

## A Comparison of Some Approximate Confidence Intervals for the Poisson Parameter

Daehak Kim<sup>1)</sup>, Hyeong Chul Jeong<sup>2)</sup>

### Abstract

In this paper, we reviewed thirteen methods for finding confidence intervals for the mean of poisson distribution. Bootstrap confidence intervals are also introduced. Two bootstrap confidence intervals are compared with the other existing eleven confidence intervals by using Monte Carlo simulation with respect to the average coverage probability of Woodroffe and Jhun(1989).

*Keywords* : Poisson distribution, Confidence interval, Bootstrap, Average coverage probability

### 1. 소개

포아송분포는 발생빈도가 드문 희귀한 사건의 횟수의 관찰 등 다양한 현상에서 발견되며 생물학 등 제 학문분야에 응용성이 높은 분포이다. 이제 평균이  $\mu$ 인 포아송분포를 따르는  $n$ 개의 랜덤표본  $X_1, X_2, \dots, X_n$  이 주어졌을 때, 기본적인 통계적 추론 중의 하나는 모수  $\mu$  에 대한 신뢰구간을 추정하는 것이다. 이에 대해 대부분의 통계학 책에서는  $\bar{X} \pm z_{1-\alpha/2} \sqrt{\bar{X}/n}$  과 같은 신뢰구간을 소개하고 있는데 이를 Wald 유형의 신뢰구간이라 부른다. 여기서  $z_{1-\alpha}$  는 표준정규분포의  $100(1-\alpha)\%$  분위수이다. Wald 유형의 신뢰구간은 포아송분포가 지니는 이산성 때문에 주어진 명목수준을 하향 추정하는 경향이 있다.

이와는 달리 대조적인 신뢰구간으로 “정확”(exact) 신뢰구간을 들 수 있다. Agresti and Coull(1998)가 이항분포에서 실험한 결과를 살펴보면 대부분의 정확신뢰구간은 주어진 명목수준을 상당부분 상향 추정하는 경향이 있다는 것이다. 또한 Efron(1979)의 붓스트랩 방법을 이용하여 모수  $\mu$ 의 신뢰구간을 구축할 수 있다. Singh(1981)은 포아송분포와 같이 격자분포를 따르는 분포에

---

1) Associate Professor, Department of Statistical Information, Catholic University of Daegu, Kyungbuk, 712-702, Korea E-mail : dhkim@cuth.cataegu.ac.kr

2) Assistant Professor, Department of Computational Science and Statistics, Pyoungtaek University, Kyunggi, 450-701, Korea E-mail : jhc@ptuniv.ac.kr

서 붓스트랩분포와 실제분포와의 차이가  $O(1/\sqrt{n})$ 이며, 정규근사를 이용한 분포와 실제분포와의 차이 역시  $O(1/\sqrt{n})$ 으로 정규근사와 붓스트랩근사 간에 큰 차이가 없음을 보인다. 하지만, Woodroffe and Jhun(1989)은 평균포함확률(average coverage probability) 측면에서 붓스트랩 분포와 실제분포와의 차이를 실질적으로  $\alpha(1/\sqrt{n})$ 까지 줄일 수 있음을 제시하였다. 이와 같은 사실에 따라 모수  $\mu$ 의 신뢰구간 구축에 붓스트랩을 사용하는 것도 의미가 있을 것이다.

본 연구에서는 Sahai and Khurshid(1993)의 포아송분포에 대한 여러 신뢰구간을 살펴보고, 붓스트랩 신뢰구간과 기존의 여러 신뢰구간들을 평균포함확률 측면에서 비교하여, 최적의 신뢰구간을 구축하는 방법을 살펴보고자 한다. 2장에서는 모수  $\mu$ 에 대한 여러 신뢰구간들을 살펴보고 3장에서는 이들을 모의실험으로 비교하였다.

