

Confidence Intervals for High Quantiles of Heavy-Tailed Distributions

Ji-Hyun Kim^{a,1}

^aDepartment of Statistics and Actuarial Science, Soongsil University

(Received March 7, 2014; Revised April 10, 2014; Accepted April 11, 2014)

Abstract

We consider confidence intervals for high quantiles of heavy-tailed distribution. The asymptotic confidence intervals based on the limiting distribution of estimators are considered together with bootstrap confidence intervals. We can also apply a non-parametric, parametric and semi-parametric approach to each of these two kinds of confidence intervals. We considered 11 confidence intervals and compared their performance in actual coverage probability and the length of confidence intervals. Simulation study shows that two confidence intervals (the semi-parametric asymptotic confidence interval and the semi-parametric bootstrap confidence interval using pivotal quantity) are relatively more stable under the criterion of actual coverage probability.

Keywords: Delta method, bootstrap, peaks-over-threshold, generalized Pareto distribution.

1. 서론

주식의 수익률이나 환율과 같은 금융 관련 자료 또는 강수량과 풍속 같은 기후 관련 자료의 분포를 보면 중심에서 먼 값들이 정규분포로 설명이 되지 않을 정도로 많이 관측되는 경향이 있다. 이러한 분포를 흔히 두꺼운 꼬리를 갖는 분포라 한다 (두꺼운 꼬리를 갖는 분포에 대한 다양한 정의가 있으나 여기서는 유한 차수의 적분이 무한대가 되는 분포, 즉 $E(|X|^k) = \infty$ 인 k 가 존재하는 분포로 정의하기로 한다.). 극단적 상황에서 관측될 가능성이 있는 값을 나타내는 고분위수(high quantile)의 추정은 금융이나 기후 관련 문제에서 위험 관리를 위해 중요하다. 극단적 상황에서의 손실이나 강수량을 점추정하는 것이 중요한 만큼 점추정량의 정확도를 가늠해볼 수 있는 구간추정도 중요하다. 하지만 두꺼운 꼬리를 갖는 분포의 고분위수의 점추정에 대한 연구는 많으나 구간추정에 대한 연구는 상대적으로 많지 않은데, 본 연구는 구간추정이 주목적이다.

누적확률분포함수의 역함수인 분위수를 모수의 함수로 표현할 수 있다면 모수의 최대가능도추정량(maximum likelihood estimator)에 델타 방법(delta method)을 적용하여 분위수에 대한 신뢰구간을 구할 수 있는데 Coles (2001)는 이런 방법들을 정리하였다 (Coles, 2001, 3.3.3절과 4.3.3절 참조). Kim (2011)은 고분위수의 신뢰구간을 구하는 방법 중에서 붓스트랩 방법들을 고려하였는데, POT(Peaks-Over-Threshold) 방법을 적용한 준모수적 붓스트랩 방법을 중심으로 다른 붓스트랩 방법

¹Department of Statistics and Actuarial Science, Soongsil University, Sangdo-Ro 369, Dongjak-Gu, Seoul 156-743, Korea. E-mail: jxk61@ssu.ac.kr

들과 그 성능을 비교하였다. Kim (2011)의 후속연구인 이번 연구는 표본 재추출을 하지 않고 추정량의 극한분포를 이용하는 점근적(asymptotic) 방법을 포함시켜 붓스트랩 방법과 비교하였다. 그리고 붓스트랩 신뢰구간만 고려한 Kim (2011)에서는 표본추출방법에 따른 차이를 비교하기 위해 표본추출방법을 제외한 나머지 과정을 동일하게 두었다. 후속 연구인 이 연구에서는 표본추출 과정뿐만 아니라 분위수를 추정하는 과정과 추정량의 정밀도로부터 신뢰구간을 구하는 과정도 다양하게 두어 더 많은 방법들을 고려하였다. 꼬리가 두꺼운 분포의 고분위수의 신뢰구간에 대한 연구에서 점근적 신뢰구간과 붓스트랩 신뢰구간을 같이 고려하는 연구가 없었으므로 이 연구가 적절한 신뢰구간을 선택할 때 도움을 줄 수 있을 것으로 기대한다.

2. 고분위수의 구간추정 방법

두꺼운 꼬리를 갖는 분포의 고분위수에 대한 신뢰구간을 구하는 가능한 방법들을 정리해 보기로 한다. 크게 추정량의 극한분포를 이용하는 점근적 방법과 붓스트랩 방법으로 나눌 수 있는데, 점근적 방법은 표본 재추출이 필요 없다는 장점이 있고, 붓스트랩 방법은 표본의 크기가 충분히 크지 않거나 조건부꼬리평균(CTE; Conditional Tail Expectation)과 같이 점근적 분포를 알기 어려운 통계량의 경우에도 적용할 수 있다는 장점이 있다. 한편 모집단의 분포함수에 대한 가정에 따라 신뢰구간을 구분해볼 수도 있는데, 모수적 분포함수를 가정하는 모수적 방법과 경험적 분포함수를 이용하는 비모수적 방법, 그리고 절충 형태인 준모수적 방법으로 신뢰구간을 구분할 수도 있다. 점근적 방법과 붓스트랩 방법에 이 세 가지 방법을 각각 적용해 볼 수 있다.

이 연구에서 고려한 여러 신뢰구간을 구체적으로 설명하기 위해 다음과 같은 기호를 쓰기로 한다.

F : 모집단 분포함수

\hat{F} : F 의 추정량에 대한 총칭

\hat{F}_n : 경험적(empirical) 분포함수

\hat{F}_θ : 모수적 분포함수 F_θ 를 가정하고 표본으로부터 모수 θ 를 추정하여 얻은 F_θ 의 추정량. 정확하게는 F_θ 로 표기해야 하지만 다른 추정량들의 표기와 일관성을 유지하기 위해 \hat{F}_θ 로 표기하기로 한다.

f : 모집단의 확률밀도함수

\hat{f} : POT (Peaks-Over-Threshold) 방법에 의한 f 의 추정량. 오른쪽 꼬리 부분의 분포는 일반화 파레토 분포(GPD; Generalized Pareto Distribution)에 가까워진다는 Pickands (1975)의 정리에 근거한 POT 방법에 대해서는 Kim (2011)의 2절에서 요약하여 설명하였다. 확률밀도함수 f 를 POT 방법을 적용해서 추정할 수 있는데, Kim (2011)의 2절에 있는 두 식

$$F(x) = (1 - F(u))F_u(x - u) + F(u),$$

$$\hat{F}(x) = (1 - \hat{F}_n(u))G_{\hat{\xi}, \hat{\beta}}(x - u) + \hat{F}_n(u)$$

으로부터

$$\hat{f}(x) = (1 - \hat{F}_n(u))g_{\hat{\xi}, \hat{\beta}}(x - u)$$

이다. 이 때 F_u 는 분계점(threshold) u 를 초과하는 값에 대한 조건부 분포이며 분계점 u 를 0.9 표본분위수로 정하면 $(1 - \hat{F}_n(u)) = 0.1$ 이다. G 와 g 는 GPD의 분포함수와 확률밀도함수이며 $(\hat{\xi}, \hat{\beta})$ 는 표본으로부터 추정된 GPD의 형상모수와 척도모수이다.

\hat{f}_θ : 모수적 분포를 가정했을 때 f 의 추정량. 표본으로부터 모수를 추정해 얻을 수 있다.

