimates (posterior Mode) are very sensitive to the prior distribution, *viz*, they are $(\theta_1, \theta_2) = (11.90, 14.00)$ for all cases. They are not very sensitive to the data. This means that the prior (what the engineer believes before the experiments) is so strong that the posterior (what the engineer believes after the experiments) is not influenced much at all by the actual data.

ACKNOWLEDGMENT

We are pleased to thank Mr. Wanzhu Tu for performing the necessary computations in Table 2. Dennis Lin was partially supported by NSF Grant DMS-9204007 and the Professional Development Award at the University of Tennessee.

REFERENCES

[1] F. Guess, J. Usher, T. Hodgson, "Estimating system and component reliabilities under partial information on the cause of failure", *J. Statistical Planning & Inference*, vol 29, 1991, pp 75-85.

[2] D.K.J. Lin, F. Guess, "System life data analysis with dependent partial knowledge on the exact cause of system failure", *4th Int'l Research Conf. Reliability*, 1991 Jun; Columbia, Missouri.

[3] D.K.J. Lin, J. Usher, F. Guess, "Exact maximum likelihood estimation using masked system data", IEEE Transactions on Reliability, Vol. 42, No. 4, December 1993.

[4] H.F. Martz, R.A. Waller, *Bayesian Reliability Analysis*, 1992; John Wiley & Sons.

[5] M. Miyakawa, "Analysis of incomplete data in competing risks model", *IEEE Trans. Reliability*, vol R-33, 1984 Oct, pp 293-296.

[6] W. Nelson, *Applied Life Data Analysis*, 1982; John Wiley & Sons.

[7] J.S. Usher, F.M. Guess, "An iterative approach for estimating component reliability from masked system life data", *Quality & Reliability Eng'g Int'l*, vol 5, 1989, pp 257-261.

[8] J. Usher, T. Hodgson, "Maximum likelihood analysis of component reliability using masked system life-test data", *IEEE Trans. Reliability*, vol 37, 1988 Dec, pp 550-555, 1988.

[9] J. Usher, S. Alexander, J. Thompson, "System reliability prediction based on historical data", *Quality and Reliability Eng'g Int'l*, vol 6, num 3, 1990, pp 209-218.

AUTHORS

Dr. Dennis K.J. Lin; Mgmt. Science and Information Systems; Penn State Univ; University Park, Pennsylvania 16802-1913 USA.

Dennis K.J. Lin is an Associate Professor of Statistics at the Pennsylvania State University. He is also an instructor for the Experimental Design Institute for Productivity Through Quality, at the University of Tennessee in Knoxville. Dr. Lin received his PhD (1988) in Statistics from University of Wisconsin, Madison. His research interests are design of experiments, quality engineering, and reliability. He has published in *Technometrics*, *J. Quality Technology*, *J. Statistical Planning and Inference*, *Computational Statistics and Data Analysis*, *etc*. He is an elected member of the International Statistical Institute, a Senior Member of the American Society for Quality Control, a member of the American Statistical Association, the Institute of Mathematical Statistics, and a Fellow of the Royal Statistical Society.

Dr. John S. Usher, PE; Dept. of Industrial Eng'g; Univ. of Louisville; Louisville, Kentucky 40292 USA.
Internet (e-mail): jsushe01@ulkyvm.louisville.edu

John S. Usher: For biography, see *IEEE Trans. Reliability*, vol 45, 1996 Jun.

Dr. Frank M. Guess; Dept. of Statistics; Univ. of Tennessee; Knoxville, Tennessee 37996-0532 USA.

Frank M. Guess, a Full Professor of Statistics at the University of Tennessee, received his PhD (1984) in Statistics from Florida State University; in that same year he won the Ralph A. Bradley Award for outstanding achievement for a PhD student from the Florida State University, Dep't of Statistics. He has published on reliability, censored data, and non-parametric statistics in a wide variety of journals, including *Quality and Reliability Eng'g Int'l*, *IEEE Trans. Reliability*, *Annals of Statistics*, *Biometrika*, *Biometrics* He is a member of the American Statistical Association, the Institute of Mathematical Statistics, and the Institute of Electrical and Electronics Engineers. Dr. Guess was on the Program Committee for the 4th International Reliability Research Conference in 1991 at the University of Missouri, and the 1st International Research Conference on Lifetime Data Models in 1994 at Harvard University.

Manuscript received 1995 November 11.

Publisher Item Identifier S 0018-9529(96)04777-X

◄TR►