## 2. $\mu$ 의 신뢰구간 구축 및 비교방법

### 2.1 정확신뢰구간

$X_1, X_2, \dots, X_n$  이 평균  $\mu$ 인 포아송 분포로부터 얻은 랜덤표본일 때 그 합  $X = \sum X_i$ 는 평균이  $\nu = n\mu$ 인 포아송분포를 따른다.  $\nu$ 의  $100(1-\alpha)\%$  정확신뢰구간을 구하려면, 관찰값  $x = \sum x_i$ 에 대해 유의수준  $\alpha/2$ 에서 귀무가설  $H_0: \nu = \nu_0$ 을 기각하지 않는 모든  $\nu_0$ 를 계산해야 된다. 즉, 주어진  $x$ 에 대해, 양측검정에서, 어떤  $\nu_0$ 를 사용하여야 귀무가설을 채택할 수 있는가하는 문제이다. 이때 얻어지는 모든  $\nu_0$  중 최소값을 정확신뢰구간의 하한으로 하고, 최대값을 정확신뢰구간의 상한으로 놓는다. 즉, 최소값  $\nu_L$ 과 최대값  $\nu_U$ 는 다음과 같다.

$$P(X \geq x | \nu = \nu_L) = \sum_{k=x}^{\infty} e^{-\nu_L} \nu_L^k / k! \leq \alpha/2$$

$$P(X \leq x | \nu = \nu_U) = \sum_{k=0}^x e^{-\nu_U} \nu_U^k / k! \leq \alpha/2$$

이같은 신뢰구간을 구하기 위해서 상당한 계산이 요구된다. 그러나 포아송분포의 누적확률  $P(X \leq x | \nu)$ 는 자유도가  $\nu = 2(1+x)$ 인 카이제곱분포 형태로 표현하면 다음과 같다.

$$\sum_{k=0}^x (e^{-\nu} \nu^k) / k! = P\{\chi_{\nu}^2 > 2\nu\}$$

이를 이용하면 정확신뢰구간을 쉽게 유도할 수 있다.

$\mu$ 에 대한 정확신뢰구간의 상한과 하한은  $\nu$ 에 대해 표본의 크기  $n$ 으로 나누어 구한다. 이러한 정확신뢰구간(EX)은 다음과 같다.

$$EX(L) : \mu_L = \frac{1}{2n} \chi_{2(x+1), \alpha/2}^2 \quad EX(U) : \mu_U = \frac{1}{2n} \chi_{2(x+1), 1-\alpha/2}^2 \quad (2.1)$$

정확신뢰구간을 구할 때는 카이제곱분포를 이용할 수 있으나, 다음과 같은 카이제곱분포의 정규근사에 의하여 정규분포로부터 신뢰구간을 구할 수 있다. 즉,

$$\chi_{\nu, \alpha}^2 = \nu \left( 1 - \frac{2}{9\nu} + z_{\alpha} \sqrt{\frac{2}{9\nu}} \right)^3$$

위의 근사관계식에 의해, 다음과 같은 신뢰구간을 얻을 수 있다. 이를 CH로 표기하자.

$$CH(L) : \mu_L = \bar{x} \left[ 1 - \frac{1}{9x} - \frac{z_{1-\alpha/2}}{3\sqrt{x}} \right]^3 \quad CH(U) : \mu_U = \left( \bar{x} + \frac{1}{n} \right) \left[ 1 - \frac{1}{9(x+1)} + \frac{z_{1-\alpha/2}}{3\sqrt{x+1}} \right]^3 \quad (2.2)$$

위의 신뢰구간은 근본적으로 정확신뢰구간(EX)에서 카이제곱분포를 정규근사 한 것에 불과하므로 정확신뢰구간과 큰 차이를 보이지 않을 것으로 예상된다.

Wald 유형 신뢰구간(WA)은 모수  $\mu$ 에 대한 최대우도추정량  $\bar{X} = \sum X/n$  와 정규근사를 이용하여 쉽게 얻을 수 있다. 즉,

$$WA : \bar{X} \pm z_{1-\alpha/2} \sqrt{\bar{X}/n} \quad (2.3)$$

포아송분포는 평균과 분산이 일치하는 성질을 이용하면 분산을 1차 적률로 추정한 신뢰구간을 구축할 수 있으나, 분산을 2차 적률로 추정하면 또 다른 신뢰구간을 얻을 수 있다. 즉 표본분산  $S^2 = \sum (X_i - \bar{X})^2 / (n-1)$  을 이용하여 다음과 같은 Wald 신뢰구간(W2)을 구축할 수 있다.

$$W2 : \bar{X} \pm z_{1-\alpha/2} S/\sqrt{n} \quad (2.4)$$

그런데, Wald 신뢰구간(WA)처럼 분산을 추정하는 것이 아니라, Wilson(1927)이 지적한 바처럼, 분산을 그대로 모수  $\mu$  로 추정하여, 2차 방정식을 풀면 신뢰구간을 구할 수 있다. 이를 Score 신뢰구간(SC)이라 한다. Score 신뢰구간(SC)은 다음과 같다.