$\tau_p = F^{-1}(p)$: 모집단 분위수.

$\hat{\tau}_p$: 모집단 분위수의 추정량에 대한 총칭.

$\hat{\tau}_{p,n}$: 표본크기 n 일 때 τ_p 의 표본분위수(sample quantile).

$\hat{\tau}_{p,t}$: POT 추정량 (Coles, 2001, 식 (4.13)), t 는 분계점(threshold)의 첫 알파벳.

$\hat{\tau}_{p,\theta}$: 모수적 분포를 가정했을 때 τ_p 의 추정량. 표본으로부터 모수를 추정하고 분포함수의 역함수로부터 τ_p 의 추정량을 구할 수 있다.

붓스트랩 표본의 추출 여부와 모수적 분포의 가정 여부에 따라 다양한 방법이 가능한데 이 연구에서는 다음과 같은 11개의 신뢰구간을 고려하였다.

- (1) (SMPLQ) 표본분위수를 점추정량으로 하고, 표본분위수 $\hat{\tau}_{p,n}$ 의 극한분포에 대한 Bahadur (1966)의 정리

$$\sqrt{n}(\hat{\tau}_{p,n} - \tau_p) \text{의 분포는 근사적으로 } N\left(0, \frac{p(1-p)}{f^2(\tau_p)}\right)$$

를 이용하여 모분위수에 대한 근사적 신뢰구간을 구하는 방법이다. 이 방법을 적용하려면 분위수에서의 확률밀도함수값인 $f(\tau_p)$ 의 추정값이 필요하다. 확률밀도함수 f 는 POT 방법을 적용해서 추정한다. $\hat{f}(\hat{\tau}_{p,n})$ 을 $f(\tau_p)$ 의 추정값으로 쓰기로 하고 다음과 같이 $100(1-\alpha)\%$ 신뢰구간을 구한다.

$$\left(\hat{\tau}_{p,n} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n\hat{f}^2(\hat{\tau}_{p,n})}}, \hat{\tau}_{p,n} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n\hat{f}^2(\hat{\tau}_{p,n})}} \right),$$

여기서 $z_{1-\alpha/2}$ 는 표준정규분포의 $1-\alpha/2$ 분위수이다. 이 방법으로 구한 신뢰구간은 표본분위수에 근거한 신뢰구간이므로 SMPLQ라 부르기로 한다. SMPLQ는 모집단에 대해 모수적 분포를 가정하지 않는다는 점에서 비모수적이고, 붓스트랩 표본을 추출하지 않고 추정량의 극한분포를 이용한다는 점에서 점근적 신뢰구간이다.

- (2) (POTBH) POT 추정량을 점추정량으로 하고 Bahadur (1966) 정리를 이용해 신뢰구간을 구하는 방법이다. Bahadur 정리는 원래 표본분위수에 적용하는 정리이지만 POT 추정량에도 적용해 성능을 살펴보기로 한다. POTBH라 부르기로 하는 이 신뢰구간의 식은 다음과 같다:

$$\left(\hat{\tau}_{p,t} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n\hat{f}^2(\hat{\tau}_{p,t})}}, \hat{\tau}_{p,t} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n\hat{f}^2(\hat{\tau}_{p,t})}} \right)$$

POTBH는 준모수적이며 점근적 신뢰구간이다 (POT 추정량은 모집단 분포의 오른쪽 꼬리 부분에 대해서만 GPD라는 모수적 분포를 가정한다. 이런 점에서 준모수적이다.).

- (3) (PARDL) 모집단의 모수적 분포 함수를 안다고 가정하고 델타 방법으로 구하는 방법이다 (Coles, 2001, 3.3.3절 참조). 이 연구에서는 모수적 분포함수로서 GEV(Generalized Extreme Value) 분포를 가정하고 모수의 추정량으로 최대가능도추정량을 쓴다. GEV 분포에 대한 설명은 다음 절에서 하기로 한다. PARDL은 모수적이면서 점근적인 신뢰구간이다.

지금까지 설명한 세 방법(SMPLQ, POTBH, PARDL)은 붓스트랩 표본을 추출하지 않고 추정량의 점근적 분포를 이용한 신뢰구간이라는 공통점이 있다. 이 방법 외에 모집단 분포함수나 추정량의 점근적

분포를 알지 못해도 적용할 수 있는 붓스트랩 방법이 있는데, 붓스트랩 표본을 추출하는 방법과 추정량의 종류에 따라 다양한 방법으로 적용할 수 있다. 고분위수에 대한 붓스트랩 신뢰구간을 구하는 방법을 Kim (2011)의 2절에 정리하였다. Kim (2011)에서 고려했던 방법에 대해서는 간략히 언급하고 이 연구에서 새로이 고려하는 방법에 대해서만 구체적으로 설명하기로 한다.

- (4) (NBPCT) 붓스트랩 표본을 경험적 분포함수 \hat{F}_n 에서 단순확률추출하고, 표본분위수를 점추정량으로 쓰며, 백분위수방법(percentile method)으로 신뢰구간을 구하는 단순한(naive) 방법이다. Kim (2011)의 NB1과 동일한 방법이다.
- (5) (NBBCA) NBPCT와 붓스트랩 표본의 추출방법과 점추정량은 같지만, 백분위수방법 대신 Efron과 Tibshirani (1993)의 편향수정가속(Bias Corrected and Accelerated) 방법을 적용한다. Kim (2011)의 NB2와 동일한 방법이다. NBBCA는 NBPCT와 마찬가지로 비모수적 붓스트랩 신뢰구간이다.
- (6) (SBONE) 붓스트랩 표본을 POT 방법을 적용해서 얻는다는 점에서 NBPCT와 다르며 나머지 과정은 NBPCT와 같다. POT 방법으로 붓스트랩 표본을 추출하는 방법은 Kim (2011)에서 구체적으로 설명하였다. 이 방법은 비모수적 분포인 \hat{F}_n 과 모수적 분포인 일반화 파레토 분포를 결합한 분포에서 붓스트랩 표본을 추출하는 셈이므로 준모수적(semi-parametric) 방법이라 할 수 있다. SBONE은 Kim (2011)의 SB와 동일한 방법이다.
- (7) (SBTWO) 붓스트랩 표본은 SBONE과 같이 POT 방법을 적용해서 얻은 다음, 붓스트랩 표본에 대해 POT 방법을 다시 적용해 분위수를 추정한다. SBTWO는 이 연구에서 새로이 적용해 보는 방법이므로 조금 더 자세히 설명해보면, $\hat{\tau}_p^* = \hat{F}_s^{*-1}(p)$ 이며, \hat{F}_s^* 는 붓스트랩 표본에 POT 방법을 적용해 얻은 분포함수의 추정량으로서 Kim (2011)의 \hat{F}_s 의 붓스트랩 버전이다. 이로부터 준모수적 붓스트랩 방법의 p 분위수의 추정량 $\hat{\tau}_p^* = \hat{F}_s^{*-1}(p)$ 은

$$\hat{\tau}_p^* = u^* + \frac{\hat{\beta}^*}{\hat{\xi}^*} \left(\left(\frac{1-p}{N_{u^*}/n} \right)^{-\hat{\xi}^*} - 1 \right)$$

와 같이 유도된다는 것을 보일 수 있다 (McNeil 등 (2005)의 식 (7.18), (7.19) 참조. 위 식에서 N_{u^*} 는 n 개의 붓스트랩 표본 관측값 중에서 임계값 u 를 초과하는 관측값의 수임. $\hat{\beta}^*$ 와 $\hat{\xi}^*$ 은 붓스트랩 표본으로부터 추정된 일반화 파레토 분포의 모수에 대한 추정량임. 일반화 파레토 분포의 정의와 성질은 Kim (2011)을 참조). 이 과정을 B 번 반복하여 얻은 $\hat{\tau}_p^{*(1)}, \dots, \hat{\tau}_p^{*(B)}$ 에 백분위수방법을 적용하여 신뢰구간을 구한다. 붓스트랩 신뢰구간을 구하는 과정을 붓스트랩 표본추출, 추정량 구하기, 반복적으로 구한 추정량으로부터 신뢰구간 구하기의 세 과정으로 구분해서 볼 때 SBTWO는 SBONE과 두 번째 과정만 다르고 나머지 두 과정은 같다.

