중도절단된 생존함수의 신뢰구간 비교연구

이경화[1)], 이재원[2)]

요 약

중도절단된 자료와 표본수가 적은 자료를 가지는 생존분석에서 생존율을 추정하거나 두 집단의 생존율을 비교할 때 정규분포 근사를 가정한 신뢰구간을 이용하는 데는 많은 어려움이 생긴다. 생존함수의 신뢰구간에 대한 중도절단율, 표본의 크기에 따른 다양한 상황의 모의실험을 통하여 Kaplan-Meier, Nelson, 적률 추정량 그리고 cox model의 β을 가지고 붓스트랩을 이용한 신뢰구간과 비모수 신뢰구간, 우도비 신뢰구간의 실제 포함 확률을 비교해보고자 한다.

주요용어 : 중도절단, Kaplan-Meier 추정량, Nelson 추정량, 적률 추정량, Cox regression model

1. 서론

중도 절단된 자료와 표본수가 적은 자료를 가지는 생존분석에서 생존율을 추정하거나 두 집단의 생존율을 비교할 때 정규분포 근사를 가정한 신뢰구간을 이용하는 것 데 많은 어려움이 생긴다. 따라서 생존함수 신뢰구간의 많은 연구가 시도 되었다. 그 중 Kani Chen 와 Shaw Hwa LO(1996)은 붓스트랩 신뢰구간과 정규 신뢰구간을 Kaplan-Meier, Nelson, 적률 추정량에 대해 포함확률을 비교하였고, David R. Tomas와 Gray L. Grunkemeier(1975)는 Kaplan-Meier 추정량에 대해 우도비 신뢰구간과 정규 근사에 근거한 신뢰구간에 대해 포함확률을 비교하였다. Ron Brookmeyer 와 John Crowley(1982)는 중간 값에 대한 비모수적 신뢰구간을 정규 신뢰구간과 비교하였다.

기존의 연구를 살펴보면 부분적인 조건하에서 신뢰구간 구축방법을 연구하거나, 연구된 신뢰구간과 정규 신뢰구간의 비교가 대부분이었다. 이에 본 논문에서는 앞에서 언급한 중도절단된 생존함수의 신뢰구간을 비교하기 위한 여러 가지 방법을 소개하고, 여러 가지 상황 하에서 모의실험을 실시하여 이들 신뢰구간의 포함확률을 비교하였다.

section 2에서 생존함수의 정의와 신뢰구간을 소개하고 section 3에서는 모의실험 구조 및 결과에 대해서 살펴본다. section 4에서 결론을 내린다.

2. 생존함수와 신뢰구간

생존 함수에 대해 살펴보면, 생존시간 $T_1, T_2, \ldots\ldots, T_n$은 서로 독립이고 동일한 분포 F를 갖는 확률 변수이며, 중도절단시간 $C_1, C_2, \ldots., C_n$은 역시 서로 독립이고 동일한 분포 G를 갖는다고 하자. 우리는 실험으로부터 $T_1, T_2, \ldots\ldots, T_n$을 모두 관측하지 못하고, $X_i = min(T_i, C_i)$ $\delta_i = \begin{cases} 1, & T_i \quad \text{관측}(uncensored) \\ 0 & T_i \quad \text{중도절단}(censored) \end{cases}$ 로 정의되는

1) (주) 씨엔알 리서치 통계분석담당
2) 고려대학교 통계학과 교수

$(X_1, \delta_1), (X_2, \delta_2), \ldots\ldots (X_n, \delta_n)$ 을 대신 관측하게 된다.

이러한 임의 중도절단의 경우 일반적으로 생존시간 T_i와 C_i가 서로 독립이라는 가정을 해야 한다. 생존분석에서 필요한 기초적인 함수들을 정의하면, 생존시간을 $T(\geq 0)$라하고 T가 확률밀도함수 f(t)와 분포함수 F(t)를 갖는다고 하자. T의 생존함수 S(t)와 위험함수 h(t)는 각각

$$S(t) = P(T > t) = 1 - F(t),$$

$$h(t) = \lim_{dt \to 0} P(t < T < t + dt \mid T \geq t)/dt = f(t)/S(t)$$

로 정의된다. 즉, 생존함수는 환자가 t시점까지는 생존했다가 t시점 바로 직후에 사망하게 되는 순간위험률을 의미한다.