$$SC : \bar{x} \left[ 1 + \frac{1}{2x} z_{1-\alpha/2}^2 \left[ 1 \pm \left( 1 + \frac{4x}{z_{1-\alpha/2}^2} \right)^{1/2} \right] \right] \quad (2.5)$$

한편, 위의 Score 신뢰구간에 연속성 수정을 가하여 다른 신뢰구간을 구축할 수 있다. 비교적 큰 모수  $\nu$ 에 대해 포아송 확률을 다음의 정규근사로 표현할 수 있다.

$$\sum_{k=0}^x (e^{-\nu} \nu^k) / k! = P \left( Z < \frac{x + 1/2 - \nu}{\sqrt{\nu}} \right)$$

위의 관계로부터 아래와 같은 다른 형태의 Score 신뢰구간(S2)을 얻을 수 있다. 즉,

$$\begin{aligned} S2(L) : \bar{x} - \frac{1}{2n} + \frac{1}{2n} z_{1-\alpha/2}^2 - z_{1-\alpha/2} \sqrt{\frac{\bar{x}}{n} - \frac{1}{2n^2} + \frac{z_{1-\alpha/2}^2}{4n^2}} \\ S2(U) : \bar{x} + \frac{1}{2n} + \frac{1}{2n} z_{1-\alpha/2}^2 + z_{1-\alpha/2} \sqrt{\frac{\bar{x}}{n} + \frac{1}{2n^2} + \frac{z_{1-\alpha/2}^2}{4n^2}} \end{aligned} \quad (2.6)$$

또한, 연속성 수정의 근사보다 좀 더 정확한 근사식은 Sahai and Khurshid(1993)의 Molenaar 근사이다. 이에 의하면, Sahai and Khurshid(1993)의 Molenaar 신뢰구간(MO)은 다음과 같다.

$$MO : \bar{x} + \frac{2z_{\alpha/2}^2 + 1}{6n} \pm \left[ \frac{1}{2n} + \frac{1}{n} \sqrt{z_{\alpha/2}^2 \left( x \pm \frac{1}{2} + \frac{z_{\alpha/2}^2 + 2}{18} \right)} \right] \quad (2.7)$$

Molenaar 근사식에 의해 신뢰구간을 구축하면 연속성수정보다 포아송분포의 누적확률에 더 유사하므로 정확신뢰구간과 매우 비슷한 신뢰구간을 구축하게 될 것이다. 신뢰구간을 구하는 또 다른 방법으로 테일러 전개와 분산안정화 방법을 활용하는 것이다. 이에 대해 본 연구에서는 Sahai and Khurshid(1993)의 Bartlett 방법(BA), Anscombe 방법(AN), Freeman and Tukey 방법(FT), Hald 방법(HA)을 고려하였다. 각 방법의 신뢰구간은 아래와 같다.

$$BA : \frac{1}{n} \left( \sqrt{x} \pm \frac{1}{2} z_{\alpha/2} \right)^2 \quad (2.8)$$

$$AN : \frac{1}{n} \left( \sqrt{x + \frac{3}{8}} \pm \frac{1}{2} z_{\alpha/2} \right)^2 - \frac{3}{8n} \quad (2.9)$$

$$FT : \frac{1}{4n} \left[ \sqrt{x + \sqrt{x+1}} \pm z_{\alpha/2} \right]^2 - 1 \quad (2.10)$$

$$HA : \frac{1}{n} \left[ \sqrt{x - \frac{1}{2}} \pm \frac{z_{\alpha/2}}{2} \right]^2 + \frac{1}{2n} \tag{2.11}$$

마지막으로, 붓스트랩 신뢰구간을 생각해 보자. 모수  $\mu$ 에 대한 붓스트랩 추정량을  $\hat{\mu}^*$  라고 하면, B 회 붓스트랩 반복에 의해 얻은 붓스트랩 분포로부터 붓스트랩 신뢰구간을 얻을 수 있다. 즉,  $100(1-\alpha)\%$  붓스트랩 신뢰구간(BT)은 다음과 같다.

$$BT : [ \bar{X} - q_{(1-\alpha/2)} \sqrt{\bar{X}/n}, \bar{X} - q_{(\alpha/2)} \sqrt{\bar{X}/n} ] \tag{2.12}$$

여기서,  $q_{(\alpha)}$ 는 다음의 B회 모수적 붓스트랩 반복에 의해 얻은  $Z^*$ 의  $100(1-\alpha)\%$ 분위수이다.

