AIM

Association for Information and Image Management

1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
301/587-8202

Centimeter

Inches

MANUFACTURED TO AIM STANDARDS
BY APPLIED IMAGE, INC.

1 of 1

Conf-9406224-1

LA-UR- 94-2275

Title:

MAXIMUM LIKELIHOOD ESTIMATION AND THE MULTIVARIATE
BERNOULLI DISTRIBUTION: AN APPLICATION TO RELIABILITY

Author(s):

P. H. Kvam

Submitted to:

Lifetime Data Models in Reliability & Survival Analysis

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Form No. 836 R5
ST 2629 10/91

MASTER

MAXIMUM LIKELIHOOD ESTIMATION AND THE MULTIVARIATE BERNOULLI DISTRIBUTION : AN APPLICATION TO RELIABILITY

Paul H. Kvam
Los Alamos National Laboratory

Abstract: We investigate systems designed using redundant component configurations. If external events exist in the working environment that cause two or more components in the system to fail within the same demand period, the designed redundancy in the system can be quickly nullified. In the engineering field, such events are called common cause failures (CCFs), and are primary factors in some risk assessments. If CCFs have positive probability, but are not addressed in the analysis, the assessment may contain a gross overestimation of the system reliability. We apply a discrete, multivariate shock model for a parallel system of two or more components, allowing for positive probability that such external events can occur. The methods derived are motivated by attribute data for emergency diesel generators from various U. S. nuclear power plants. Closed form solutions for maximum likelihood estimators exist in many cases; statistical tests and confidence intervals are discussed for the different test environments considered.

1. Introduction

This study is motivated by the following reliability problem. In the nuclear industry, redundancy in component groups is employed to provide additional reliability when an individual component reliability is considered inadequate for the operation. For instance, emergency diesel generators (EDGs) are used as a safeguard against loss of power in case offsite power at a nuclear power plant is temporarily interrupted. Once power is lost, two or more diesels are ready to be called upon to start up and provide emergency power until the problem is alleviated, even though only one working EDG is needed to resupply power. Practitioners in the nuclear industry realize that maintaining high reliability among the component groups significantly reduces the risk of critical events to the power plant.

If failure mechanisms work independently on each component in the system, the group reliability can be greatly increased using this strategy. However, not all failure modes are necessarily attributable to independent causes. External events that potentially cause the simultaneous failure of two or more of items, for instance natural disasters like a lightning strike or flood, are called common cause failures (CCFs), and are critical factors in reliability analysis, since the designed redundancy used to increase system reliability can be negated by a single common cause event. CCF events also include errors from internal sources, such as poor maintenance procedures or design flaws, both of which can induce the simultaneous failure of several components. If such events have a nontrivial probability but are excluded from the fault trees for a probabilistic risk assessment, the reliability of the system may be significantly overestimated.

If CCF can be readily identified after the component failures occur, modeling the reliability is straightforward. Unfortunately, failure modes are not always explained in plant records, and multiple failures are not partitioned into categories for independent failure or CCF. In this paper, we assume that only the attribute failure data are available for the analysis of component reliability, so it is not generally known which failures, if any, are due to common cause.

Problems of detecting and estimating common cause failure probabilities are analogous to problems described by competing risk models. See Chiang (1968) or Gail (1975) for expositions. Competing risk models are constructed to apply to lifetime distributions of items prone to failure

from two or more 'competing' sources. They commonly arise in actuarial science and epidemiology, and typical applications of competing risk methods presume that information about the modes of failure is included within the data. Models of competing risk and CCF differ mostly in their application, rather than an underlying theory.