- (8) (SBPVT) SBONE과 SBTWO 모두 세 번째 과정에서 백분위수방법을 쓰는데 이 과정을 달리 해서 축량(pivotal quantity)을 쓰는 또 다른 신뢰구간을 시도해 보았다. 축량을 이용한 신뢰구간이 반드시 좋은 성능을 갖는 것은 아니지만 (Efron과 Tibshirani, 1993, 12.5절) 그 가능성을 확인해 보고자 하였다. 좀 더 구체적으로 설명하면 $(\hat{\tau}_p - \tau_p)/\sqrt{\widehat{\text{var}}(\hat{\tau}_p)}$ 의 근사적 분포를 $(\hat{\tau}_p^* - \hat{\tau}_p)/\sqrt{\widehat{\text{var}}(\hat{\tau}_p^*)}$ 의 분포로부터 구하는데, 붓스트랩 표본추출과 분위수의 추정량은 SBTWO와 같이 POT 방법으로 구한다. $\hat{\tau}_p^*$ 는 붓스트랩 표본에 POT 방법을 적용해서 얻은 추정량인데, $\widehat{\text{var}}(\hat{\tau}_p^*)$ 을 구할 때 표본분위수의 분산에 관한 식인 Bahadur 정리의 식을 이용하였다. B 개의 붓스트랩 축량 $(\hat{\tau}_p^* - \hat{\tau}_p)/\sqrt{\widehat{\text{var}}(\hat{\tau}_p^*)}$ 을 계산하고 이로부터 백분위수방법으로 $(\hat{\tau}_p - \tau_p)/\sqrt{\widehat{\text{var}}(\hat{\tau}_p)}$ 에 대한 $100(1 - \alpha)\%$ 신뢰구간인 (L^*, U^*) 를 구한 다음, τ_p 에 대한

$100(1 - \alpha)\%$ 신뢰구간 $(\hat{\tau}_p - U^* \sqrt{\text{var}(\hat{\tau}_p)}, \hat{\tau}_p - L^* \sqrt{\text{var}(\hat{\tau}_p)})$ 를 얻는다. SBPVT는 SBTWO와 처음 두 과정이 같으며, 마지막 세 번째 과정만 다르다. SBONE, SBTWO, SBPVT는 모두 준모수적 붓스트랩 신뢰구간이다.

- (9) (PBONE) 모수적 붓스트랩 신뢰구간이다. 모집단이 모수적 분포 F_θ 를 따른다고 가정하고, 주어진 표본으로부터 모수를 추정하여 얻은 분포 \hat{F}_θ 로부터 붓스트랩 표본을 추출한다. 분위수의 추정량으로 표본분위수를 쓰며 백분위수방법으로 신뢰구간을 구한다. PBONE은 붓스트랩 표본을 추출하는 첫 번째 과정이 SBONE이나 NBPCT와 다르지만 나머지 두 과정은 같으며, Kim (2011)의 PB와 동일하다.
- (10) (PBTWO) 모수적 붓스트랩 신뢰구간으로서 표본추출방법은 PBONE과 동일하다. 붓스트랩 표본으로부터 분위수를 추정할 때 모수적 분포함수를 이용하여 추정하며, 이 과정을 B 번 반복하여 얻은 추정량들부터 백분위수방법으로 신뢰구간을 구한다. 추정량을 구하는 두 번째 과정을 제외하고는 PBONE과 같다.
- (11) (PBPVT) 붓스트랩 표본을 추출하고 추정량을 구하는 두 과정은 PBTWO와 같으며 마지막 세 번째 과정은 SBPVT와 같이 측량을 이용한다. 측량을 구할 때 분위수 추정량의 분산이 필요한데, PARDL과 같이 델타방법을 적용한다. (Coles, 2001, 3.3.3절 참조.) PBONE, PBTWO, PBPVT는 모두 모수적 붓스트랩 신뢰구간이다.

지금까지 서술한 11개의 신뢰구간 중에서 처음 3개는 붓스트랩 표본을 추출하지 않는 방법인데, PARDL을 제외하고는 모집단의 분포함수를 가정할 필요가 없다. 나머지 8개의 붓스트랩 신뢰구간은 표본추출방법에 따라 비모수적 붓스트랩 신뢰구간(NBPVT, NBBCA), 모수적 붓스트랩 신뢰구간(PBONE, PBTWO, PBPVT), 그리고 준모수적 신뢰구간(SBONE, SBTWO, SBPVT)으로 구분할 수 있으며, 추정량을 구하는 과정과, 반복해서 구한 추정량으로부터 신뢰구간을 구하는 과정의 차이에 따라 다시 세부적으로 구분할 수 있다. 이들 중에서 점근적이면서 준모수적인 POTBH와 준모수적 붓스트랩 신뢰구간인 SBTWO와 SBPVT, 그리고 모수적 붓스트랩 신뢰구간인 PBTWO와 PBPVT는 이 연구에서 새로이 제안하는 신뢰구간이다.

3. 모의실험

앞 절에서 설명한 11종류의 신뢰구간의 성능을 모의실험을 통해 비교하고자 한다. 신뢰구간의 성능은 모집단의 분포 또는 모의실험의 조건에 따라 달라질 것으로 예상된다. 모집단의 분포족을 알고 있는 경우는 이 사실을 이용하는 신뢰구간들인 PARDL, PBONE, PBTWO, PBPVT가 더 나은 성능을 보일 것으로 예상할 수 있다. 하지만 가정한 모집단의 분포와 참 분포가 다를 경우 (즉, 가정한 모집단의 분포족에 참 분포함수가 속하지 않을 경우) 성능이 떨어질 것으로 예상할 수 있다. 그래서 가정한 모집단의 분포와 참 분포의 일치 정도에 따라 Kim (2011)과 같이 4가지 경우로 나누어 모의실험을 실시하였다. 또 표본의 크기와 오른쪽 꼬리의 극단 정도를 나타내는 $F^{-1}(p)$ 의 p 의 값에 따라 신뢰구간들의 성능이 달라질 것으로 예상하여 이 실험조건도 같이 고려하였다.

신뢰구간의 성능을 나타내는 기준으로 두 가지를 고려하였다. 신뢰구간이 모집단의 분위수를 포함할 실제 확률(actual coverage probability)과 신뢰구간의 길이이다. 첫 번째 기준을 실제신뢰수준(actual confidence level)이라 부르기로 하는데 500번의 반복실험을 통해 추정할 수 있다. 두 번째 기준도 반복 실험으로 얻은 500개의 신뢰구간의 평균 길이로 추정한다. 추정량의 표준오차도 구하여 같이 보고하기로 한다.