$$Z^*(b) = \frac{\sqrt{n}(\hat{\mu}^* - \bar{X})}{\sqrt{\bar{X}/n}}, \quad b=1, \dots, B$$

또한, 식 (2.4)와 같은 방법으로 다음과 같은 붓스트랩 신뢰구간(B2)을 구할 수 있다.

$$B2 : [ \bar{X} - q_{(1-\alpha/2)}^* S/\sqrt{n}, \bar{X} - q_{(\alpha/2)}^* S/\sqrt{n} ] \tag{2.13}$$

여기서,  $q_{(\alpha)}^*$ 는 다음의 B회 모수적 붓스트랩 반복에 의해 얻은  $Z^{**}$ 의  $100(1-\alpha)\%$ 분위수이다.

$$Z^{**}(b) = \frac{\sqrt{n}(\hat{\mu}^* - \bar{X})}{S^*/\sqrt{n}}, \quad b=1, \dots, B$$

이때  $S^*$  는 붓스트랩표본에서 얻은 붓스트랩 표준편차를 의미한다. 식 (2.4)와 식 (2.13)의 신뢰구간은 평균포함확률측면에서 두 신뢰구간이 지니는 특징을 살피는데 활용될 것이다.

### 2.2 상대오차

본 연구에서는 여러 가지 유형의 신뢰구간들을 비교하는 척도로서 상대오차(relative error)를 계산하였다. 비교의 기준이 되는 방법으로는 식 (2.1)의 정확신뢰구간을 선정하였다.  $EX_*$ 를 정확신뢰구간의 상한이라 하고  $P_*$ 를 각 방법의 신뢰구간의 상한이라고 할 때 정확신뢰구간과 다른 신뢰구간과의 상대오차를 다음과 같이 계산하였다.

$$\text{Relative error} = 100 \left[ \frac{(1 + P_*) - (1 + EX_*)}{(1 + EX_*)} \right]$$

신뢰구간의 하한에 대한 상대오차도 같은 방법으로 계산된다. 각 신뢰구간에 1을 더한 이유는 1이하의 신뢰구간이 존재하기 때문이다.

### 2.3 평균포함확률

표본의 크기가 1일 때, 임의의 모수  $\mu$  에 대해, 식 (2.4)와 식 (2.13)를 제외한 11가지 방법에 대해 정확한 포함확률을 계산할 수 있다. 즉, 포함확률은 다음과 같다.

$$\text{Coverage} : C(\mu) = \sum_{k=0}^{\infty} I(k, \mu) \frac{e^{-\mu} \mu^k}{k!}$$

여기서,  $I(k, \mu)$  는  $X=k$  에서, 신뢰구간이 주어진 모수  $\mu$ 를 포함하면 1, 그렇지 않으면 0이 되는 지시함수이다. 본 연구에서는 임의의 모수  $\mu$ 에 대한 평균적 수행능력을 평가하기 위해, 모수  $\mu$ 에 사전 분포  $g(\mu)$ 를 가정하여 다음을 계산하였다.

$$\bar{C}(\mu) = \int_{\Omega} C(\mu)g(\mu)d\mu \tag{2.14}$$

Agresti and Coull(1998)은 식 (2.14)를 평균포함확률(mean coverage probability)이라 칭하였으며, 이를 이항분포 신뢰구간의 비교에 활용한 바 있다. 평균포함확률측면에서 붓스트랩이 지니는 이론적 특징을 살펴보자.  $X_1, X_2, \dots, X_n$  을 평균이  $\theta$  이고 분산이  $\sigma^2$ 인 격자분포에서 생성된 확률변수라 하자.  $Z_n = \sqrt{n}(\bar{X}_n - \theta)/\sigma$  의 분포함수를  $G_n(\omega, t) = P_{\omega}\{Z_n \leq t\}$ 로 나타내자. 여기서  $\omega$ 는 모수공간  $\Omega$ 에서 정의되는 알려지지 않은 모수이다. 이제,  $t \in \mathbf{R}$  의 정수부분을  $[t]$ , 소수부분을  $\langle t \rangle = t - [t]$ 로 표시하면,  $G_n(\omega, t)$ 은 다음과 같이 Edgeworth 연장된다.

$$G_n(\omega, t) = \Phi(t) + \frac{1}{\sqrt{n}} \varphi(t) \left[ \frac{\rho}{6} (1-t^2) + R_n(\omega, t) \right] + o\left(\frac{1}{\sqrt{n}}\right)$$