The reliability literature contains numerous articles and applications of CCF analysis. Mosleh, et al. (1988) reviews much of the current literature, and serves as an excellent reference for useful estimation techniques. The binomial failure rate model, investigated in Atwood (1986), models CCFs as external system shock parameters, independent of the individual component failure probabilities, similar to competing risk models. Like models in competing risk, however, estimation is unrealizable at times due to nonidentifiable parameters. That is, in some settings different values of parameters will lead to the same underlying lifetime distribution, making the parameters impossible to identify uniquely. As in Atwood (1986), we model CCF as an external, stochastic event, occurring independently of the individual component failures, but identifiability is maintained in all instances.

In this paper it is our goal to model the reliability (or unreliability) of a system of two or more components using demand failure data, assuming the presence of possible common cause events that could fail all the items in the same group. Implicit in the use of the model is the assumption that multiple failures of all components on test are of main concern. The statistical model discussed in this paper is an application and extension of the multivariate Bernoulli model introduced by Boyles and Samaniego (1983), where procedures are introduced to model the (possibly unequal) reliability of a group of components, leading to derivation of a general method for obtaining numeric solutions to reliability parameters. The Bernoulli model is a discrete version of the multivariate exponential model by Marshall and Olkin (1967), and the methods of estimation in section 2 correspond to those by Proschan and Sullo (1976). We discuss specific applications of this model, derive explicit solutions and obtain statistical properties.

We examine three particular scenarios to illustrate the estimation methods. First, we look at the case in which two components share a common cause group, both having identical underlying reliability. This is a typical system set-up for EDGs in a power plant, for example. Next, the two system components are not assumed to have the same reliability. This parameterization is useful in cases where assumptions of homogeneous components and systems are deemed unrealistic; e.g., if plant maintenance procedures change significantly between stations or if there exists obvious points in time when reliability significantly increases (replacement, for instance) or decreases (aging or imperfect repair). Finally, we consider the model for three system components with the same reliability.

For each of the three examples, we apply the method of maximum likelihood to derive parameter estimates. Each component demand (or start-up) is assumed to be independent of past and future demands, and we further assume the reliability remains constant throughout the time period corresponding to the data. Tests for decreasing reliability using attribute data are discussed in Martz, et al. (1993). The resultant MLEs from the CCF model are consistent, and we compare the corresponding asymptotic variances with those of similar estimators based on models for which CCF is not admitted.

2. The Model

For the general model of m components operating in a parallel system, the component unreliability can be numerically computed, using techniques described in Boyles and Samaniego (1983). We focus on specific applications for which explicit solutions are available, and investigate various properties of the estimates.

The unreliability of each component in the system is modeled in terms of the parameter p , where p is the failure-on-demand probability. All components are tested simultaneously, so the chance exists that more than one items fail in a single test. In addition to the unreliability parameter, the probability of a common cause failure disabling the entire group of components on any given demand is denoted q . The corresponding parameter space is $\{(p, q) : 0 \leq p < 1, 0 \leq q \leq 1\}$. Occurrences of CCFs and individual failures are assumed to be

independent, thus we model the probability distribution for the number of system components failing in any given demand trial as

$$P(j \text{ failures of } m \text{ components}) = \binom{m}{j} (1-q)p^j(1-p)^{m-j}, 0 \leq j < m.$$

$$P(\text{all } m \text{ components fail}) = q + (1-q)p^m$$

The data are recorded from n tests of both test items, and x_i represents the number of times (out of n) in which i items failed, $i=0, \dots, m$. Note that $x_0 + \dots + x_m = n$. Based on this observed failure data, the corresponding likelihood function is expressed

$$L(p, q | \mathbf{x}) = \binom{n}{x_0, \dots, x_m} (1-q)^{\sum_{j=0}^{m-1} x_j} p^{\sum_{j=0}^{m-1} j x_j} (1-p)^{\sum_{j=0}^{m-1} (m-j) x_j} (q + (1-q)p^m)^{x_m}.$$