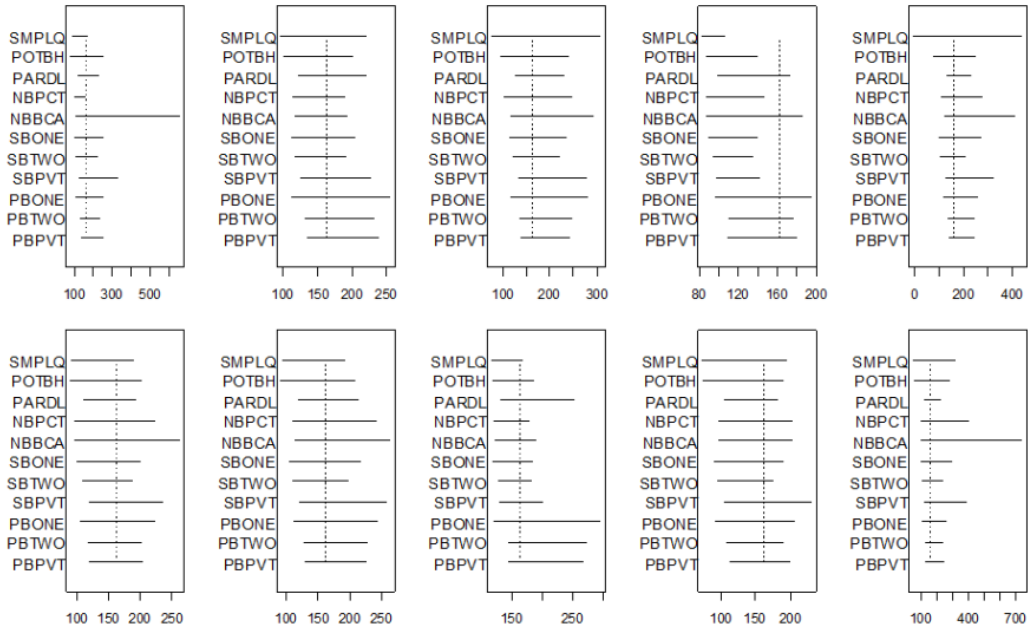


Figure 3.1. Simulation 1: 95% conf. intervals for the .99 quantile

3.1. 모의실험 1

참 모집단 분포와 가정한 모집단의 분포가 일치하는 경우이다. 모집단의 분포로 일반화극단값(GEV; Generalized Extreme Value) 분포를 사용하였다. GEV 분포는 형상모수 ξ , 위치모수 μ , 척도모수 σ 라는 3개의 모수를 갖는 융통성 있는 분포로서 $GEV(\xi, \mu, \sigma)$ 의 분포함수는 다음과 같다.

$$F(x) = \begin{cases} \exp\left(-\left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right), & \xi \neq 0, \\ \exp\left(-e^{-\frac{(x - \mu)}{\sigma}}\right), & \xi = 0. \end{cases}$$

GEV 분포는 “표준화시킨 최대값의 유일한 비퇴화 극한분포”라는 성질을 갖고 있으나 (McNeil 등, 2005, 7.1절) 이 연구에서는 형상모수 ξ 가 양일 때 오른쪽 꼬리가 두꺼운 분포가 되며 세 모수의 값에 따라 다양한 분포를 만들 수 있다는 이유 때문에 사용하였다. ξ 의 값이 커질수록 오른쪽 꼬리가 두꺼워지며 $\xi = 0.5$ 일 때 분산이 존재하지 않게 된다. 그리고 $\xi \neq 0$ 일 때 분위수함수(quantile function)는 $q(F) = \mu - \sigma(1 - (-\ln F)^{-\xi})/\xi$ 이다. Kysely (2010)와 Kim (2011)의 연구에서와 같이 참 분포로서 $GEV(\xi, 30, 10)$, $\xi = 0.1, 0.2, 0.3, 0.4, 0.5$ 를 차례로 사용하여 표본을 생성한 다음, 모집단의 분포를 GEV 분포로 가정하고 분위수를 추정한다. (위치모수와 척도모수는 실제 강수량 자료와 비슷하도록 설정하였으며, 형상모수의 여러 값에 대응하는 GEV 분포의 확률밀도함수는 Kim (2011)의 그림 1을 참조하면 된다.) 참 모집단과 추정할 때 가정하는 모집단이 일치하기 때문에 모수적이며 점근적 방법인 PARDL, 그리고 붓스트랩 표본추출에 모수적 방법을 적용한 PBONE, PBTWO, PBPVT 에게 유리한 실험조건이다. 모의실험 1뿐만 아니라 나머지 세 모의실험에서 각각 표본 크기를 100, 300, 500, 1000으로 두고, $F^{-1}(p)$ 에서 p 의 값을 0.99, 0.995로 두

Table 3.1. Simulation 1: Performance of 95% conf. intervals for the .99 quantile

ξ	성능기준	SMPLQ	POTBH	PARDL	NBPCT	NBBCA	SBONE	SBTWO	SBPVT	PBONE	PBTWO	PBPVT
	RMSE	8.7	7.1	6.5								
0.1	실제신뢰수준	0.854	0.956	0.944	0.782	0.940	0.946	0.932	0.938	0.976	0.950	0.962
	신뢰구간길이	35.2	37.5	26.0	27.5	45.0	34.6	28.8	36.7	33.6	26.1	27.2
	(표준오차)	(0.8)	(0.5)	(0.2)	(0.6)	(1.3)	(0.5)	(0.4)	(0.6)	(0.3)	(0.2)	(0.2)
	RMSE	14.3	11.8	10.7								
0.2	실제신뢰수준	0.824	0.936	0.932	0.814	0.944	0.926	0.904	0.934	0.964	0.932	0.934
	신뢰구간길이	55.5	57.5	39.8	46.1	74.3	53.3	43.3	57.4	53.5	40.3	41.6
	(표준오차)	(1.5)	(0.9)	(0.4)	(1.2)	(2.4)	(0.9)	(0.6)	(1.0)	(0.6)	(0.4)	(0.4)
	RMSE	21.5	17.6	16.1								
0.3	실제신뢰수준	0.832	0.928	0.936	0.830	0.940	0.926	0.888	0.940	0.974	0.932	0.940
	신뢰구간길이	94.7	90.7	60.9	80.1	146	85	66.1	95	86	62.4	63.5
	(표준오차)	(2.7)	(1.5)	(0.7)	(2.5)	(5.7)	(1.5)	(1.0)	(1.9)	(100)	(0.7)	(0.7)
	RMSE	34.6	26.7	24.4								
0.4	실제신뢰수준	0.768	0.896	0.936	0.800	0.930	0.898	0.858	0.924	0.986	0.938	0.948
	신뢰구간길이	139	131	93.4	119	236	124	94.2	140	138	95.8	97.2
	(표준오차)	(4.4)	(2.3)	(1.1)	(3.8)	(13.4)	(2.3)	(1.5)	(2.9)	(1.7)	(1.1)	(1.1)
	RMSE	61.7	43.3	38.2								
0.5	실제신뢰수준	0.818	0.908	0.970	0.830	0.954	0.920	0.858	0.942	0.994	0.960	0.950
	신뢰구간길이	260	210	147	231	482	199	146	234	230	153	154
	(표준오차)	(10.6)	(4.0)	(1.8)	(14.6)	(44.5)	(4.0)	(2.5)	(5.3)	(3.1)	(1.9)	(1.9)