여기서  $\Phi$ 는 표준정규분포함수이고,  $\varphi$ 는 표준정규확률밀도함수이며,  $\rho$ 와  $R_n(\omega, t)$ 는 다음과 같다.

$$\rho = \rho(\omega) = \varphi'''(\omega)/\sigma^3, R_n(\omega, t) = \frac{1}{\sigma} \left\{ \frac{1}{2} - \langle n\theta + \sqrt{n}\sigma t \rangle \right\}$$

이제,  $G_n^*(\omega, t) = \Phi(t) + \frac{\rho}{6\sqrt{n}}(1-t^2)\varphi(t)$ 라 하면, 다음이 성립한다.

$$\lim_n \int_{\Omega} \sqrt{n} [G_n(\omega, t) - G_n^*(\omega, t)] \xi(\omega) d(\omega) = 0$$

모수  $\omega$ 의 최대우도추정량을  $\hat{\omega}_n$ 이라 할 때  $Z_n$ 의 붓스트랩 분포함수를  $G_n(\hat{\omega}_n, t)$ 과 같이 표현하면, Woodrooffe and Jhun(1989)에 의해 붓스트랩분포와  $G_n^*(\omega, t)$ 와 차의 기대값은 다음과 같다.

$$E_{\omega}\{G_n(t) - G_n^*(\omega, t)\} = \frac{1}{\sigma\sqrt{n}} \varphi(t) e\left(\langle t\sqrt{n}\sigma \rangle, \frac{1}{2} \rho\sigma t\right) + o\left(\frac{1}{\sqrt{n}}\right)$$

여기서  $e(m, r) = \sum_{k=1}^{\infty} \frac{1}{k\pi} \int_{\mathbf{R}} \sin(2\pi kt) \frac{1}{r} \varphi\left(\frac{t-m}{r}\right) dt = \sum_{k=1}^{\infty} \frac{1}{k\pi} \sin(2\pi mk) e^{-2r^2\pi^2 k^2}$  이며, 실질적인  $e(m, r)$ 의 크기는  $\sup_m \left| e\left(m, \frac{1}{2}\right) \right| < 0.01, \sup_m |e(m, 1)| < 10^{-9}$ 으로  $r$ 이 증가함에 따라 매우 작아지는 함수이다. 그러므로 붓스트랩분포와  $G_n^*(\omega, t)$ 와 차이의 기대값은 실제적으로  $o(1/\sqrt{n})$ 에 근사하게 된다. 이는 평균포함확률을 계산하면, 붓스트랩근사가 정규근사에 비해 이론적 우월성을 격자분포인 경우에서도 그대로 지니고 있음을 나타내는 결과라 하겠다.

본 연구에서는  $g(\mu)$ 의 분포로 Uniform(1,50)분포를 사용하였다. 그런데, 표본의 크기가 1 이상일 때 정확한 포함확률을 계산하기 어려우므로 정확한 포함확률을 추정하여 평균포함확률을 계산하였다. 이에 대해서는 사전분포로 Normal(25, 25)와 Uniform(1,50)을 사용하였다.

### 2.3 신뢰구간의 기대 폭

신뢰구간의 폭이란 신뢰구간의 상한과 하한의 차이이다. 본 연구에서 신뢰구간의 기대 폭(expected widths)을 다음과 같이 계산하였다.

$$E_{\mu}(w) = \sum_{k=0}^{\infty} W_k \frac{e^{-\mu} \mu^k}{k!}$$

여기서,  $W_k$  는  $X=k$  일 때 신뢰구간의 폭을 나타낸다. 신뢰구간의 성격상 카이제곱분포를 정규 근사로 활용한 신뢰구간(CH), Molenaar 신뢰구간(MO) 그리고 정확신뢰구간(EX)의 기대 폭은 일치하게 된다.

### 3. 모의실험 결과

#### 3.1 평균포함확률

[표 1]은 사전분포가 Uniform(1,50)일 때, 다양한 표본의 크기에 대해, 평균포함확률을 명목수준 95%에서 계산한 결과를 제시하고 있다.

