Summary: Exact Maximum Likelihood Estimation Using Masked System Data

Alexander Towell
atowell@siue.edu

1 Introduction

Often, systems (like patients) in reliability studies can be reasonably approximated as series systems consisting of n components (1-out-of-n). Estimating these component reliabilities is of practical interest since they can be used to help predict the likely cause of failure.

Often, when performing this estimation based on a given sample, the failures are masked, i.e., the exact cause of failure is only known to be contained in some subset of the components. Thus, the statistics are system lifetime and partial information on the cause of failure.

This paper extends previous results by deriving an exact, closed-form maximum likelihood estimates (MLE) for a 1-out-of-3 system, where each component is exponentially distributed and is either in a failed state or non-failed state. Additionally, the MLE is based on samples of size n in which the systems are i.i.d. and s-independent failure masking is provided.

2 Maximum likelihood estimation

Given a sample from a population and a parametrized distribution, MLE is a technique one can use to find the parameters for the distribution that maximizes the likelihood function of the selected model, i.e., maximizes the compatibility of the selected model with the sample.

For exact, uncensored observations, the likelihood function L for a sample of size n is:

$$L(\vec{\theta}, \vec{t}) = \prod_{i=1}^{n} f_s(t_i; \vec{\theta}) \tag{1}$$

where f_s is the pdf of the system such that $Pr\left(t_a < T < t_b\right) = \int_{t_a}^{t_b} f_s(w; \vec{\theta}) dw$, i.e., the probability that the system will fail at some time $t \in (t_a, t_b)$.

Maximum likelihood estimation solves the following optimization problem:

$$\arg\max_{\vec{\theta}} L(\vec{\theta}, \vec{t}) \tag{2}$$

In general, this is an intractable (non-linear programing) problem requiring numerical techniques to find approximate (suboptimal) solutions. Moreover, the numerical technique can even fail to converge. The problem is further compounded by the presence of failure masking.

However, we are interested in the 1-out-of-3 system described previously. The likelihood function L for this system for parameters $\lambda_1, \lambda_2, \lambda_3$ is:

$$L = \prod_{i=1}^{n} \left[\sum_{s \in S_i} \left(f_s(t_i) \prod_{j \neq s} R_j(t_i) \right) \right] = \prod_{i=1}^{n} \left[\sum_{s \in S_i} \left(\lambda_s \exp(-\lambda_s t_i) \prod_{j \neq s} \exp(-\lambda_s t_i) \right) \right] \tag{3}$$

where $R_j(t) = \exp(-\lambda_j t)$ is the survival function for component j and S_i is the subset of components known to contain the cause of failure for system i.

To find $\arg\max_{\vec{\theta}} L$, in this case we can find the stationary points of $\ln(L)$ (since the parameters that maximize L also maximize $ln(L)$:

$$\frac{\partial}{\partial \theta_k} \ln L = -T + \frac{n_k}{\lambda_k} + \frac{n_{12}^k}{\lambda_1 + \lambda_2} + \frac{n_{13}^k}{\lambda_1 + \lambda_3} + \frac{n_{23}^k}{\lambda_2 + \lambda_3} + \frac{n_{123}}{\lambda_1 + \lambda_2 + \lambda_3} = 0, k = 1, 2, 3 \tag{4}$$

where n_k is the number of failures in which $S_i = k, k = 1, 2, 3$, n_{jk} is the number of failures in which $S_i = j, k, j, k = 1, 2, 1, 3, 2, 3$, n_{123} is the number of failures in which $S_i = 1, 2, 3$, $T = \sum_{i=1}^n t_i$, and $n_{ij}^k = n_{ij}$ if $k = i$ or $k = j$, 0 otherwise.

Finding a closed-form solution for the stationary points $\vec{\lambda}$ is not obvious. However, the problem can be transformed into the equivalent problem:

$$\begin{aligned}
\frac{n_1}{\alpha_1} - \frac{n_{23}}{\alpha_1 + \alpha_3} &= 1 \\
\frac{n_1}{\alpha_1} - \frac{n_{23}}{\alpha_1 + \alpha_3} &= 1 \\
\frac{n_1}{\alpha_1} - \frac{n_{23}}{\alpha_1 + \alpha_3} &= 1 \\
A = 1 + \frac{n_{12}}{\alpha_1 + \alpha_2} \frac{n_{13}}{\alpha_1 + \alpha_3} \frac{n_{23}}{\alpha_2 + \alpha_3} \frac{n_{123}}{\alpha_1 + \alpha_2 + \alpha_3} & \\
\alpha_k = \lambda_k \frac{T}{A}, k = 1, 2, 3 &
\end{aligned} \tag{5}$$

Solving eq 5 for the three unknown a_k values is an easier problem that can be expressed in a closed-form solution (finding the roots of a 4^{th} degree polynomial). Once these are solved, the MLE estimate for the parameters is:

$$\vec{\lambda} = \frac{A}{T}[\alpha_1, \alpha_2, \alpha_3] \tag{6}$$

Several simplifying cases, such as $n_{13} = n_{23} = 0$, further simplify the closed-form solution. Thus, unlike in previous efforts, this closed-form solution will provide an exact MLE in a single step (so lack of convergence is a non-issue).