Table 3.2. Simulation 2: Performance of 90% and 95% conf. intervals for the .99 quantile

	성능기준	SMPLQ	POTBH	PARDL	NBPCT	NBBCA	SBONE	SBTWO	SBPVT	PBONE	PBTWO	PBPVT
	RMSE	42.1	33.7	27.7								
90% 신뢰 구간	실제신뢰수준	0.782	0.870	0.754	0.766	0.772	0.846	0.778	0.874	0.860	0.796	0.822
	신뢰구간길이	142	136	77.3	128	135	123	96.7	137	107	79.5	79.8
	(표준오차)	(4.4)	(2.5)	(0.8)	(4.3)	(4.7)	(2.3)	(1.5)	(2.9)	(1.3)	(0.9)	(0.9)
95% 신뢰 구간	실제신뢰수준	0.826	0.914	0.828	0.780	0.930	0.920	0.854	0.940	0.928	0.872	0.894
	신뢰구간길이	169	162	92.1	146	283	153	117	173	131	95.2	96.0
	(표준오차)	(5.3)	(2.9)	(1.0)	(4.8)	(12.4)	(2.9)	(1.9)	(3.7)	(1.6)	(1.1)	(1.0)

어 실험하였다. $n = 300$ 이고 $p = 0.99$ 인 경우의 결과만 여기서 보고하고 나머지 경우는 별도의 파일(<http://bayes.ssu.ac.kr/jhkim/Publication/SimulationResult.txt>)에 보고하였다.

11종류의 신뢰구간의 평균적 성능을 알아보기 전에 개별적 특성을 파악하기 위해 10번의 모의실험에서 11종류의 신뢰구간을 각각 구해서 Figure 3.1에 나타내었다. 이 때 GEV 분포의 형상모수 ξ 는 0.4로 두어 비교적 꼬리가 두꺼운 분포로 설정 하였으며, 한 번의 모의실험마다 11개의 신뢰구간을 구하고, GEV 분포의 분위수함수로부터 계산한 0.99 분위수 값인 162.4를 그래프에 세로 방향 점선으로 표시하여 각 신뢰구간이 참 분위수를 포함하는지를 알 수 있게 하였다. Figure 3.1을 보면 우선, SBONE보다 SBTWO가, 또한 PBONE보다 PBTWO가 더 짧은 길이를 갖는다는 것을 알 수 있다. 각 쌍의 방법들은 붓스트랩 표본추출 방법은 같고 추출된 붓스트랩 표본에서 분위수의 추정량을 어떻게 구하느냐만 다른데, SBONE과 PBONE은 모두 표본분위수를 점추정량으로 쓴다. 따라서 SBONE과 PBONE 두 신뢰구간의 길이가 더 긴 이유는 점추정량인 표본분위수의 변동이 심해서 생긴 결과로 보인다. Kim (2011)에서 붓스트랩 표본추출 방법의 차이를 비교하기 위해 점추정량은 표본분위수로 동일하게 두었으므로 Kim (2011)은 SBTWO나 PBTWO를 아예 고려하지도 않았다. 따라서 500번의 모의실험에 의한 평균적 결과에서 재확인하게 될 이 사실, 즉 붓스트랩 방법에서 표본분위수를 점추정량을 쓰면 신뢰구간의 길이가 길어진다는 사실의 발견은 이 연구에서 새로 얻게 된 것으로서 실용적 의미가 있다. 한편 NBBCA의 길이는 가끔 지나치게 길다는 사실도 Figure 3.1의 처음과 마지막 그림으로부터 알 수 있다.

Table 3.3. Simulation 3: Performance of 90% and 95% conf. intervals for the .99 quantile

	성능기준	SMPLQ	POTBH	PARDL	NBPCT	NBBCA	SBONE	SBTWO	SBPVT	PBONE	PBTWO	PBPVT
90%	RMSE	23.9	18	14.8								
신뢰	실제신뢰수준	0.778	0.874	0.750	0.804	0.800	0.866	0.816	0.850	0.830	0.788	0.802
신뢰	신뢰구간길이	78	74.8	37.7	72.9	78.4	67.7	53	75.3	51.7	39.4	38.9
구간	(표준오차)	(2.7)	(1.4)	(0.4)	(2.4)	(3.6)	(1.3)	(0.9)	(1.6)	(0.6)	(0.4)	(0.4)
95%	실제신뢰수준	0.806	0.906	0.806	0.830	0.928	0.908	0.872	0.916	0.900	0.852	0.882
신뢰	신뢰구간길이	93	89.2	44.9	81.7	170	83.8	64	95.4	62.7	47.1	46.8
구간	(표준오차)	(3.2)	(1.6)	(0.5)	(3.0)	(8.3)	(1.6)	(1.0)	(2.1)	(0.7)	(0.5)	(0.5)

신뢰구간의 길이가 짧더라도 명목신뢰수준을 보장하지 못하면 좋은 신뢰구간이라 할 수 없다. 각 모의 실험 조건에서 500번의 모의실험을 통해 실제신뢰수준을 구해 Table 3.1에 보고하였다. 90%와 95% 두 개의 신뢰수준을 갖는 신뢰구간을 구하였으나, 여기서는 95% 신뢰구간만 Table 3.1에 결과를 보고 하고 90% 신뢰구간은 따로 보고하기로 한다 (<http://bayes.ssu.ac.kr/jhkim/Publication/SimulationResult.txt>).

신뢰구간의 성능을 비교하기 전에 점추정량의 성능을 500번의 반복실험에서 구한 제곱근평균제곱오차(RMSE; Root Mean Square Error)를 기준으로 비교하였다. 고분위수의 점추정량으로 표본분위수 $\hat{\tau}_{p,n}$, POT 추정량 $\hat{\tau}_{p,t}$, 모수적 분포함수에 의한 추정량 $\hat{\tau}_{p,\theta}$ 등 세 개를 고려하였는데, 표본분위수는 SMPLQ, NBPCT, NBBCA, SBONE, PBONE에 쓰이며, POT 추정량은 POTBH, SBTWO, SBPVT에, 그리고 $\hat{\tau}_{p,\theta}$ 는 PARDL, PBTWO, PBPVT에 쓰인다. 점추정량의 성능을 신뢰구간의 성능을 비교하는 표에 같이 나타내기 위해 세 개의 점추정량이 쓰이는 SMPLQ, POTBH, PARDL 밑에 표시하였다. 모든 형태모수의 값에서 $\hat{\tau}_{p,\theta}$ 의 RMSE 값이 가장 작았고, $\hat{\tau}_{p,n}$ 의 RMSE 값이 가장 컸다. 모의실험 1에서는 모집단의 분포를 제대로 가정하고 있으므로 예상대로 모수적 추정량인 $\hat{\tau}_{p,\theta}$ 가 가장 정확하며, 표본의 크기가 크지 않을 때 고분위수의 추정량으로 표본분위수 $\hat{\tau}_{p,n}$ 는 그 변동이 커서 좋지 않음을 알 수 있다.