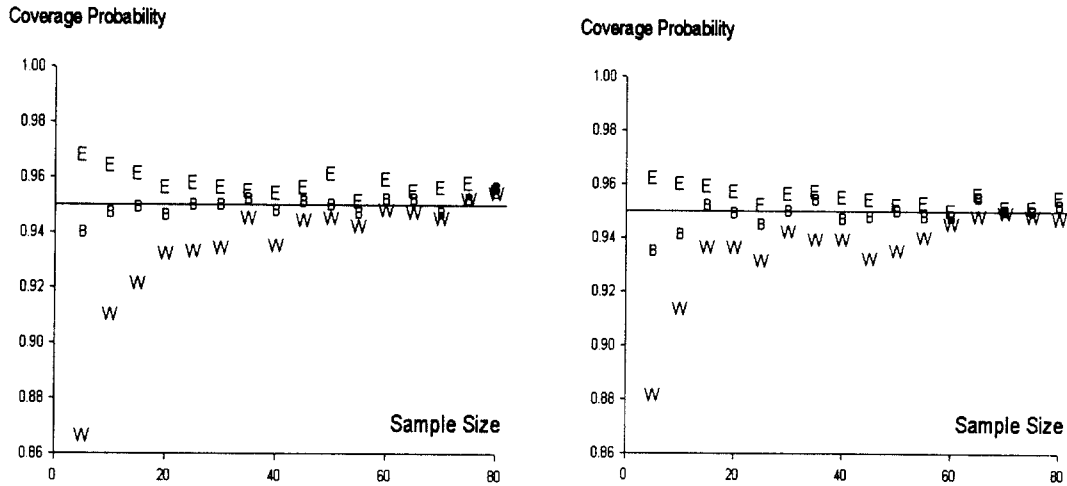
[표 1] 각 방법들에 대한 명목수준 95%에서의 평균포함확률  
(모의실험횟수 1000회, 붓스트랩 반복 2000회)

표본(n)	WA	W2	EX	SC	S2	MO	CH	BA	AN	FT	HA	BT	B2
5	.948	.881	.962	.953	.961	.962	.962	.953	.955	.953	.951	.952	.939
15	.951	.936	.959	.958	.963	.959	.959	.949	.949	.951	.949	.955	.952
30	.953	.942	.956	.954	.955	.956	.956	.955	.955	.956	.955	.951	.950
50	.946	.935	.952	.950	.952	.952	.952	.950	.950	.952	.950	.950	.950

[표 1]을 살펴보면, 2차 적률을 이용하여 분산을 추정하는 Wald 방법(W2)의 평균포함확률이 명목수준을 가장 하향 추정하고 있음을 볼 수 있다. 즉 n=5에서 0.881, n=15에서 0.936등 하향 추정되는 경향이 보인다. 반대로 정확신뢰구간(EX)은 주어진 명목수준보다 과대추정 되고 있음을 볼 수 있다. 이는 정확신뢰구간이 지니는 보수적인 성향을 보여준다고 하겠다. 또한, 정확방법(EX)의 평균포함확률이 Molenaar의 근사방법(MO), 그리고 식 (2.2)의 방법(CH)과 일치함을 볼 수 있다. 의외의 결과로, 이항분포에서 가장 정확한 방법으로 알려진 Score 방법(SC)이 포아송 모수 추정에서는 약간 과대추정 되는 경향을 보이고 있음을 볼 수 있다. 위의 표에서는 붓스트랩 방법(B2)이 명목수준에 상당히 근사함을 볼 수 있다.

평균포함확률 측면에서 붓스트랩 방법이 지니는 의미를 보다 명확히 파악하여 보자. Woodrooffe and Jhun(1989)의 결과는, 표본의 크기가 커질수록, 2차 적률을 이용하여 분산을 추정한 붓스트랩 주축통계량의 분포가 정규분포에 기초한 분포보다 수렴이 빠름을 의미하므로, 본 연구에서는 분산을 2차 적률로 추정한 Wald 방법(W2)와 이와 유사한 통계량인 붓스트랩 방법(B2)을 서로 비교하여 제시하였다. [그림 1]은 표본의 크기가 커질수록 Wald 방법(W2), 정확방법(EX) 그리고 붓스트랩방법(B2)의 평균포함확률을 보여주고 있다. [그림 1]에 의하면, 붓스트랩 방법이 명목수준에 수렴하는 속도가 정규근사에 의한 방법보다 빠름을 보여 주고 있다. 또한 정확방법(EX)은 주어진 95% 명목수준보다 위에서 수렴하는 형태를 보여주고 있다. [그림 1]에서 붓스트랩 방법이 지니는 이론적 우월성과 더불어 붓스트랩의 활용가능성을 살펴볼 수 있겠다. 그런데 표본의 크기가 클 때, 즉 실제 문제의 응용에서는 W2 보다 WA의 방법이 또 B2 보다는 BT의 방법이 계산의 편리성과 함께 수렴속도가 더 빠르다고 하겠다. 특히 WA는 W2보다 훨씬 수렴속도가 빠름을 여러 실

험과 경험을 통하여 파악 할 수 있었다.

