소표본인 경우 신뢰성 순위 척도의 고찰
- Overview of Reliability Rank Measures for Small Sample -

최 성 운 *
Choi Sung Woon

Abstract

This paper presents three methods for expression of reliability measures for large and small data. First method is to express parametric estimation of cardinal reliability measure data for large sample, which requires numerous sample. Second is to obtain nonparametric distribution classification of ordinal reliability measure data for small sample. However it is difficult for field user to understand this method. Last method is to acquire parametric estimation of ordinal reliability measure data for small data. Because this method requires small sample and is comprehensive, we recommend this one among the proposed methods. Various reliability rank measures are presented.
Keywords : Reliability Measures, Parametric, Nonparametric, Cardinal, Ordinal, Small and Large Sample

1. 서 론

다양한 고객의 사용조건하에서 동적인 신뢰성(Reliability) 데이터를 구하는 것은 시간과 비용의 관점에서 쉬운 일이 아니다. 정적인 관점에서의 품질(Quality)과 같이 수명 데이터를 구하기가 용이한 대표본(Large Sample)의 신뢰성 기수(Cardinal) 데이터인 경우 구간별 신뢰성 척도의 연관 공식에 의해 쉽게 구할 수 있다. 분포를 가정한 경우 신뢰성 척도는 분포의 미분, 적분 관계에 의해 연계 공식으로 표현되며, 분포를 모르는 경우 도수분포표의 구간데이터에 의한 신뢰성 척도 공식에 의하여 쉽게 구할

* 경원대학교 산업공학과 교수
2007년 3월 접수; 2007년 4월 수정본 접수; 2007년 4월 게재확정

수 있으며 분포의 χ^2 및 Bartlett 적합도 검정을 실시한다. 대표본의 경우 사용되는 신뢰성 척도로는 고장 pdf(Probability Density Function) $f(t)$, 고장 cdf(Cumulative Distribution Function, Unreliability Function) $F(t)$, 신뢰도 함수(Reliability Function) $R(t)$, 위험률 함수(Hazard Rate, Instantaneous Failure Rate) $\lambda(t)$등이 있으며 분포의 모수적(Parametric) 추정방법을 사용한다.[1.5]

그러나 제품이나 시스템의 개발, 생산, 사용, 폐기등의 전생애과정에서의 완전한 수명 데이터(Complete Data)를 구하는 것은 비효율적인 경우가 많아 가혹한 조건에서 시간을 단축하기 위한 가속수명 시험(ALT : Accelerated Life Test)이나 정시 중도중단시험(Type I Censored Test),정수 중도중단시험(Type II Censored Test), 랜덤 중도중단시험(Randomly Censored Test) 등을 수행한다.

이와같은 소표본(Small Sample)인 경우 서수(Ordinal) 데이터를 사용하여 순위(Rank)형 신뢰성 척도를 구하고 분포의 비모수적(Nonparametric) 분류와 비모수적 적합도 검정을 수행한다. 소표본, 순위형 신뢰성 척도 데이터인 경우 분포의 비모수적 분류에서 사용되는 신뢰성 척도는 위험률 함수 $\lambda(t)$, 누적 위험률 함수(Cumulative Hazard Rate Function) $H(t)$, 평균 잔여 수명(Mean Residual Life) $m(t)$, 신뢰도 함수 $R(t)$ 등이 있다. [2] 그러나 비모수적 분류방식은 신뢰성 척도함수들의 단조성 또는 기타 성질에 의하여 몇 개의 집단으로 분류하는 방법으로 현장의 신뢰성 실무자들이 이해하기에는 어려운 점이 있으며 객관성과 명확성에 있어서도 분포를 이용하는 모수적 방법에 비해 문제점이 많다.

따라서 본 연구에서는 소표본 분포의 모수적 추정방식을 위한 다양한 순위형 신뢰성 척도의 연관공식을 고찰한다. 지수분포, 정규분포, 와이블 분포 등의 신뢰성 분포의 모수를 추정하기 위해 LSE(Least Square Estimator), MLE(Maximum Likelihood Estimator), MM(Moment Method), 평균법 등의 대수적인 방법을 사용할 수도 있으나 확률용지(Probability Paper)에 의한 직선의 기하학적인 방법을 이용하는 것이 실무자에게 용이할 수 있다.

직선은 분포의 적합성 여부를 판단하는 기준으로 콜모고로프(Kolmogorov) 적합도 검정과 같은 효과를 지니면서도 동시에 효율적인 방법이다. 따라서 실무에서는 확률용지를 이용한 방법을 적극 권장하며 이 경우 사용되는 것이 순위형 신뢰성 척도 데이터이며 주로 고장 cdf $F(t_i)$가 사용된다.

