Summary: Estimating Component Reliabilities from Incomplete System Failure Data

Alexander Towell
atowell@siue.edu

1 Introduction

Consider a system that consists of k components. If the system's components' test data is provided, then such data may be used to estimate their respective reliability distributions. Being able to perform this estimation is of practical interest, e.g., if a system fails, in what order should its components be analyzed to identify the cause of failure? That is, given a system's failure time, what is the probability that a specified component was the cause of the system's failure?

If component test data is not provided, component reliabilities may be estimated from system failure data instead. The article[1] primarily concerns itself with estimating component reliabilities from masked system failure data.

An approach to performing this estimation is outlined in which the following is true:

- The system is series or parallel.
- The system's components follow arbitrary reliability distributions.
- The sample consists of pairs of masked system failures and failure times, in which masked system failures are a subset of the components that may have caused the failure to occur, and failure time is an exact failure time, interval failure time, or right-censored failure time.

2 Series systems

2.1 System and component distributions

Assuming the k component reliability distributions, $R_1, ..., R_k$, are known, system reliability (eq 1) for a series system can be derived. Once this is known, the system cdf (eq 2) and system pdf (eq 3) can also be derived.

$$R_s(t) = \prod_{j=1}^{k} R_j(t) \tag{1}$$

$$F_s(t) = 1 - R_s(t) \tag{2}$$

$$f_s(t) = -\frac{\mathrm{d}}{\mathrm{d}t} R_s(t) = \sum_{j=1}^{k} \left(f_j(t) \prod_{p \neq j}^{k} R_p(t) \right) \tag{3}$$

Note that in eq 3, $f_j(t) = -\frac{\mathrm{d}}{\mathrm{d}t} R_j(t)$.

2.2 Likelihood functions

If $R_1, ..., R_k$ are not known, but a sample of masked system failures are available, then maximum likelihood estimation (MLE) may be used.

Let C_i be the set of components that is possibly causing the i^{th} system failure. If the i^{th} system failure specifies an exact failure time t_i, then the likelihood function for the i^{th} observation is:

$$l_i = \sum_{j \in C_i} \left[f_j(t_i) \prod_{p \neq j}^{k} R_p(t_i) \right] \tag{4}$$

Note that, on the one hand, if C_i has only a single component, then eq 4 simplifies to a case in which the exact cause of failure is known (i.e., an instance of test component data). On the other hand, if C_i has k components, then eq 4 reduces to an instance in which one only knows at what time the i^{th} failure occurred.

If the i^{th} system failure specifies an interval failure time $t \in [t_{i_L}, t_{i_U}]$, then the likelihood function for the i^{th} observation is:

$$l_i = \sum_{j \in C_i} \left[\int_{t_{i,L}}^{t_{i,U}} f_j(t_i) \prod_{p \neq j}^{k} R_p(t_i) dt \right] \tag{5}$$

If the i^{th} system failure specifies a suspension (right censored failure time), then the likelihood function for the i^{th} observation is:

$$l_i = R_s(t_i) = \prod_{p=1}^{k} R_p(t_i) \tag{6}$$

The complete likelihood function (for the i.i.d. sample of n failures) is:

$$L = \prod_{i=1}^{n} l_i \quad (7)$$

For each component j, R_j must be its parametrized reliability distribution so that that the MLE of L can be found. Typically, it is easier to find the MLE of $\ln(L)$, since the parameters that maximize L also maximize $\ln(L)$. Note that many optimization techniques may be used to find the MLE (since no restriction is made on the types of distributions, numerical techniques may be necessary).

With these MLE estimates of $R_1, ..., R_k$, the probability that a system failure at time t was caused by component j is:

$$P_j = \frac{f_j(t) \prod_{p \neq j} R_p(t_i)}{\sum_{i=1}^{k} \left[f_i(t) \prod_{p \neq i} R_p(t_i) \right]} \quad (8)$$

And, likewise, the probability that a system failure at time $t \in [t_U, t_L]$ was caused by component j is:

$$P_j = \frac{\int_{t_L}^{t_U} f_j(t) \prod_{p \neq j} R_p(t_i) dt}{\sum_{i=1}^{k} \left[\int_{t_L}^{t_U} f_i(t) \prod_{p \neq i} R_p(t_i) dt \right]} \quad (9)$$

With probability estimates P_j (eqs 8 and 9), one can perform tasks like failure analysis by inspecting the components in an optimal order for a system's given failure time.

3 Parallel systems

3.1 System and component distributions

Assuming the k component reliability distributions, $R_1, ..., R_k$, are known, system reliability (eq 10) for a parallel system can be derived. Once this is known, the system cdf (eq 11) and system pdf (eq 12) can also be derived.

$$R_s(t) = 1 - \prod_{j=1}^{k} (1 - R_j(t)) \quad (10)$$

$$F_s(t) = 1 - R_s(t) = \prod_{j=1}^{k} (1 - R_j(t)) \quad (11)$$

$$f_s(t) = \sum_{j=1}^{k} \left(f_j(t) \prod_{p \neq j}^{k} R_p(t) \right) \quad (12)$$

3.2 Likelihood functions

If $R_1, ..., R_k$ are not known, but a sample of masked system failures are available, then MLE may be used, as in section 2. If the i^{th} system failure specifies an exact failure time t_i, then the likelihood function for the i^{th} observation is:

$$l_i = \sum_{j \in C_i} \left[f_j(t_i) \prod_{p \neq j}^{k} (1 - R_p(t_i)) \right] \quad (13)$$

If the i^{th} system failure specifies an interval failure time $t \in [t_{i_L}, t_{i_U}]$, then the likelihood function for the i^{th} observation is:

$$l_i = \sum_{j \in C_i} \left[\int_{t_{i,L}}^{t_{i,U}} f_j(t_i) \prod_{p \neq j}^{k} (1 - R_p(t_i)) dt \right] \quad (14)$$

If the i^{th} system failure specifies a suspension (right censored failure time), then the likelihood function for the i^{th} observation is:

$$l_i = R_s(t_i) = 1 - \prod_{p=1}^{k} (1 - R_p(t_i)) \quad (15)$$

Similar to the series case, the MLE of the joint likelihood function L is generated and the MLE for each component, $R_1, ..., R_k$, can be found. Subsequently, the probability that a system failure was caused by a given component can be computed.

References

[1] H. Guo, F. Szidarovszky, and P. Niu. Estimating component reliabilities from incomplete system failure data.