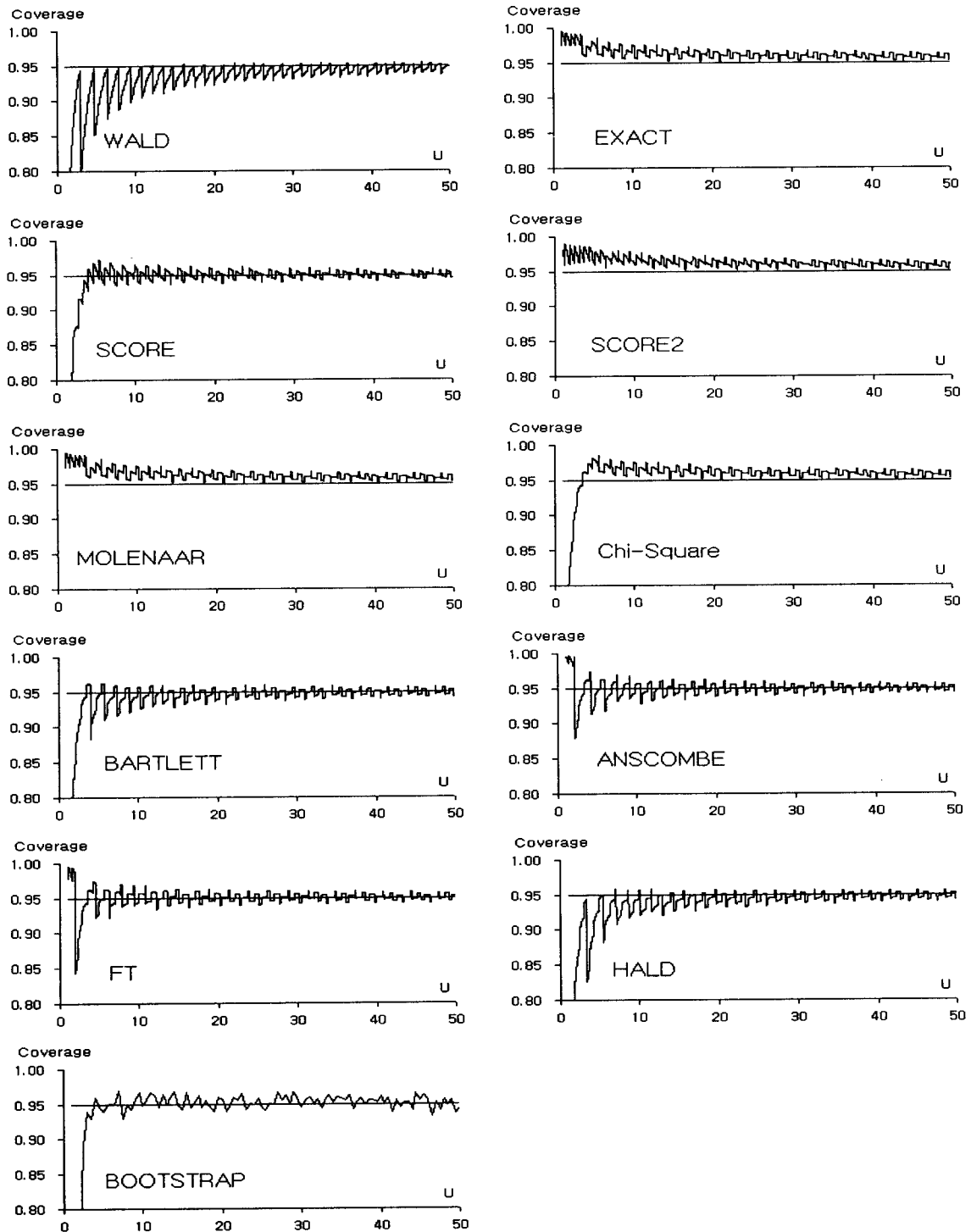


(a) 사전분포 : Normal(25,25)

(b) 사전분포 : Uniform(1,50)

[그림 1] 표본의 크기에 따른 평균포함확률, 명목수준 95%  
( 정확방법:E, Wald 방법 :W, B2의 붓스트랩방법:B )