2.1 Method of Maximum Likelihood

For any $m > 1$, the likelihood function has a unique stationary point if $x_j > 0$ for some $j < m-1$. With small samples, the MLE may lie outside the parameter space, in which case ad-hoc estimators can be formulated. In these instances, the MLE for q is negative, so we intuitively suggest estimating q with $\hat{q} = 0$, and estimating p using the partial likelihood $L(p | \mathbf{x}, q = 0)$, which yields the solution

$$\tilde{p} = \frac{1}{mn} \sum_{j=0}^{m-1} j x_j.$$

We can show the MLE is contained in the parameter space if the solution $(\hat{\beta})$ to the likelihood equations satisfies

$$\left(\sum_{j=0}^{m-1} (k-j) x_j \right) p^m - \left(k \sum_{j=0}^{m-1} x_j \right) p + \sum_{j=0}^{m-1} j x_j < \tilde{p}.$$

In this case, we solve for the common cause parameter in terms of $\hat{\beta}$:

$$\hat{q} = \frac{x_m - n \hat{p}^m}{(1 - \hat{p}^m) n}.$$

Specific cases and examples are detailed in the following sections. Closed form solutions are available for the specific models we discuss in detail. In these cases, $m \leq 3$. For larger m , we can easily obtain numeric solutions using one of several iterative techniques.

2.2 Test of the Common Cause Parameter

Hypothesis tests for the common cause failure parameter are of special interest. Specifically, we are interested in testing $H_0: q = 0$ vs. $H_a: q > 0$ using the information from attribute data. The test statistic derived from the Neyman-Pearson ratio lacks a simple distribution, so exact tests are not easily tractable. Instead, a simple Chi-square test for goodness-of-fit can be used based on $(m+1)$ multinomial parameters, themselves a function of one unknown parameter (p). The Chi-square statistic, below, is based on $(m-1)$ degrees of freedom:

$$X_{m-1}^2 = \sum_{i=0}^m (x_i - \hat{E}(x_i))^2 / \hat{E}(x_i)$$

where $\hat{E}(x_i) = E(x_i | p=\hat{p}, q=0)$ for $i = 0, \dots, m$. In practice, the approximate test does well with large samples, but lacks power with small m and n , especially if q is close to zero. This is apparent in the test's non-centrality parameter; under H_a , the test statistic is approximately distributed as a chi-square random variable with $(m-1)$ degrees of freedom and non-centrality parameter equal to $n q^2 (1 - p^{-m})$.

3. Inference for small systems

In this section we focus on specific applications where $m = 2$ or 3 . Though explicit formulation of MLEs is not possible for general m , we do obtain closed-form solutions in these special cases, hence the properties of the estimators are elementary. Proofs for the derivation of the following MLEs are messy, but are easily obtainable with a symbolic program and sufficient computing facilities. Computations in this paper were performed using Mathematica.

3.1 A system of two identical components

With two components tested in parallel, we observe $\{x_0, x_1, x_2\}$. In the region $X^* = \{x : x_1^2 < 4x_0x_2\}$, L has a unique stationary point at $\hat{p} = x_1/(2x_0+x_1)$, and $\hat{q} = (4x_0x_2 - x_1^2)/4nx_0$, since the Hessian matrix of the log likelihood is shown to be negative definite around the unique solution to the log likelihood equations. If $x \notin X^*$, the MLE does not necessarily exist, and the solutions that do exist fall outside the parameter space. In this situation, a plausible ad-hoc rule is suggested: estimate q to be zero, and solve for p via the marginal likelihood, which is multinomial based on binomial probabilities, with a unique solution to the likelihood equations at $\bar{p} = (x_1 + 2x_2)/2n$.