본 논문에서 모의실험 시 쓰인 신뢰구간은 총 7가지로 정규 근사에 의한 신뢰구간(①), 일반적인 붓스트랩 신뢰구간(②), 백분위 신뢰구간(③), Hybrid 신뢰구간(④), BCa 신뢰구간(⑤), 우도비 신뢰구간(⑥), 비모수 신뢰구간(⑦)이다.

가장 일반적인 정규근사에 의한 신뢰구간 표준정규분포의 $100\% \times (\alpha)$백분위수를 z_α라고하면, θ에 대한 $100\% \times (1-\alpha)$ 근사 신뢰구간은

$$(\hat{\theta} - z_{\alpha/2}\hat{\sigma} \quad , \quad \hat{\theta} - z_{(1-\alpha/2)}\hat{\sigma})$$ (단, $z_\alpha = \Phi^{-1}(\alpha)$)이 된다. -①

$\hat{\theta}^*$을 붓스트랩 표본으로부터 얻어진 추정량이라고 하면, $\hat{\theta}^*$의 붓스트랩 분포를 K라고 표기한다. 이런 가정을 통해 일반적인 붓스트랩 신뢰구간은 B개의 붓스트랩 표본 각각에 대하여 $T^* = (\hat{\theta}^* - \hat{\theta})/\hat{\sigma}^*$를 구한 후 T^*의 $100 \times \alpha$백분위수를 $\sum I\{T^* \leq t_\alpha\}/B = \alpha$를 만족하는 t_α로 추정한다. $100\% \times (1-\alpha)$인 붓스트랩 신뢰구간은

$$(\hat{\theta} - t_{(1-\alpha/2)}\hat{\sigma} \quad , \quad \hat{\theta} - t_{(\alpha/2)}\hat{\sigma})$$ 이 된다. -②

붓스트랩 추정량 $\hat{\theta}^*$의 값들을 순서대로 작은 것부터 나열했을 때 $100\% \times (1-\alpha)$ 백분위 값을 $K^{-1}(1-\alpha/2)$ 라고 하면, 모수 θ에 대한 신뢰도가 $100\% \times (1-\alpha)$ 인 붓스트랩 백분위 신뢰구간은

$$(K^{-1}(\alpha/2) \quad , \quad K^{-1}(1-\alpha/2))$$ 이 된다. -③

모수 θ에 대한 $100\% \times (1-\alpha)$인 Hybrid 신뢰구간은

$$(2\hat{\theta} - K^{-1}(1-\alpha/2) \quad , \quad 2\hat{\theta} - K^{-1}(\alpha/2))$$ 이다. -④

BCa방법은 acceleration이라 불리는 $\hat{a}$와 편향 정정에 해당하는 $\widehat{z_0}$을 활용한다.
모수 θ에 대한 신뢰도가 $100\% \times (1-\alpha)$인 BCa 신뢰구간은

$(\theta^{*(\alpha_1)}$, $\theta^{*(\alpha_2)})$이 된다. -⑤

생존율 추정치 $\hat{S}$라고 하고 [0, a]구간에서는 최소한 한사람의 사망이 이루어졌다고 가정하면, 귀무가설을

H(S); $\prod_{i=1}^{k(a)} s_i = S$ 라고 하면, $\prod_{i=1}^{k(a)} \tilde{s_i} = \prod_{i=1}^{k(a)} [1 - D_i/(N_i + \lambda)] = S$ 을 통해,

모수 λ에 대한 신뢰도가 $100\% \times (1-\alpha)$ 인 우도비 신뢰구간은 $\{-2R \le \chi^2_\alpha(1)\}$이 된다. S는 λ의 증가함수이므로 모수 S에 대한 신뢰도가 $100\% \times (1-\alpha)$ 인 우도비 신뢰구간은

$P_L = \sum_{i=1}^{K(a)} [1 - D_i/(N_i + \lambda_L)]$, $P_U = \sum_{i=1}^{K(a)} [1 - D_i/(N_i + \lambda_U)]$이 된다. -⑥

생존율 추정치 $\hat{S}$로부터 얻은 $\widehat{N_*}$에 근거한 $100\% \times (1-\alpha)$ 비모수 신뢰구간을 만들 수 있는데, 이는 B(n, 1/2)에 근사한 통계량이다.