참 모집단과 가정 모집단이 일치하는 모의실험 1의 특성을 고려했을 때 예상대로 PARDL의 성능이 가장 좋았다. $2\sqrt{0.95 \times 0.05/500} \approx 0.02$ 이므로 실제신뢰수준(의 추정량)이 0.93 이상이면 명목신뢰수준 0.95에서 벗어나지 않는다고 볼 수 있는데, PARDL은 이 조건을 만족하면서도 신뢰구간의 길이가 다른 신뢰구간의 길이보다 짧아 가장 만족스럽다. 한편, 붓스트랩 신뢰구간 중에서는 역시 예상대로 PBTWO나 PBPVT가 좋은 성능을 보였다. NBPCT는 명목신뢰수준을 보장하지 못하며, NBBCA는 실제신뢰수준이 명목신뢰수준에 가깝지만 길이가 지나치게 길어 좋은 신뢰구간이라 할 수 없다. (꼬리가 두꺼운 분포의 고분위수를 추정할 때 비모수적 붓스트랩 방법이 적절하지 못한 방법이라는 사실은 Kysely (2010)와 Kim (2011)에서도 지적된 바 있지만 여기서는 다른 방법과의 비교를 위해 다시 고려하였다.) 준모수적 붓스트랩 방법 중에서는 SBPVT가 모수적 붓스트랩 방법에 비해 길이는 다소 길지만 실제신뢰수준이 명목신뢰수준에 가까웠다.

3.2. 모의실험 2

두 번째 모의실험부터 참 분포와 적합 분포를 다르게 하였다. 참 분포로 일반화로지스틱(GLO; general logistic) 분포를 사용하였다. GLO(ξ, μ, σ) 분포의 분포함수는 다음과 같이 정의되는데,

$$F(x) = \begin{cases} 1 / \left(1 + \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right), & \xi \neq 0, \\ 1 / \left(1 + e^{-\frac{x - \mu}{\sigma}} \right), & \xi = 0. \end{cases}$$

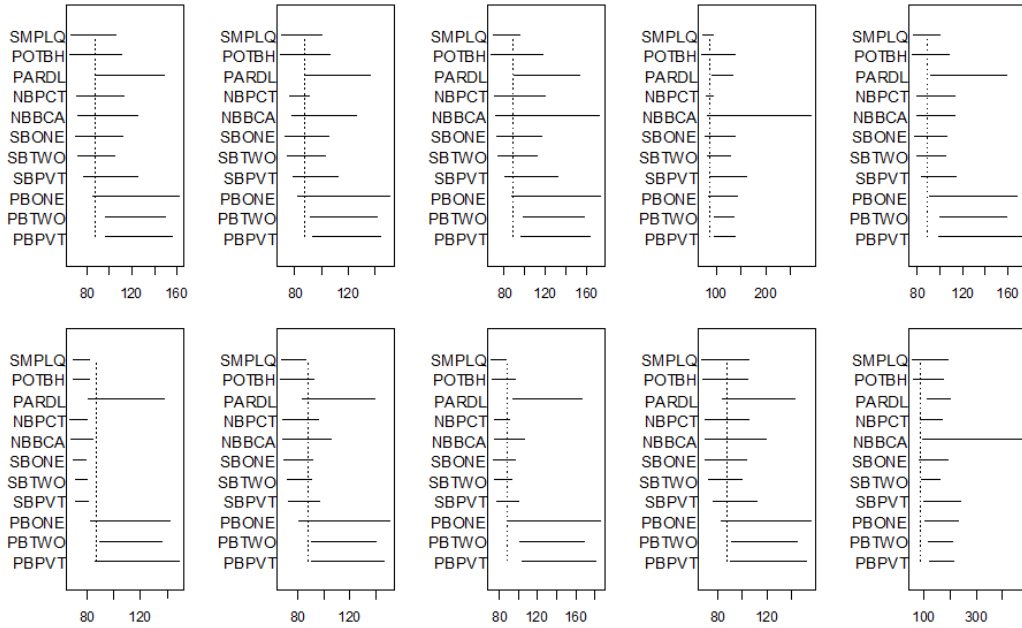


Figure 3.2. Simulation 4: 95% conf. intervals for the .99 quantile

일반화극단값(GEV) 분포와 함께 강수량에 대한 분포로 자주 쓰인다. Kysely (2010)와 Kim (2011)의 모의실험에서와 같이 참 분포를 GLO(0.4, 42, 15)로 하고, 적합 분포로 GEV 분포를 가정하였다.

GLO 분포를 GEV 분포로 적합 시키면 매우 유사한 분포를 얻게 되는데 (Kim, 2011, 그림 5 참조) 따라서 참 분포와 적합 분포가 동일할 때와 같이 모수적 방법들의 성능이 그리 나쁘지 않을 것으로 예상된다. 모의실험 1에서와 마찬가지로 점추정량으로는 모수적 점추정량인 $\hat{\tau}_{p,\theta}$ 의 RMSE가 제일 작았다. 하지만 Table 3.2의 신뢰구간의 성능을 보면 SBPVT를 제외한 나머지 방법들의 실제신뢰수준이 낮아 적절한 신뢰구간이 되지 못함을 알 수 있다. 특히 모의실험 1에서 좋은 성능을 보인 PARDL의 길이는 여전히 짧았으나 실제신뢰수준이 0.754, 0.828로서 명목신뢰수준인 0.9와 0.95를 각각 한참 밑돌았으며 모수적 붓스트랩 방법에 의한 신뢰구간들도 마찬가지로 신뢰수준에 문제를 보였다.

3.3. 모의실험 3

세 번째 모의실험에서는 또 다른 참 분포를 사용해 보았는데, Kysely (2010)에서와 같이 GEV 분포와 GLO 분포의 혼합분포(mixture distribution)를 참 분포로 가정하였다. 즉, f_1 을 GEV(0.1, 30, 10) 분포의 확률밀도함수라고 하고, f_2 를 GLO(0.4, 45, 12) 분포의 확률밀도함수라고 할 때, 참 분포의 확률밀도함수를 $0.75f_1(x) + 0.25f_2(x)$ 로 두었다. 적합 분포는 모의실험 2에서와 같이 GEV 분포를 가정하였다 (Kim, 2011, 그림 7 참조).

Table 3.3을 보면 잘못된 분포를 가정하고 있음에도 불구하고 모수적 점추정량인 $\hat{\tau}_{p,\theta}$ 의 RMSE가 제일 작았다. 실제신뢰수준을 기준으로 11개 신뢰구간의 성능을 비교해 보면, 90% 신뢰구간에서는 점근적이면서 준모수적 방법인 POTBH가 명목신뢰수준에 가까운 성능을 보였으나, 95% 신뢰구간에서는 그렇지 않았다. 전반적으로 만족스러운 추정량은 없으나 준모수적 방법인 SBPVT는 90%와 95% 신뢰구간 모두에서 상대적으로 나쁘지 않은 결과를 보였다. 신뢰구간의 길이라는 기준에서 보면 점근적이면서 모

Table 3.4. Simulation 4: Performance of 90% and 95% conf. intervals for the .99 quantile

	성능기준	SMPLQ	POTBH	PARDL	NBPCT	NBBCA	SBONE	SBTWO	SBPVT	PBONE	PBTWO	PBPVT
90%	RMSE	10.7	8.0	30.6								
신뢰 구간	실제신뢰수준	0.766	0.890	0.304	0.772	0.780	0.858	0.836	0.874	0.558	0.122	0.120
	신뢰구간길이 (표준오차)	34.8 (1.3)	33.6 (0.6)	48.2 (0.5)	31.2 (1.0)	32.1 (1.1)	30.5 (0.6)	24.1 (0.4)	33.4 (0.8)	60.9 (0.7)	44.7 (0.4)	49.7 (0.5)
95%	실제신뢰수준	0.820	0.926	0.554	0.790	0.940	0.926	0.884	0.936	0.772	0.240	0.224
신뢰 구간	신뢰구간길이 (표준오차)	41.4 (1.6)	40.1 (0.8)	57.5 (0.5)	33.6 (1.1)	78.3 (7.1)	37.7 (0.8)	29.0 (0.5)	42.4 (1.0)	74.8 (0.8)	53.7 (0.5)	59.8 (0.6)

수적 방법인 PARDL과 모수적 붓스트랩 방법인 PBTWO, PBPVT가 나은 성능을 보였으나 명목신뢰수준을 보장하지 못하는 상황에서 길이를 비교하는 것은 적절하지 않다.