If enough tests are run, and n is large, the MLEs are expected to perform well. Properties are discussed below. Problems may exist for small samples, however. If no simultaneous failures occur ($x_2 = 0$), L is maximized at $q = 0$, and the MLE for p is $\bar{p}$. If x_1 or x_0 are zero, the likelihood is not maximized within the parameter space, so the ad-hoc rule is again suggested. With highly reliable components, small samples may yield few or no multiple failures. The sample size requirement for $(1 - \alpha)100\%$ assurance that $x_2 > 0$, given component reliability $1 - p = 0.95$ and $q = \{0.00, 0.05, 0.10, 0.20\}$ is listed below, for illustrative purposes:

q	$\alpha = 0.20$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
0.00	643	920	1197	1840
0.05	30	43	56	86
0.10	15	22	28	43
0.20	8	11	14	21

We can make use of the asymptotic properties of maximum likelihood estimates to create approximate confidence intervals and tests for the model parameters, provided the sample size is adequate. Since regularity conditions of multivariate maximum likelihood are satisfied by $L(p, q)$ (see Chapter 6 of Lehmann (1983), for instance), the MLEs are consistent as well as

asymptotically efficient in that $\sqrt{n}(\hat{q} - q) \rightarrow N(0, [I(p, q)]_{(1,1)}^{-1})$ and $\sqrt{n}(\hat{p} - p) \rightarrow N(0, [I(p, q)]_{(2,2)}^{-1})$,

$$\text{where } [I(p, q)]^{-1} = \begin{bmatrix} \frac{(1-q)(q(1-p)^2 + p^2)}{n(1-p)^2} & \frac{-p^2}{n(1-p)} \\ \frac{-p^2}{n(1-p)} & \frac{p(1+p)}{2n(1-q)} \end{bmatrix}$$

and I denotes the information matrix for the likelihood function. Tests and intervals can be easily constructed using parameter estimates and the asymptotic variances ($\sigma_{\hat{q}}^2 = (1-q)(q(1-p)^2 + p^2)/(n(1-p)^2)$, $\sigma_{\hat{p}}^2 = p(1+p)/(2n(1-q))$) from the information matrix inverse.

If common cause failure is not admitted in the model (so $q = 0$), the failure probabilities are assumed to be binomial, hence our estimate of p is $\tilde{p} = (x_1 + 2x_2)/2n$, and its calculated variance reduces to $p/2n$. Note that $\sigma_{\tilde{p}}^2/\sigma_{\hat{p}}^2 = (1-q)/(1+p) \leq 1$ for all values (p, q) in the parameter space, and $\hat{p} \leq \tilde{p}$, with equality only if $x_1 = x_2 = 0$. As a consequence, if the model is incorrect and the possibility of common cause failures exists for the system components in the working environment, reliability estimates are biased, and the estimate's uncertainty is grossly underestimated.

The bias inherent in $\tilde{p}$ makes asymptotic comparisons of the estimators pointless, because the variances are $O(n^{-1})$, and the bias is $O(1)$. The relationship between $\hat{p}$ and the underlying CCF parameters concerns the practitioner, since the uncertainty associated with the unreliability estimate (p) changes as q varies in $(0, 1)$. The asymptotic variance for $\hat{p}$ increases only gradually for small values of q , which we realistically expect to be much less than 0.5.

3.2. Two Components with Unequal Reliability

The second scenario is the same as the first except that the system components may have unequal reliability. This parameterization is useful if we observe data from two plants with significantly different maintenance procedures, two component brands that may differ in quality, or two greatly separated time periods in which the data are obtained. For this situation, we define parameters p_1 and p_2 as the demand failure probabilities for the first and second test items, respectively. Without the need for generality, we simplify the notation: from n tests of the system, we observe

- y_0 = Number of instances in which no components failed.
- y_1 = Times in which only the first component failed.
- y_2 = Times in which only the second component failed.
- y_3 = Times in which both components failed.