$\widehat{N_*} = max(n(1-\hat{S}(\theta))$, $n\hat{S}(\theta))$ -⑦

$\widehat{N_*}$을 계산한 후 그에 해당하는 구간의 P-value를 기준으로 하여 $\hat{S}$와 비교하여 신뢰구간을 결정한다.

3. 모의실험 구조 및 결과

Kaplan-Meier 추정량, Nelson추정량, 적률 추정량, 준모수적 방법의 추정량인 Cox regression의 β을 가지고 모의실험을 하였다. 각 추정량의 기저 분포는 지수분포와 와이블 분포를 이용하였고, 중도절단율에 대한 분포는 균일분포를 이용하여 생성하였다. 이를 10%, 20%, 35%, 50%로 놓아 중도절단율이 낮을 때와 높을 때를 골고루 배치하였다. 또한 생존함수에서 유의수준은 0.05인 경우로 표본의 크기를 50, 100, 150인 경우로 배치하였으며 1000번의 모의실험을 하였다. (단, 붓스트랩 반복수는 1000번)

Kaplan-Meier 추정량일 경우, 생존 분포가 지수분포이고 중도절단율이 10%일때 정규신뢰구간과, 일반적인 붓스트랩 신뢰구간, BCa신뢰구간의 포함확률이 0.945~0.952사이에 퍼져있어 다른 신뢰구간에 비하여 0.95에 근접하는 것을 볼 수 있었다. 백분위 신뢰구간과, Hybrid 신뢰구간의 포함확률은 0.936보다 낮아 이 경우에는 부적합하다. 비모수 신뢰구간도 표본의 크기가 50일때에는 포함확률이 0.964보다 크게 나타나 다른 신뢰구간을 사용하는 것이 바람직하다. 중도절단율이 50%일 때 비모수 신뢰구간, 붓스트랩 신뢰구간, 백분위 신뢰구간, BCa신뢰구간의 포함확률이 0.95에 가장 근접하여 중도절단율이 높을 때 유용하다. 생존 분포가 와이블일 때 지수

분포인 경우와 비슷한 형태를 보이고 있었다. 중도절단율이 10%일 경우 표본의 크기가 50일 때에는 비모수 신뢰구간을 제외한 다른 신뢰구간의 포함확률이 0.931이하로 나타났다. 중도절단율이 50%일때 일반적인 붓스트랩 방법과 백분위 신뢰구간 BCa신뢰구간, 비모수 신뢰구간의 포함확률이 0.942~0.95사이에 있어, 다른 신뢰구간에 비해 0.95에 가장 근접한 것으로 나타났다.

Nelson 추정량일 경우, 비모수 신뢰구간은 중도절단율이 20%일 때는 표본의 크기가 클때 포함확률이 0.953으로 0.95에 근접하고, 중도절단율이 35%이상일 때는 표본의 크기가 50과 100일 때 0.48로 0.95에 근접한다. 와이블 분포일때, 중도절단율이 10%일 경우 지수분포와 마찬가지로 비모수 신뢰구간만이 포함확률이 0.95에 가까웠고, 중도절단율이 20%일 경우 우도비 신뢰구간과 비모수 신뢰구간의 포함확률이 0.953과 0.955로 다른 신뢰구간에 비해 포함확률이 0.95에 가장 근사함을 볼 수 있다. 중도절단율이 35%와 50%일 경우에는 비모수 신뢰구간의 포함확률이 0.95에 가장 근접하며 표본의 크기가 100이상일때는 BCa, 우도비 신뢰구간의 포함확률도 0.945~0.960으로 0.95에 근접한다.

적률 추정량일 경우, 생존 분포가 지수 분포일 때 정규신뢰구간의 포함확률이 붓스트랩 신뢰구간의 포함확률보다 높지만 적정 범위(0.936~0.964) 안에는 속하지 못하고 있다. 중도절단율이 20%이상일 경우 우도비, 비모수 신뢰구간만이 0.95에 근사하는 것을 알 수 있다. 우도비 신뢰구간보다는 비모수 신뢰구간의 포함확률이 0.95에 근접하다. 생존 분포가 와이블일 때 우도비, 비모수 신뢰구간의 포함확률이 0.95에 가장 근접하여 사용하는 데 유용하다.