3.4. 모의실험 4

모의실험 3에서 가정한 참 분포는 혼합분포이긴 하지만 단봉분포(unimodal distribution)이다. 적합 분포인 GEV 분포는 세 개의 모수의 값에 따라 다양한 단봉분포에 좋은 근사분포가 될 수 있어 잘못된 분포를 가정한 모수적 방법의 성능이 그리 나쁘지 않았을 수 있다. 모의실험 4에서는 참 분포로 이봉분포를 갖는 GEV(0.1, 30, 5)와 GLO(0.4, 55, 5)의 혼합분포를 가정하였고, 적합 분포로 앞의 모의실험에서와 마찬가지로 단봉분포를 갖는 GEV 분포를 가정하였다 (Kim, 2011, 그림 9 참조). 이봉분포를 단봉분포로 가정하는 것이 다소 비현실적일 수 있지만 표본 크기가 크지 않을 때 모집단 분포의 이봉 여부를 제대로 판단하지 못할 수 있다. 하지만 이 연구에서는 그런 현실성 여부를 떠나 모집단의 참 분포를 가정하는 모수적 방법과, 그런 가정을 하지 않는 준모수적 방법이나 비모수적 방법이 상황에 따라 얼마나 차이가 날 수 있는가를 보고자 Kim (2011)에서와 같이 이런 실험조건을 고려하기로 한다.

먼저 각 신뢰구간의 특성을 파악하기 위해 모의실험 1과 같이 10번의 모의실험 결과를 그림으로 나타내었다. 모수적 방법인 PARDL, PBONE, PBTWO, PBPVT를 보면 참 분위수를 포함하지 않는 경우가 자주 발생하고 길어도 길어 적절한 신뢰구간이 아님을 알 수 있다. 반면 준모수적 방법이나 비모수적 방법들은 길이에 차이가 있긴 하지만 대부분 참 0.99 분위수를 포함하는 것을 알 수 있다.

모의실험을 500번 반복해서 얻은 평균적 성능을 Table 3.4에 보고하였다. 먼저 세 개의 점추정량을 비교해 보면 모수적 점추정량인 $\hat{\tau}_{p,\theta}$ 의 RMSE가 커서 참 분포와 적합 분포와 많이 다른 경우 적절한 추정량이 될 수 없음을 확인할 수 있다.

점근적이면서 모수적 방법인 PARDL과 모수적 붓스트랩 방법인 PBONE, PBTWO, PBPVT 모두 도저히 쓸 수 없을 정도의 나쁜 성능을 보였으며, 점근적이면서 준모수적 방법인 POTBH나 준모수적 붓스트랩 방법인 SBONE과 SBPVT가 안정된 성능을 보였다. 특정한 분포를 가정하지 않는 방법들이 네 번째 모의실험에서 안정된 결과를 보일 것으로 예상하였지만, 비모수적 방법들은 명목신뢰수준을 보장하지 못한 반면 준모수적 방법들은 안정된 결과를 보인 점은 주목할 점이다.

지금까지 네 종류의 모의실험에서 표본크기를 300으로 고정하고 0.99 분위수에 대해 살펴보았다. 참 분포와 적합 분포가 일치할 경우 모수적 방법인 PARDL와 PBPVT의 성능이 좋았지만, 두 분포가 달라질수록 그 성능이 떨어졌다. 반면 준모수적 방법인 POTBH와 SBPVT 등은 비교적 안정된 결과를 보였다. 또한 SBTWO와 비교했을 때 SBPVT가 실제신뢰수준의 기준에서 나은 성능을 보여 추량을 이용하는 것이 도움이 됨을 알 수 있었다. 이러한 정성적(qualitative) 결과는 표본크기가 300에서 500과 1000으로 달라지거나 0.99 분위수가 아닌 0.995 분위수에 대해서도 그대로 유지되었는데, Table 3.5에 신뢰구간의 성능 기준 중에서 신뢰수준만 단순화하여 보고하였다. (실제신뢰수준의 정확

Table 3.5. Actual coverage probabilities of 95% conf. intervals (x: <0.85, Δ : 0.85~0.9, \diamond : 0.9~0.93, \circ : 0.93~0.97, \uparrow : >0.97)

실험	p	n	SMPLQ	POTBH	PARDL	NBPCT	NBBCA	SBONE	SBTWO	SBPVT	PBONE	PBTWO	PBPVT	
1	.99	300	x	Δ	\circ	x	\circ	Δ	Δ	\diamond	\circ	\circ	\circ	
		500	x	\diamond	\circ	Δ	\diamond	\diamond	Δ	Δ	\uparrow	\circ	\circ	
		1000	Δ	\circ	\circ	\circ	\circ	\circ	\diamond	\circ	\uparrow	\circ	\circ	
	.995	300	x	\diamond	\circ	x	x	\diamond	Δ	\circ	\uparrow	\circ	\circ	
		500	x	\diamond	\circ	Δ	\diamond	\diamond	Δ	Δ	\uparrow	\circ	\circ	
		1000	Δ	\diamond	\circ	Δ	\diamond	\circ	Δ	\diamond	\uparrow	\circ	\circ	
2	.99	300	x	\diamond	x	x	\circ	\diamond	Δ	\circ	\diamond	Δ	Δ	
		500	Δ	\circ	Δ	Δ	\circ	\circ	Δ	\circ	\circ	Δ	\diamond	
		1000	Δ	\circ	x	\diamond	\diamond	\circ	\diamond	\diamond	\circ	x	Δ	
	.995	300	x	Δ	x	x	x	Δ	Δ	\diamond	\circ	Δ	Δ	
		500	x	\circ	x	Δ	\circ	\diamond	Δ	\circ	\circ	x	Δ	
		1000	x	\circ	x	Δ	\circ	\circ	\diamond	\diamond	\circ	x	x	
3	.99	300	x	\diamond	x	x	\diamond	\diamond	Δ	\diamond	\diamond	Δ	Δ	
		500	Δ	\circ	x	Δ	\circ	\circ	\diamond	\diamond	\diamond	x	Δ	
		1000	Δ	\circ	x	\diamond	\diamond	\circ	Δ	\diamond	Δ	x	x	
	.995	300	x	\diamond	x	x	x	Δ	Δ	\diamond	Δ	x	Δ	
		500	x	Δ	x	Δ	\diamond	Δ	Δ	Δ	x	x	x	
		1000	x	\diamond	x	Δ	\circ	\circ	Δ	\diamond	x	x	x	
4	.99	300	x	\diamond	x	x	\circ	\diamond	Δ	\circ	x	x	x	
		500	x	\circ	x	Δ	\diamond	\circ	\diamond	\circ	x	x	x	
		1000	Δ	\circ	x	\circ	\circ	\circ	\diamond	\diamond	x	x	x	
	.995	300	x	\diamond	x	x	x	Δ	Δ	Δ	Δ	x	x	x
		500	x	\diamond	x	x	Δ	Δ	Δ	Δ	x	x	x	
		1000	x	\circ	x	Δ	\diamond	\circ	\diamond	Δ	x	x	x	