2장에서는 대표본, 기수형 신뢰성 척도 데이터인 경우 분포의 모수적 추정 및 χ^2 적합도 검정방법을 소개하고 3장에서는 소표본, 순위형 신뢰성 척도 데이터인 경우 $\lambda(t)$, $H(t)$, $m(t)$, $R(t)$에 의한 분포의 비모수적 분류 방식 및 비모수적 적합도 검정 방법을 나타낸다. 4장에서는 소표본 분포의 모수적 추정을 수행하는 경우 다양한 순위형 신뢰성 척도의 연관공식 등을 제안하며 5장에서 결론을 맺는다.

2. 대표본, 기수(Cardinal)형 신뢰성 척도 데이터, 분포의 모수적 (Parametric) 추정

2.1 기수형 신뢰성 척도

확률분포 즉 고장 pdf $f(t)$를 알고 있는(Known) 경우 고장 cdf $F(t)=\int_0^t f(t)dt$, 신뢰도 함수 $R(t)=\int_t^\infty f(t)dt=1-F(t)$, 위험률 함수 $\lambda(t)=\frac{f(t)}{R(t)}$ 등을 유도할 수 있다. 신뢰도 척도의 미분, 적분관계에 의해 $f(t)=\frac{dF(t)}{dt}=-\frac{dR(t)}{dt}$, $R(t)=e^{-\int_0^t \lambda(t)dt}=e^{-H(t)}$ 로 표시되며 $H(t)=\int_0^t \lambda(t)dt$ 로서 누적 위험률 함수(Cumulative Hazard Rate)라고 한다. $f(t)$는 전구간에서 t시간째 고장나는 확률(상대도수)이며, $\lambda(t)$ 는 특정시간 또는 특정구간에서 생존한 것 중에서 t시간째 고장나는 조건부 확률로서 비율을 나타낸다. 신뢰성 데이터의 경제성 관점에서 $f(t)$보다는 $\lambda(t)$를 많이 사용한다. $F(t), R(t)$는 t시간까지 각각 고장, 생존하는 확률이다.

확률분포 즉 고장 pdf $f(t)$를 모르는(Unknown) 경우 대표본(Large Sample)의 신뢰성 데이터에 대한 신뢰성 척도의 공식은 다음과 같다. n은 전체 샘플 데이터의 수이며 $n(t)$는 t시간까지의 생존수, Δt는 신뢰성 시험 시간의 증분값을 나타낸다.

$$R(t)=\frac{n(t)}{n}$$

$$F(t)=\frac{n-n(t)}{n}$$

$$f(t)=\frac{n(t)-n(t+\Delta t)}{n\cdot\Delta t}$$

$$\lambda(t)=\frac{n(t)-n(t+\Delta t)}{n(t)\cdot\Delta t}$$

2.2 분포의 모수적 추정 및 적합도 검정

지수분포인 경우 고장률 또는 평균수명(MTBF : Mean Time Between Failure, MTTF : Mean Time To Failure)의 모수(Parameter)를 완전자료(Complete Data), 정시 중도중단자료(Type I Censored Data), 정수 중도중단자료(Type II Censored Data) 등의 평균법으로 점추정, 구간추정을 구한다. 정규분포인 경우 MLE, LSE등으

로 μ, σ의 모수를 추정하고 와이블 분포인 경우 LSE, MM등으로 위치모수 $r=0$로 놓고 형상모수 m과 척도모수 η 등을 구한다. 그러나 정규분포, 와이블분포에서 MSE, LSE, MM등의 대수적인 방법으로 모수를 추정할 경우 복잡한 계산식을 간단한 확률용지의 그래프로 사용하는 것이 신뢰성 실무자에게 바람직하다. 가정된 분포의 적합도 여부를 알고자 하는 경우 χ^2분할표 검정 및 Bartlett적합도 검정을 수행한다. [3.4]

3. 소표본, 순위형(Rank) 신뢰성 척도 데이터, 분포의 비모수적 (Nonparametric) 분류 방법

3.1 분포의 비모수적 분류

분포의 비모수적 분류 방법은 $\lambda(t)$, $H(t)$, $m(t)$, $R(t)$ 등의 신뢰성 척도 함수의 단조성 또는 기타 성질에 의하여 분류하는 방법이다. 분포의 비모수적 분류된 집단간의 성질은 화살표에 따라 좌변이 우변이 되기 위한 충분조건을 만족한다.[2]