Given (y_0, y_1, y_2, y_3) , the likelihood function for p_1 , p_2 and q is

$$L(p, q | y) = \binom{n}{y_0, y_1, y_2, y_3} (1-q)^{y_0+y_1+y_2} p_1^{y_1} p_2^{y_2} (1-p_1)^{y_0+y_1} (1-p_2)^{y_0+y_2} (q + (1-q)p_1 p_2)^{y_3}.$$

In the region $Y^* = \{y : y_1 y_2 < y_0 y_3\}$, the likelihood achieves a unique maximum at $\hat{p}_1 = y_1/(y_0+y_1)$, $\hat{p}_2 = y_2/(y_0+y_2)$ and $\hat{q} = (y_0 y_3 - y_1 y_2)/ny_0$. This can be proven directly via the log-likelihood. If $y \notin Y^*$, we again estimate q to be zero, and compute the individual reliability parameter estimates as $\tilde{p}_1 = (y_1 + y_3)/n$ and $\tilde{p}_2 = (y_2 + y_3)/n$. From the information matrix, we deduce the approximate asymptotic variances to be

$$\sigma_{p_1}^2 = \frac{(1-p_1)p_1(q+(1-q)p_1 p_2)}{(1-q)(1-p_2)\phi(q, p_1, p_2)}, \quad \sigma_{p_2}^2 = \frac{(1-p_2)p_2(q+(1-q)p_1 p_2)}{(1-q)(1-p_1)\phi(q, p_1, p_2)} \quad \text{and} \\ \sigma_q^2 = \frac{(1-q)(q+(1-q)p_1 p_2)(p_1 p_2 + q(1-p_1)(1-p_2))}{(1-p_2)(1-p_1)\phi(q, p_1, p_2)}$$

where $\phi(q, p_1, p_2) = q^2 + 2q(1-q)p_1 p_2 + p_1^2 p_2^2 (q+1)^2$. A test for the common cause parameter can be deduced using the approach from section 2.1.

3.3. Three Identical System Components

In this case, we have three test items with identical reliability working in the same system. The event of a CCF will fail all three items. We define the parameters p and q , and the data (x_0, x_1, x_2) as in (1). Additionally, we define x_3 as the number of times (out of n) all three diesels fail upon demand. For three components, the likelihood is expressed as

$$L(p, q | x) = \binom{n}{x_0, x_1, x_2, x_3} (1-q)^{x_0+x_1+x_2} p^{x_1+2x_2} (1-p)^{3x_0+2x_1+x_2} (q + (1-q)p^3)^{x_3}$$

The MLEs have closed-form solutions of

$$\hat{p} = \sqrt{\frac{3x_0 + 6x_1 + 9x_2}{12x_0 + 8x_1 + 4x_2}} - \frac{1}{2}, \quad \text{and} \quad \hat{q} = \frac{x_3/n - \hat{p}^3}{1 - \hat{p}^3}.$$

Again, the parameter space is constrained by the data, thus alternative estimators are formed as before for some small samples in which $\hat{p} < 0$ or $\hat{q} < 0$. For larger samples, we can use the approximate asymptotic variances of $\hat{p}$ and $\hat{q}$:

$$\sigma_{\hat{p}}^2 = \frac{p(p^2 + p + 1)}{3n(1-q)(1+2p)}, \quad \text{and} \quad \sigma_{\hat{q}}^2 = \frac{(1-q)(p^3 + q(2p^3 - 3p^2 + 1))}{n(1-p)^2(1+2p)}.$$

4. Examples

The first set of data is found in Atwood (1980). From various power plants in the United States, single and multiple failures of the pressurized water reactor (PWR) auxiliary feedwater systems are tabulated. The test unit in this case is a train, which includes the strainer, pump and connecting valves. For purpose of illustration, we assume the reliability of the auxiliary feedwater systems are approximately homogeneous, and that we observe one system demand for every

calendar month of operation (as described by Atwood(1980)). The common cause group size is $m = 2$, and in 1641 starts, we record $x_0 = 1538$, $x_1 = 68$ and $x_2 = 35$. For the hypothesis test of $q = 0$, the Chi-square statistic is calculated to be 387.17, corresponding to a p-value far less than 0.0001. From the likelihood equations, we deduce $\hat{p} = 0.0216$ (with $\sigma_{\hat{p}} \approx 0.0113$) and $\hat{q} = 0.0209$ (with $\sigma_{\hat{q}} \approx 0.0198$).