구간의 길이 등은 별도로 보고하기로 한다.) 500번의 모의실험이므로 실제신뢰수준의 추정값이 0.93에서 0.97 사이이면 명목신뢰수준 0.95를 벗어나지 않는다고 보아 동그라미(\circ)로 표시하였고, 0.9~0.93, 0.85~0.9를 구분하여 다른 기호(\diamond , Δ)로 표시하였다. 0.85에 미치지 못한 경우 가위표(x)로 표시하여 부적절함을 나타내었으며, 0.97을 상회하는 것들은 길이가 지나치게 긴 것들이라 역시 적절한 신뢰구간이 아니었는데 위 방향 화살표(\uparrow)로 표시하여 구분하였다. 모수적 방법이나 비모수적 방법들에는 모의 실험의 종류에 따라 가위표가 보이지만, 준모수적 방법들인 POTBH, SBONE, SBTWO, SBPVT에는 가위표가 없어 신뢰수준의 기준에서 보았을 때 안정된 성능을 가짐을 알 수 있다.

표본크기가 300에서 100으로 작아지면 준모수적 방법의 성능이 다소 떨어지게 되는데, 이는 POT 추정량의 특성 때문으로 파악된다. POT 추정량은 분계점(threshold)을 초과하는 값들을 이용하여 오른쪽 꼬리 부분을 추정하게 되는데, 분계점의 값을 크게 잡을수록 일반화 파레토 분포(GPD)가 오른쪽 꼬리 부분에 대한 좋은 근사 분포가 되는 반면 (Pickands, 1975), 이를 만족하는 표본 크기가 줄어들게 되어 GPD의 모수 추정이 어렵게 된다. 따라서 적절한 분계점을 잡아 균형을 취해야 하는데, 이 연구에서는 0.9 표본분위수를 분계점으로 잡았다. 하지만 표본크기가 100이 되면 이를 만족하는 표본크기가 10개밖에 되지 않아 GPD의 모수 추정이 부정확해지게 되며, 분계점을 0.8 표본분위수로 작게 잡으면 표본크기는 20개로 늘어나지만 GPD가 오른쪽 꼬리 부분에 대한 좋은 근사 분포가 되지 못한다. 따라서 POT 추정량을 이용하는 준모수적 방법은 표본크기가 작을 때 적용하기에는 적절하지 않아 보인다.

4. 결론

지금까지 꼬리가 두꺼운 분포의 고분위수에 대한 구간추정을 다양한 방법으로 할 수 있음을 제시하고

모의실험 결과를 보고하였다. Kysely (2010)는 비모수적 붓스트랩 방법과 모수적 붓스트랩 방법을 비교하였고, Kim (2011)은 준모수적 방법을 추가적으로 고려하였으나 표본추출방법에만 준모수적 방법을 적용하였다. 이 연구에서는 붓스트랩 방법이 아닌 Coles (2001)에 정리되어 있는 점근적 방법도 같이 고려하였으며, 준모수적 방법을 표본추출뿐만 아니라 점추정량을 구하는 데에도 적용할 수 있음을 제안하였고, 반복해서 구한 붓스트랩 점추정량으로부터 붓스트랩 신뢰구간을 구할 때에도 백분위수방법뿐 아니라 측량을 이용한 방법까지 같이 고려하여 비교하였다.

구간추정은 점추정량에 근거하는데, 먼저 세 가지 점추정량의 성능을 제곱근평균제곱오차(RMSE) 기준으로 비교하였다. 그 결과 모집단의 실제 분포인 참 분포와 추정할 때 가정하는 모집단의 분포인 적합 분포가 일치하는 경우 점근적이면서 모수적 추정량인 $\hat{\tau}_{p,\theta}$ 의 성능이 제일 좋았으나 참 분포와 적합 분포의 차이가 커지면 다른 두 추정량보다 낮은 성능을 보였다. 준모수적 추정량인 POT 추정량은 비모수적 추정량인 표본분위수보다 모든 실험조건에서 더 나은 성능을 보였으므로 참 분포에 대한 정보 여부에 따라 모수적 추정량과 준모수적 추정량 중에서 하나를 선택하면 좋을 것이다.

세 가지 점추정량에 근거해서 11개의 구간추정 방법을 제시하고 비교하였다. 세 추정량의 점근적 분포를 이용한 신뢰구간 3종류와 세 추정량에 붓스트랩 방법을 적용한 8종류의 신뢰구간을 고려하였다. 붓스트랩 신뢰구간을 다시 비모수적, 준모수적, 모수적 방법으로 구분할 수 있다. 참 분포와 적합 분포가 일치한다면, 신뢰수준과 길이라는 두 기준에서 봤을 때 점근적이면서 모수적 추정량을 이용한 PARDL이나 모수적 붓스트랩 신뢰구간인 PBPVT를 쓰는 것이 좋을 것이며, 참 분포에 대한 충분한 정보가 없을 때는 준모수적 신뢰구간인 POTBH나 SBPVT를 쓰는 것이 신뢰수준의 기준에서 봤을 때 안전하다.

References

- Bahadur, R. R. (1966). A note on quantiles in large samples, *Annals of Mathematical Statistics*, **37**, 577–580.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman and Hall.
- Kim, J. (2011). Semi-parametric bootstrap confidence intervals for high-quantiles of heavy-tailed distributions, *Communications of the Korean Statistical Society*, **18**, 717–732.
- Kysely, J. (2010). Coverage probability of bootstrap confidence intervals in heavy-tailed frequency models, with application to precipitation data, *Theoretical and Applied Climatology*, **101**, 345–361.
- McNeil, A., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management*, Princeton University Press.
- Pickands, J. (1975). Statistical inference using extreme order statistics, *Annals of Statistics*, **3**, 119–131.

꼬리가 두꺼운 분포의 고분위수에 대한 신뢰구간

김지현^{a,1}

^a승실대학교 정보통계보험수리학과

(2014년 3월 7일 접수, 2014년 4월 10일 수정, 2014년 4월 11일 채택)

요약

꼬리가 두꺼운 분포의 고분위수에 대한 신뢰구간을 연구하였다. 통계량의 극한 분포에 근거한 점근적 방법과 붓스트랩 방법을 같이 고려하였다. 이 두 방법에 모수적, 비모수적, 준모수적 기법을 각각 적용할 수 있는데, 전체 11가지 신뢰구간의 성능을 실제신뢰수준과 길이로 비교하였다. 모의실험 결과 준모수적이면서 점근적인 신뢰구간과 추량을 이용하는 준모수적 붓스트랩 신뢰구간이 실제신뢰수준의 기준에서 안정된 성능을 보인다는 것을 알 수 있었다.

주요용어: 델타 방법, 붓스트랩, 분계점 초과 방법, 일반화 파레토 분포.

¹(156-743) 서울시 동작구 상도로 369, 승실대학교 정보통계보험수리학과. E-mail: jxk61@ssu.ac.kr