성질 1 : IFR(Increasing Failure Rate : $\lambda(t)$) → IFRA(Increasing Failure Rate Average
: $H(t)$) → NBU(New Better than Used : $R(t)$) → NBUE(New Better than Used in Expectation : $m(t)$) → HNBUE(Harmonic New Better than Used in Expectation : $m(t)$)

성질 2 : DFR(Decreasing Failure Rate : $\lambda(t)$) → DFRA(Decreasing Failure Rate Average : $H(t)$) → NWU(New Worse than Used : $R(t)$) → NWUE(New Worse than Used in Expectation : $m(t)$) → HNWUE(Harmonic New Worse than Used in Expectation : $m(t)$)

성질 3 : IFR(Increasing Failure Rate : $\lambda(t)$) → DMRL(Decreasing Mean Residual Life
: $m(t)$) → NBUE(New Better than Used in Expectation : $m(t)$)

성질 4 : DFR(Decreasing Failure Rate : $\lambda(t)$) → IMRL(Increasing Mean Residual Life
: $m(t)$) → NWUE(New Worse than Used in Expectation : $m(t)$)

3.2 분포의 비모수적 적합도 검정

3.1절에서 분포의 비모수적 분류된 집단이 귀무가설이 지수분포인가를 검정하는 방법으로 IFR, IFRA, NBU,NBUE, NBU-t_0, HNBUE, DMRL, NBUE등의 검정이 있으며 DFR, DFRA, NWU, NWUE, NWU-t_0, IMRL, NWUE등의 검정은 반대 방향의 단조성을 가정한다.[2]

4. 소표본, 순위(Rank)형 신뢰성 척도 데이터, 분포의 모수적 (Parametric) 추정

4.1 근사 신뢰성 순위 척도 및 적합도 검정

메디안 순위(MR : Median Rank)를 정확히 구하는 함수는 불완비(Incomplete) 베타 분포 $B_x(a,b)$로서 $a=i$, $b=n-i+1$, $B=0.5$인 $x=MR$을 유도하면 된다. 유도식은

$B=\frac{n!}{(i-1)!(n-1)!}\int_0^{MR} t^{i-1}(1-t)^{n-i}dt$이며 계산식은

$B'=\frac{n!}{(i-1)!(n-1)!}MR^{i-1}(1-MR)^{n-i}$이며 $MR=\frac{B-0.5}{B'}$, $B'=\frac{dB}{dt}$이 된다.[7.9]

메디안 순위(MR)를 구하는 또 다른 방법은 $\sum_{k=1}^{n}\binom{n}{k}MR^{k}(1-MR)^{n-k}=0.5=50\%$를 만족하는 MR을 구하면 된다. F분포를 이용하여 메디안 순위(MR)를 구하면

$MR=\frac{1}{1+\frac{n-i+1}{i}F(2(n-i+1),2i:0.50)}$이다.[13]

F분포를 이용하여 퍼센타일 P를 구하는 하한 순위(LR : Lower Rank)

$LR_P=\frac{1}{1+\frac{n-i+1}{i}F(2(n-i+1),2i:1-P)}$이며 상한 순위(UR : Upper Rank)

$UR_p=\frac{1}{1+\frac{n-i+1}{i}F(2(n-i+1),2i:P)}$이다. 메디안 순위(MR)는 n ≤20일 경우

$MR=\frac{i-0.30685-0.3863\frac{i-1}{n-1}}{n}$이며 n ≥20 일 경우 $MR=1-2^{-\frac{1}{n}}+(\frac{i-1}{n-1})[2^{(1-\frac{1}{n})}-1]$이다. [10]

평균순위 (Mean Rank) 근사법은 다음의 기대값을 사용하여 구한다.

E[F(i,n)] $\equiv$ E(i,n) $=(-1)^{i+1}\binom{n}{i-1}\sum_{j=0}^{i-1}\binom{i-1}{j}(-1)^{i-j-1}\{-r-\ln(n-i+j+1)\}+$E($i$ -1, n), $i \geq 2$ 이며 E(i,n) = -r-lnη 으로 r은 오일러 상수(Euler Constant)이다. [6]

