Bootstrapping confidence intervals (BCa) of the maximum likelihood estimator of components in a series systems from masked failure data

Alex Towell

Abstract

We estimate the parameters of a series system with Weibull component lifetimes from relatively small samples consisting of right-censored system lifetimes and masked component cause of failure. Under a set of conditions that permit us to ignore how the component cause of failures are masked, we assess the bias and variance of the estimator. Then, we assess the accuracy of the boostrapped variance and calibration of the confidence intervals of the MLE under a variety of scenarios.

Contents

0.0.1 Conditional distribution of K_i given T_i and $\mathcal{C}_i$

This subsection is not necessary in our likelihood model, but it derives a useful result for making predictions about the component cause of failure.

Suppose we have observed a candidate set and a series system failure and we are interested in the probability that a particular component is the cause of failure.

Theorem 1. *Assuming Conditions **??** and **??**, the conditional probability of the component cause of failure is component j ($K_i = j$) given a masked component cause of failure ($\mathcal{C}_i = c_i$) and system lifetime ($T_i = t_i$) is given by*

$$\Pr\{K_i = j | T_i = t_i, \mathcal{C}_i = c_i\} = \frac{h_j(t_i|\boldsymbol{\theta_j})}{\sum_{l \in c_i} h_l(t_i|\boldsymbol{\theta_l})} 1_{\{j \in c_i\}}. \tag{0.1}$$

Proof. The conditional probability $\Pr\{K_i = j | T_i = t_i, \mathcal{C}_i = c_i\}$ may be written as

$$\Pr\{K_i = j | T_i = t_i, \mathcal{C}_i = c_i\} = \frac{\Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j, T_i = t_i\} f_{K_i,T_i}(j, t_i; \boldsymbol{\theta})}{\sum_{j=1}^{m} \Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j, T_i = t_i\} f_{K_i,T_i}(j, t_i; \boldsymbol{\theta})}.$$

By Theorem **??**, $f_{K_i,T_i}(j, t_i; \boldsymbol{\theta}) = h_j(t_i; \boldsymbol{\theta}) R_{T_i}(t_i; \boldsymbol{\theta})$. We may make this substitution and simplify:

$$\Pr\{K_i = j | T_i = t_i, \mathcal{C}_i = c_i\} = \frac{\Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j, T_i = t_i\} h_j(t_i; \boldsymbol{\theta_j})}{\sum_{j'=1}^{m} \Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j', T_i = t_i\} h_{j'}(t_i; \boldsymbol{\theta_{j'}})}.$$

Assuming Conditions **??** and **??**, we may rewrite the above as

$$\Pr\{K_i = j | T_i = t_i, \mathcal{C}_i = c_i\} = \frac{\Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j, T_i = t_i\} h_j(t_i; \boldsymbol{\theta_j})}{\Pr_{\boldsymbol{\theta}}\{\mathcal{C}_i = c_i | K_i = j, T_i = t_i\} \sum_{l \in c_i} h_l(t_i; \boldsymbol{\theta_l})} = \frac{h_j(t_i; \boldsymbol{\theta_j})}{\sum_{l \in c_i} h_l(t_i; \boldsymbol{\theta_l})}.$$

□

Frequently, we may not have any information at all about the component cause of failure. In this case, $c_i = \{1, \ldots, m\}$, and we obtain the following corollary.

Corollary 1. *The probability that the j^{th} component is the cause of system failure given only that we know a system failure occured at time t_i is given by*

$$\Pr\{K_i = j | T_i = t_i\} = \frac{h_j(t_i; \boldsymbol{\theta_j})}{\sum_{l=1}^{m} h_l(t_i; \boldsymbol{\theta_l})}.$$