Bootstrap Confidence Interval for the Steady State Availability

S. W. Shin[1], D. K. Kim[2*], J. H. Lim[3] and D. H. Park[1]
[1] Dept. of Statistics, Hallym University, Korea
[2] Dept. of Statistics, Chonbuk National University, Korea
[3] Division of Business Administration, Hanbat National University, Korea

Table of Contents

- Introduction
- Estimation of Availability
- Bootstrap Estimation of Availability
- The Simulation
- The Simulation Results

Introduction

- Availability [MIL-STD-721C(1981)]
  - *A measure of the degree to which an item is in the operable and committable state at the start of the mission, when the mission is called for at an unknown (random) time.*
  - $$A = \frac{MTBF}{MTBF + MTTR}$$

    where MTBF is interpreted as the expected value of each of the (i.i.d.) uptimes $X_1, X_2, \ldots$ *and correspondingly, MTTR is the expected value of each of the (i.i.d.) downtimes* $Y_1, Y_2, \ldots$. *MTBF stands for mean time between failures and MTTR stands for mean time to repair.*

Introduction

- Some Works on Constructing the Confidence Interval for the Steady State Availability
  - Thompson (1966)
  - Gary and Lewis (1967)
  - Masters and Lewis (1987)
  - Masters, Lewis and Kolarik (1992)
  - Chandrasekhar, Nataragjan and Sujatha (1994)

Introduction

- The Bootstrap Method
  - Effron (1979, 1982)
    - introduces and develops the nonparametric but computer intensive, estimation method called bootstrap.
  - Efron and Gong(1983), Efron and Tibshirani(1988) and Efron(1987)
    - the standard bootstrap confidence interval(SB)
    - the percentile bootstrap confidence interval(PB)
    - the bootstrap-t confidence interval(BT)
    - the bias-corrected and accelerated confidence interval(BCa).

Introduction

- The Purpose of the Paper
  - Constructing a confidence interval for the steady state availability by bootstrap method.
  - Bootstrap methods used are SB, PB and BCa.
  - Investigating the accuracy of the developed bootstrap confidence intervals by calculating the coverage probability and the average length of intervals.

Estimation of Availability

- Let $X_1, X_2, \ldots, X_n$ be a random sample from F(·) and let $Y_1, Y_2, \ldots, Y_n$ be a random sample from G(·). (Here (X_i, Y_i) represents failure time and repair time for the i-th operation period.) Then availability is defined as

$$A = \frac{MTBF}{MTBF + MTTR}$$

,where $MTBF = \int \overline{F}(t)dt$ and $MTTR = \int \overline{G}(t)dt$.

- The MLE of A

$$\hat{A}_{ML} = \frac{M\hat{T}BF_{ML}}{M\hat{T}BF_{ML} + M\hat{T}TR_{ML}}$$

, where $M\hat{T}TR_{ML}$ and $M\hat{T}BF_{ML}$ are MLE of MTBF and MTTR, respectively

Estimation of Availability

- Confidence Limits for Availability when F and G are exponential.
  - Suppose that F and G are exponential distributions with means of 1/ λ and 1/μ, respectively.

→ $A = \frac{\mu}{\lambda + \mu} = \frac{1}{1+\theta}$ and $\hat{A}_{ML} = \frac{\overline{X}}{\overline{X} + \overline{Y}} = \frac{1}{1+\hat{\theta}}$, where $\theta = \frac{\lambda}{\mu}$ and $\hat{\theta} = \frac{\overline{Y}}{\overline{X}}$

  - $2n\lambda\overline{X}$ and $2n\mu\overline{Y}$ are χ^2-distributions with 2n degrees of freedom.

→ $F^* = \frac{2n\lambda\overline{X}}{2n} / \frac{2n\mu\overline{Y}}{2n} = \frac{\theta}{\hat{\theta}}$ becomes F-distribution with 2n and 2n degrees of freedom.