4.2 $F(t_i)=\frac{i-c}{n-2c+1}$ 근사 추정 방법

$F(t_i)=\frac{i-c}{n-2c+1}$, $1<i<n$이며 $i=1$일 때 $F(t_i)=1-0.5^{\frac{1}{n}}$이며 $i=n$일 때 $F=0.5^{\frac{1}{n}}$이다. IEC 56 추정량은 $F(t_i)=\frac{i-0.5}{n+0.25}$ (c는 불일치)이며 평균 순위(Mean Rank, Weibull et al. Estimator, Herd-Johnson Estimator, Expected Rank)는 $F(t_i)=\frac{i}{n+1}$ ($c=0$를 대입)이다. 모드 순위(Mode Rank, Hazen Estimator, Modified Kaplan-Meier Estimator, Symmetrical Rank)는 $F(t_i)=\frac{i-0.5}{n}$ ($c=0.5$를 대입)이며 메디안 순위(Median Rank, Bernard et al. Estimator, Bos-Levenbach Estimator)는 $F(t_i)=\frac{i-0.5}{n+0.4}$ ($c=0.5$를 대입)이다.

Filliben 추정량은 $F(t_i)=\frac{i-0.3175}{n+0.365}$ ($c=0.3175$를 대입)이며 Jacquelin 추정량은 $F(t_i)=\frac{i-c}{n-2c+1}$($c$는변수)로서 $c=c_i+c_{n-i+1}-0.2784$, $c_k=(1-(k+1)(0.5)^{\frac{1}{k}})/(1-2(0.5)^{\frac{1}{k}})$ 이다. 위의 6가지 추정량 중 편차가 작은 순서는 Jacquelin 추정량, Filliben 추정량, 메디안 순위, 모드 순위, 평균 순위, IEC 56 추정량 순이다.[8,9,11,12]

평균 순위는 정규분포일 경우 사용되며 왜곡된 분포(Skewed Distribution)일 경우 메디안 순위를 사용한다.

IEC 56 추정량의 $R(t_i)=\frac{n-i+0.75}{n+0.25}$, $f(t_i)=\frac{1}{(n+0.25)(t_{i+1}-t_i)}$, $\lambda(t_i)=\frac{1}{(n-i+0.75)(t_{i+1}-t_i)}$이다. 평균 순위의 $R(t_i)=\frac{n-i+1}{n+1}$, $f(t_i)=\frac{1}{(n+1)(t_{i+1}-t_i)}$, $\lambda(t_i)=\frac{1}{(n-i+1)(t_{i+1}-t_i)}$이다. 모드 순위의 $R(t_i)=\frac{n-i+0.5}{n}$, $f(t_i)=\frac{1}{n(t_{i+1}-t_i)}$, $\lambda(t_i)=\frac{1}{(n-i+0.5)(t_{i+1}-t_i)}$ 이다.

메디안 순위 $R(t_i)=\frac{n-i+0.7}{n+0.4}$, $f(t_i)=\frac{1}{(n+0.4)(t_{i+1}-t_i)}$, $\lambda(t_i)=\frac{1}{(n-i+0.7)(t_{i+1}-t_i)}$이다.

Filliben 추정량 $R(t_i)=\frac{n-i+0.6825}{n+0.365}$, $f(t_i)=\frac{1}{(n+0.365)(t_{i+1}-t_i)}$,

$\lambda(t_i)=\frac{1}{(n-i+0.6825)(t_{i+1}-t_i)}$ 이다. Kaplan-Meyer 추정량(Empirical Distribution : 경험순위)은 $F(t_i)=\frac{i}{n}$ 이며 $R(t_i)=\frac{n-i}{n}$, $f(t_i)=\frac{1}{n(t_{i+1}-t_i)}$, $\lambda(t_i)=\frac{1}{(n-i)(t_{i+1}-t_i)}$ 이며 기타 추정량으로 $F(t_i)=\frac{0.5^{\frac{1}{2}}(2i-n+1)+(n-1)}{n-1}$ 이 있다.[3]

4.3 $F(t_i)=\frac{i+A}{n+2B+1}$ 근사 추정 방법

$F(t_i)$에서 $B=A$, $A=0$인 경우 평균 순위 $F(t_i)=\frac{i}{n+1}$가 되며 $A=-0.5$인 경우 모드 순위 $F(t_i)=\frac{i-0.5}{n}$가 된다. $F(t_i)$에서 $B=A$, $A=-0.3175$인 경우 Filliben 추정량 $F(t_i)=\frac{i-0.3175}{n+0.365}$가 되며 $A=-0.5$, $B=-0.375$인 경우 IEC 56 추정량 $F(t_i)=\frac{i-0.5}{n+0.25}$가 된다.[6]