Cox regression의 β에서 생존 분포가 지수 분포일 때 중도절단율이 10%인 경우 BCa와 비모수 신뢰구간을 제외한 나머지 신뢰구간의 포함확률이 적정 범위(0.936~0.964)에 속하지 않는다. 중도절단율이 35%이상일 때 우도비 신뢰구간과 비모수 신뢰구간의 포함확률이 0.939~0.946사이에 있어 0.95에 근접하여 다른 신뢰구간에 비해 좋은 것을 알 수 있다. 생존 분포가 와이블일 경우에는 중도절단율에 상관없이 BCa, 우도비, 비모수 신뢰구간의 포함확률이 가장 0.95에 근접한 것을 알 수 있다.

4. 결론

기존의 연구들은 생존함수에서 정규 신뢰구간과의 각 신뢰구간 비교가 대부분이었고, 실제 자료를 가지고 비교를 하여 다양한 상황에서 연구된 방법들을 평가하기가 어려웠었다. 이에 기존의 방법 7가지와 생존 분포, 표본의 크기, 중도절단율을 고려하여 비교해보았다.

기존 연구 결과를 잠시 살펴보면, Kani Chen과 Shaw Hwa LO(1996)은 붓스트랩 근사 신뢰구간을 추천하였고, David R. Tomas 와 Gray L. Grunkemeier(1975)는 Kaplan-Meier추정량에 대해 우도비 신뢰구간을 추천하였다. Ron Brookmeyer와 John Crowley(1982)는 중간 값에 대한 비모수적 신뢰구간을 중도절단율이 높을 때 추천하였다. 그러나 모의실험을 통해 살펴본 결과는 붓스트랩 신뢰구간이 정규 신뢰구간에 비해서는 우월하지만, 비모수 신뢰구간이나 우도비 신뢰구간의 포함확률이 0.95에 더 근접하는 것을 확인할 수 있었다. 우도비 신뢰구간은 정규 신뢰구간보다 우월하지만 비모수 신뢰구간과 비교했을 때 포함확률도 0.95에 더 근접하고 있었다. 생존 함수에 상관없이 비모수 신뢰구간이나 일반적인 붓스트랩, BCa신뢰구간의 포함확률이 0.95에 가장 근접하였고, 특히 비모수 신뢰구간은 중도절단율이 높을 때 유용하다.

기존 연구에서 분산을 고정했을 때 비모수, 우도비 신뢰구간의 포함확률보다 더 우월한 것을 밝히고 있다. 이에 분산을 최소화할 수 있는 신뢰구간을 연구한다면 여러 상황에서도 우월성을 가질 수 있으리라 생각된다.

참고문헌

Robert L. Strawderman and Martin T. Wells. (1997), Accurate bootstrap confidence limits for the cumulative hazard and survivor functions under random censoring, *Journal of American Statistical Association,* Vol.92, No.440.

Robert L. Strawderman; Michael I. Parzen; Martin T, Well. (1997), Accurate Confidence Limits for Quantiles under Random Censoring. *Biometrics,* Vol.53, No.4.

Deborah Burr. (1994), A comparison of certain bootstrap confidence interval in the cox model, *Journal of American Statistical Association* Vol.89, No.428.

Ron Brookmeyer and John Crowley.(1982), A Confidence Interval for the Median Survival Time, *Biometrics,* Vol.38, No.1, 29-41.

B. Efron and R. Tibshirani. (1986), Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy.*Statistical Science,*Vol. 1, 54-77.

Robert A. Stine, (1985), Bootstrap Prediction Intervals for Regression,*Journal of American Statistical Association* Vol.80, No.392.

Bradly Efron, (1981), Censored Data and the Bootstrap,*Journal of the American Statistical Association,* Vol.76, No.374, 312-319.

David R. Tomas and Gray L. Grunkemeier (1975), Confidence Interval Estimation of Survival Probabilities for Censored Data, *Journal of the American Statistical Association,* Vol.70, No.352.

John D. Emerson. (1982), Nonparametric Confidence Intervals for the Median in the Presence of Right Censoring, *Biometrics,* vol.38, 17-27.

Kani Chen and Shaw-Hwa Lo. (1996), On bootstrap accuracy with censored data, *The Annals of Statistics* Vol.24, No.2, 569-595.