Reliability Estimation in Series Systems: Maximum Likelihood Techniques for Right-Censored and Masked Failure Data

Alex Towell

Abstract

This paper investigates maximum likelihood techniques to estimate component reliability from masked failure data in series systems. A likelihood model accounts for right-censoring and candidate sets indicative of masked failure causes. Extensive simulation studies assess the accuracy and precision of maximum likelihood estimates under varying sample size, masking probability, and right-censoring time. The studies specifically examine the accuracy (bias) and precision of estimates, along with the coverage probability and width of BCa confidence intervals. Despite significant masking and censoring, the maximum likelihood estimator demonstrates good overall performance. The bootstrap yields reasonably well-calibrated confidence intervals even for small sample sizes. Together, the modeling framework and simulation studies provide rigorous validation of statistical learning from masked reliability data.

Contents

1 Future Work

This paper developed maximum likelihood techniques and simulation studies to estimate component reliability from masked failure data in series systems. The key results were:

- The likelihood model enabled rigorous inference from masked data via right-censoring and candidate sets.
- Despite masking and censoring, the MLE demonstrated accurate and robust performance in simulation studies.
- Bootstrap confidence intervals were reasonably well-calibrated, even for small samples.
- Estimation of shape parameters was more challenging than scale parameters.

Building on these findings, we propose the following promising areas for future work.

Relaxation of Masking Conditions Investigate relaxations of Conditions 1, 2, and 3. Condition 1 stipulates that the failed component is always in the candidate set,

$$\Pr\{K_i \in \mathcal{C}_i\} = 1.$$

Instead, we could model this as a probability, where the probability of the failed component being in the candidate set is a function of the failure time T_i and the component cause of failure K_i,

$$\Pr\{K_i \in \mathcal{C}_i | K_i = j, T_i = t_i\} = g(j, t_i).$$

Condition 2 stipulates that

$$\Pr\{\mathcal{C}_i = c_i | T_i = t_i, K_i = j\} = \Pr\{\mathcal{C}_i = c_i | T_i = t_i, K_i = j'\}$$

for all $j, j' \in c_i$. We call this an *uninformed* candidate set, since the conditional probability of the candidate set given the failure time and component cause of failure is independent of the component cause of failure.

We could relax this condition to allow for *informed* candidate sets, where the conditional probability of the candidate set given the failure time and component cause of failure is dependent on the component cause of failure.

In each of these violations or relaxations, we can either construct a new likelihood model that takes this relaxation into account, or we can use the existing likelihood model and assess the sensitivity of the estimator to this violation. A potentially interesting way to do the lattr is by using **KL-divergence** to measure the distance between, for instance, the uninformed and informed candidate set models, and then assess the sensitivity of the estimators to this distance.

Deviations from Well-Designed Systems In Section **??**, we assessed the sensitivity of the estimator to the masking probability p by fixing the sample size and right-censoring quantile and varying the masking probability p. We noted that since the series system was a well-designed system, the estimator was not very sensitive to the masking probability p.

In future work, we could assess the sensitivity of the estimator to deviations in system design. For instance, we could vary the masking probability and the shape or scale parameters of a component to see how the estimator behaves when the system deviates from the well-designed system. We did some preliminary investigation of this, and we found that the estimator was quite sensitive to deviations in system design, particularly when the masking probability was large. This suggests that the estimator is sensitive to deviations in system design, and so we should be careful when applying the estimator to real-world systems.

In the same vein, we could also assess the trade-off between using the homogenous shape parameter model and the full model. The homogenous shape parameter model assumes that the shape parameters are equal, which is a simplification of the full model. We could assess the sensitivity of the estimator to deviations in the homogenous shape parameter assumption. By the bias-variance trade-off, we expect that the homogenous shape parameter model will have lower variance, but higher bias than the full model, but if the assumption of homogenuity is reasonable, then the homogenous shape parameter model may be quite useful.

Semi-Parametric Bootstrap We used the non-parameteric bootstrap to construct confidence intervals, but we could also use the semi-parametric bootstrap. In the semi-parametric bootstrap, instead of resampling from the original data, we sample component lifetimes from the parametric distribution fitted to the original data and sample candidate sets from the empirical distribution of the conditional candidate sets in the original data. This is a compromise between the non-parametric bootstrap and the fully parametric bootstrap.[1]

Data Augmentation Assess the robustness of Data Augmentation (DA) as an implicit prior. For example, we may adopt the prior that the system is well-designed and augment particularly small samples with synthetic data sampled from a reduced model (with homogenous shape parameters) fitted to the original data.

Unlike a full Bayesian approach, where we would need to specify a prior for the parameters, DA is an implicit prior that need not be explicitly specified. It is a form of regularization that reduces the variance of the estimator by leveraging the structure of the model and the data.

Penalized Likelihood For Homogenous Shape Parameters Assess the use of penalized likelihood methods instead of DA as a form of regularization. For instance, we can add a penalty term to the log-likelihood function that penalizes the likelihood when the shape parameters are not close to each other. Instead of using a reduced model, we can use a penalized likelihood approach to encourage the shape parameters to be close to each other, but not necessarily equal.

General Likelihood Model with Predictors In this paper, we focused on a likelihood model that assumed Weibull components in a series configuration. We can extend this model by generalizing the hazard functions in two ways:

[1] The fully parametric bootstrap is not appropriate for our likelihood model because we do not assume a parametric form for the distribution of the candidate sets.

1. Let the hazard model be a function of predictors $w_1, \ldots, w_n$, where w_i is a vector of predictors for the ith observation. Then, the hazard function for the jth component is

$$h_j(t_i|w_i; \beta_j),$$

for instance we might make the shape and scale parameters of the Weibull component model be a function of the predictors w_i.

2. Replace the Weibull hazard function with a more general hazard function. For instance, in the Cox proportional hazards model [1], the hazard function for the jth component is given by

$$h_j(t_i|w_i; \beta_j) = h_0(t_i) \exp({\beta_j}^T w_i),$$

where $h_0(t_i)$ is a baseline hazard function shared by all components and β_j is the parameter vector for the jth component. A more general model would allow the component hazard functions to take any valid form, namely non-negative and integrable.

In either case, by the relation

$$R_j(t_i|w_i; \beta_j) = e^{-H_j(t)},$$

where

$$H_j(t_i) = \int_0^{t_i} h_j(u|w_i; \beta_j) du$$

is the cumulative hazard function for the j^{th} component, we can plug these component hazard and reliability functions into the likelihood contribution model in Theorem **??** to obtain a general likelihood model with predictors for the series system.

Assess the Calibration of Other Related Bootstrapped Statistics The calibration of the bootstrapped confidence intervals were evaluated and shown to be quite robust. We could do a similar analysis for other bootstrapped statistics. For instance, we could assess the bootstrapped 95% prediction interval for the probability that component j is the component cause of the next system failure given the data $\mathcal{D}_n$,

$$\Pr\{K_{n+1} = j|\mathcal{D}_n\}.$$

The current results provide a solid foundation for extensions like these that can further refine the methods and expand their applicability. By leveraging the rigorous likelihood framework and simulation techniques validated in this study, future work can continue advancing the capability for statistical learning from masked reliability data.

References

[1] D. R. Cox, *Regression models and life-tables*, Journal of the Royal Statistical Society: Series B (Methodological) **34** (1972), no. 2, 187–202.