Estimation of Availability

- 100(1-α)% Confidence Limits for A

$$1-\alpha/2 = P(F^* \ge F_{\alpha/2}(2n,2n)) = P(\frac{\theta}{\hat{\theta}} \ge F_{\alpha/2}(2n,2n)) = P(\theta \ge \hat{\theta}F_{\alpha/2}(2n,2n))$$

$$= P(\frac{1}{1+\theta} \le \frac{1}{1+\hat{\theta}F_{\alpha/2}(2n,2n)}) = P(A \le \frac{1}{1+\hat{\theta}F_{\alpha/2}(2n,2n)})$$

- The upper limit of 100(1-α) confidence interval for A

$$\frac{1}{1+\hat{\theta}F_{\alpha/2}(2n,2n)}$$

- Analogously, the lower limit of 100(1-α) confidence interval for A

$$\frac{1}{1+\hat{\theta}/F_{\alpha/2}(2n,2n)}$$

Bootstrap Estimation of Availability

Let $\mathbf{X}_i^*$ be a bootstrap resample of size m from $X_1, X_2, \dots, X_n$ the original sample and let $\mathbf{Y}_i^*$ be a bootstrap sample of size m from $Y_1, Y_2, \dots, Y_n$, where i=1, 2, ⋯, B. Note that $m \le n$.

- Estimator of A based on the bootstrap resamples $\mathbf{X}_i^*$ and $\mathbf{Y}_i^*$

$$\hat{A}_i^* = \frac{\overline{X}_i^*}{\overline{X}_i^* + \overline{Y}_i^*} \quad \text{for } i = 1, 2, \dots, B.$$

- Bootstrap estimator of A

$$\hat{A}_B = \sum_{i=1}^{B} \hat{A}_i^* / B$$

- Standard deviation of the estimator of A

$$sd(\hat{A}_B) = \left\{ \sum_{i=1}^{B} (\hat{A}_i^* - \hat{A}_B)^2 / (B-1) \right\}^{\frac{1}{2}}$$

Bootstrap Estimation of Availability

- Bootstrap Confidence Interval of Availability
  - Standard Bootstrap(SB) Confidence Interval
    - If the distribution of is approximately normal, the 100(1-2α)% SB confidence interval for A is $\hat{A} \pm Z_{\alpha} \cdot Sd(\hat{A}_B)$, where Z_α is the upper α–th quantile of the of standard normal distribution.
  - Percentile Bootstrap(PB) Confidence Interval
    - Let $\hat{A}^*(i)$ be the i-th ordered bootstrap estimator among $\hat{A}_1^*, \cdots, \hat{A}_B^*$. Then 100(1-2α)% PB confidence interval for A is given by

$$\left[\hat{A}^*(\alpha B), \hat{A}^*((1-\alpha)B)\right]$$

Bootstrap Estimation of Availability

- Bootstrap Confidence Interval of Availability
  - BCa Confidence Interval

Let $\hat{z}_0 = \Phi^{-1}(\#(\hat{A}_i^* < \hat{A})/B)$, where $\Phi^{-1}(\cdot)$ is the inverse function of $\Phi(\cdot)$ and $\Phi(\cdot)$ is the c.d.f. of standard normal distribution and let $\mathbf{X}(i)$ and $\mathbf{Y}(i)$ be the original samples with the i th point $x(i)$ and $y(i)$ deleted. And let $\hat{A}(i)$ be the estimator of A based on the samples $\mathbf{X}(i)$ and $\mathbf{Y}(i)$ and let $\hat{A}(\cdot) = \sum_{i=1}^{n} \hat{A}(i)/n$. Then the acceleration is defined

$$\hat{a} = \sum_{i=1}^{n} (\hat{A}(\cdot) - \hat{A}(i))^3 / 6 \left\{ \sum_{i=1}^{n} (\hat{A}(\cdot) - \hat{A}(i))^2 \right\}^{\frac{3}{2}}$$

Bootstrap Estimation of Availability

- Bootstrap Confidence Interval of Availability
  - BCa Confidence Interval

    Note that $\hat{z}_o$ and $\hat{a}$ are called as bias-correction and acceleration, respectively. Then the 100(1-2α)% BCa confidence interval is given by

$$(\hat{A}^*(\alpha_1), \hat{A}^*(\alpha_2))$$

, where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + Z_\alpha}{1 - \hat{a}(\hat{z}_0 + Z_\alpha)}\right) \qquad \alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + Z_{1-\alpha}}{1 - \hat{a}(\hat{z}_0 + Z_{1-\alpha})}\right)$$