Next, We consider a three component parallel system of EDGs. In this example, we suppose that common cause events are typically lethal (i.e., they cause all the components to fail), thus failure events involving exactly two of the three diesels are relatively rare and not usually due to CCF. If such events do occur, we treat them as lethal and accept a slightly more conservative solution for the analysis. In 80 attempted start-ups, suppose that we observe $x_0 = 48$ successes, $x_1 = 23$ failures of a single component, $x_2 = 4$ failures of two components, and $x_3 = 5$ simultaneous failures of the entire system (all three EDGs). For the test of $q = 0$, $X^2 = 38.755$, which corresponds to a p-value of less than 0.0001, indicating with strong confidence that $q > 0$. The MLEs and approximate standard deviations are $\hat{p} = 0.1402$, ($\sigma_{\hat{p}} \approx 0.0237$), $\hat{q} = 0.0599$, (and $\sigma_{\hat{q}} \approx 0.0272$).

5. Conclusion

Based on common cause failure models applied in the nuclear industry, we derive and illustrate simple estimation and testing methods using discrete (fail vs. no fail) data. Methods for small component groups, where closed form estimators can be derived, are emphasized in section 3. The results provide insight for practitioners who may infrequently have to deal with common cause data for which CCF is not identified, thus the methods proposed are not dependent on being able to distinguish CCF from multiple independent failures.

The model does not include a Bayesian treatment. Similar methods based on input prior information are discussed in Kvam and Martz (1994).

References

- Atwood, C. L., (1980), "Common cause and individual failure and fault rates for licensee event reports of pumps at U. S. commercial nuclear power plants," Technical Report EGG-EA-5289, EG&G Idaho, Inc., Idaho Falls, Idaho.
- Atwood, C. L. (1986), "The binomial failure rate common cause model," *Technometrics*, No. 2, Vol. 28, 139-148.
- Boyles, R. A.. and Samaniego, F. J. (1983), "Maximum Likelihood Estimation for a Discrete Multivariate Shock Model," *Journal of the American Statistical Association*, vol. 78, 445-448.
- Chiang, C. L. (1968), Introduction to Stochastic Processes in Biostatistics, John Wiley and Sons, New York.
- Gail, M. (1975), "A review and critique of some models used in competing risk analysis," *Biometrics*, 31, 209-222.
- Kvam, P. H. and Martz, H. F. (1994) "Bayesian inference in a discrete common cause shock model," Los Alamos National Laboratory Technical Report, LA-UR-94-1414.
- Lehmann, E. L. (1983), Theory of Point Estimation, John Wiley and Sons, New York.

- Marshall, A. W., and Olkin, I. (1967), "A multivariate exponential distribution," Journal of the American Statistical Association, 62, 30-44.
- Martz, H. F., Abramson, L. R., Tietjen, G. L., Kvam, P. H. (1993), "On monitoring nuclear power plant emergency diesel generator reliability," American Statistical Association 1993 Proceedings of the Section of Physical Sciences and Engineering.
- Mosleh, A., Fleming, K. N., Parry, G. W., Paula, H. M., Worledge, D. H., and Rasmuson, D. M.(1988), "Procedures for treating common cause failures in safety and reliability studies," NUREG/CR-4780, EPRI NP-5613, Volume 1 and Volume 2, prepared for the Nuclear Regulatory Commission and the Electric Power Research Institute by Pickard, Lowe, and Garrick, Inc.
- Proschan, F., and Sullo, P. (1976), "Estimating the parameters of a multivariate exponential distribution," Journal of the American Statistical Association, 77, 465 - 472.

DATE

FILMED

10 / 3 / 94

END