4.4 PLE(Product Limit Estimator) 근사 추정 방법

랜덤 중도중단자료(Randomly Censored Data)인 경우 PLE(Product Limit Estimator)추정량으로 Kaplan-Meyer 추정량 $F(t_i)=1-\prod_{j=1}^{i}(\frac{n-j}{n+1-j})^{\delta_{(j)}}$, Nelson 추정량 $F(t_i)=1-\exp[-\sum_{j=1}^{i}(\frac{\delta_{(j)}}{n-1-j})]$, 순위법(Rank Method)으로 $F(t_i)=1-\prod_{j=1}^{i}(\frac{n_j+1-r_j}{n_j+1})$ 이며 $\delta_{(j)}$는 t_j가 고장시간일 경우는 1을 대입하고 시험 중단시간일 경우 0을 대입하며 n_j, r_j는 j시점에서 생존 개수, 고장 개수이다.[2]

5. 결 론

본 연구에서는 대표본, 소표본인 경우 기수, 서수 신뢰성 척도 데이터의 모수적, 비모수적 추정 방법 세가지를 제시하였다. 첫번째 방법은 대표본인 경우 기수형 신뢰성 척도 데이터에 대한 분포의 모수적 추정방법으로 도수분포표, 히스토그램을 이용하여

손쉽게 작성할 수 있으나 많은 데이터를 필요로 하는 단점을 가지고 있다. 두번째 방법은 소표본인 경우 순위형 신뢰성 척도 데이터에 대한 비모수적 분류 방법으로 척도 함수의 단조성 또는 기타 성질등의 분류방법이 실무자에게는 이해하기 쉽지 않은 면이 있다. 세번째 방법으로 소표본인 경우 순위형 신뢰성 척도 데이터에 대한 분포의 모수적 추정방법으로 메디안 순위, 모드 순위, 평균, 경험순위 등 다양한 추정량을 구할 수 있는 근사식, 일반식과 랜덤 임의 중도절단자료등에 적용되는 Kaplan-Meyer추정량 등을 소개하였다. 이 방법은 데이터면에서 경제적이고 실무자들에게 이해하기 쉽기 때문에 세가지 제안된 방법 중 권장되는 방법이다.

6. 참 고 문 헌

[1] 박경수, 신뢰도 및 보전 공학, 영지문화사, 1999.
[2] 박동호, 수명 분포 개념과 응용, 영지문화사, 2006.
[3] 이상용, 신뢰성공학, 형설출판사, 2003.
[4] 정해성 외, 신뢰성 시험 분석 평가, 영지출판사, 2005.
[5] 최성운, "신뢰성 척도 및 분포의 적용", 대한안전경영과학회지, 7(5)(2005) 175-184.
[6] Cacciari M., Montanari G.C., "Discussion on Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper", IEEE Transactions on Electrical Insulation, 26(6)(1991)1224-1229.
[7] Fothergill J.C., "Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper," IEEE Transactions on Electrical Insulation, 25(3)(1990)489-492.
[8] Jacquelin J., "Generalization of the Method of Maximum Likelihood", IEEE Transactions on Electrical Insulation, 28(1)(1993)65-72.
[9] Jacquelin J., " A Reliable Algorithm for the Exact Median Rank Function", IEEE Transactions on Electrical Insulation, 28(2)(1993)168-171.
[10] Keccioglu D., Reliability and Life Testing Handbook, PTR Prentice Hall Inc., 1993.
[11] Shimokawa T.,Liao M., "Goodness-of-Fit Tests for Type-I Extreme -Value and 2-Paramenter Weibull Distributions", IEEE Transactions on Reliability, 48(1) (1999) 79-86.
[12] http://www.minitab.com
[13] http://www.weibull.com

저 자 소 개

최 성 운 : 현 경원대학교 산업공학과 교수. 한양 대학교 산업공학과에서 공학사, 공학석사, 공학박사 학위를 취득하고, 1994년 한국과학재단 지원으로 University of Minnesota에서 1년간 Post-Doc을 수행하였으며, 2002년부터 1년 8개월 동안 University of Washington에서 Visiting Professor를 역임하였음. 주요 관심분야는 경영품질시스템, 서비스 사이언스, 자동화 생산 및 장치 산업에서의 품질관리이며, 컴퓨터·정보통신시스템의 신뢰성 설계 및 분석, RFID 시스템에도 관심을 가지고 있음. swchoi@kyungwon.ac.kr

저 자 주 소

최 성 운 : 경기도 성남시 수정구 복정동 산65번지 경원대학교 산업정보시스템공학과