  - Note that $\hat{z}_0$ measures median bias of $\hat{A}_i^*$'s, that is the discrepancy between median of $\hat{A}_i^*$'s and $\hat{A}$, in normal units

The Simulation

- Performance Comparison of Bootstrap Confidence Intervals
  - The coverage probability and the length of the interval
  - The failure time distribution : $\bar{F}(t) = Exp[-t^\alpha / \theta]$
  - The repair time distribution : $\bar{G}(t) = Exp[-t]$
  - Values of α = 2, 1 and 0.5 so that failure time distribution becomes IFR, CFR and DFR, respectively.
  - For a given value of α, values of θ are chosen so that the ratios of MTBF to MTTR are 10:1, 50:1 and 90:1.
  - Five values of availability used in simulation study.

MTBF : MTTR	10:1	50:1	90:1
Availability	0.9091	0.9804	0.9890

The Simulation

- Performance Comparison of Bootstrap Confidence Intervals
  - Size of random number = 10
  - B(bootstrap resamples) = 1000 (each of size 10)
  - A 90% confidence interval is constructed by each of 3 methods.
  - N(the number of replications) =1000
  - Coverage probability = the proportion of times each of 90% confidence interval contains the true value of availability
  - The average length of intervals and corresponding standard deviation.
  - All simulations are run on PC with random number generation accomplished using SAS

The Simulation Results

- Coverage probability for the 90% bootstrap confidence intervals

		10:1	50:1	90:1
DFR	SB	0.960	0.982	0.985
	PB	0.889	0.889	0.889
	BCa	0.795	0.727	0.715
Exponential(CFR)	SB	0.869	1.000	0.89
	PB	0.853	1.000	0.861
	BCa	0.842	1.000	0.842
IFR	SB	0.970	0.968	0.854
	PB	0.973	0.973	0.849
	BCa	0.980	0.983	0.850

The Simulation Results

- Note on Coverage Probability
  - It is noted from the table that, for DFR case, as the ratios of MTBF to MTTR increases, the coverage probability increases for SB confidence interval while it decreases for BCa confidence interval.
  - For CFR case, the coverage probability shows its trend change for all confidence interval as the ratios of MTBF to MTTR increases.
  - For IFR case, the coverage probability tends to decrease for all confidence interval as the ratios of MTBF to MTTR increases.

The Simulation Results

- Average length and standard deviation of the 90% bootstrap confidence intervals

		10:1	50:1	90:1
DFR	SB	0.263 (0.100)	0.087 (0.047)	0.053 (0.032)
	PB	0.242 (0.094)	0.069 (0.037)	0.040 (0.023)
	BCa	0.352 (0.157)	0.151 (0.122)	0.098 (0.095)
Exponential (CFR)	SB	0.129 (0.054)	0.030 (0.011)	0.018 (0.010)
	PB	0.124 (0.052)	0.028 (0.009)	0.017 (0.009)
	BCa	0.141 (0.062)	0.035 (0.016)	0.021 (0.012)
IFR	SB	0.087 0.028	0.020 (0.007)	0.012 (0.005)
	PB	0.086 (0.027)	0.020 (0.006)	0.012 (0.004)
	BCa	0.089 (0.029)	0.021 (0.007)	0.012 (0.005)

The Simulation Results

- Note on average length
  - As expected from the theory, PB confidence interval has the smallest average length (and standard deviation of the length) for all types of failure time distributions. And SB confidence interval has a little longer average length than PB interval.
  - BCa confidence interval is somewhat relatively longer average length than SB and PB confidence interval.

Conclusion & Further Studies

- Conclusions
  - We construct several Bootstrap confidence intervals for the steady state availability. Bootstrap methods used are SB, PB and BCa.
  - The Bootstrap confidence intervals are not good as much as we expected in the sense of the coverage probability. This results seem to be caused by the fact that the exponential and Weibull distribution are skewed.
- Further Studies
  - We will further investigate the reason of such results.
  - We will find other Bootstrap confidence intervals which are suitable for skewed distributions.