[그림 2]는  $n=1$ 에서 모수  $\mu$ 에 따른 W2와 B2를 제외한 11가지 방법에 대한 실제 포함확률(actual coverage probability)을 보여주고 있다. [그림 2]를 통하여 각 방법들의 특징을 살펴보자. [그림 2]는 많은 내용을 시사하고 있다. 우선 정규근사에 기초한 Wald 방법의 포함확률은 톱날 모양을 취하면서 모수가 커짐에 따라 아래에서 수렴하는 형태를 취하고 있다. 즉 대부분의 통계학 책에서 추천하는 정규근사에 의한 Wald 방법은 하향추정하는 문제점을 확실히 보여주고 있다. 이와 비슷한 모양은 분산안정화를 통한 정규근사에 의존하는 Sahai and Khurshid(1993)의 Bartlett 방법(BA)과 Hald(HA) 방법이다. 이 두 가지 방법들 역시 명목수준보다 아래에서 수렴하는 형태이나 톱날의 모양이 Wald 방법에 비해 날카롭지 않음을 볼 수 있다. Wald 방법에 비해, Sahai and Khurshid(1993)의 Anscombe 방법(AN)이나 Freeman and Tukey 방법(FT)은 평균적으로 양 톱날이 균형이 잡힌 형태를 보이고 있으나 모수가 작을 때는 상향추정과 하향추정을 급격히 반복하는 형태를 보여주고 있다. 이들 방법은 모수가 어느 정도 클 때는 좋은 방법으로 보인다. 정확방법(EX)은 그림에서 나타나듯이 주어진 명목수준을 상향추정하면서 모수가 커짐에 따라 명목수준에 수렴하는 형태를 보여주고 있다. 정확방법과 완벽히 같은 모양을 취하는 것은 Sahai and Khurshid(1993)의 Molenaar 방법이다. Molenaar 방법은 누적 포아송분포에 가장 근사한 수렴식을 이용하여 신뢰구간을 구축하는 방법이므로 정확방법과 거의 같은 모양을 취함을 알 수 있다. 식 (2.2)의 카이제곱분포를 정규근사에 이용하는 방법(CH) 역시 모수 5이상에서 정확방법과 거의 일치함을 볼 수 있다. 단 모수가 작을 때는 수렴에 문제가 있어서 기존의 정규근사에 의존하는 방법과 비슷하게 하향추정하고 있다. Score 방법은 모수가 5이상부터 매우 안정적임을 볼 수 있다. 단 모수가 작을 때는 정규근사에 의존하는 수렴의 형태처럼 하향추정하는 경향을 보여주고 있다. 이를 연속성 수정한 방법(S2)은 정확방법과 매우 유사해짐을 볼 수 있다. 그러므로 특별한 경우 Score 방법에 연속성 수정을 할 필요는 없어 보인다.



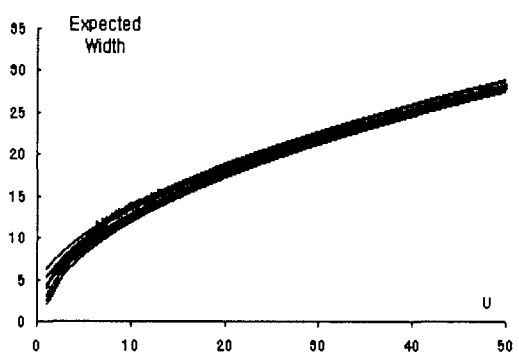
[그림 2] 11가지 방법에 대한 명목수준 95%에서 실제포함확률  
 (연속성수정을 한 SCORE 방법:SCORE2, 식(2.2) 방법:Chi square, Freeman and Tukey 방법 :FT)



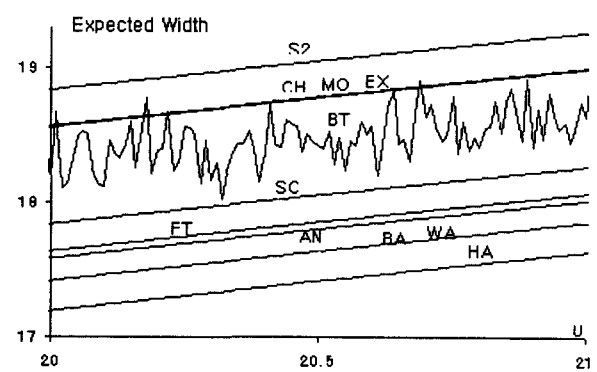
마지막으로 붓스트랩 방법은 특별한 모양을 보여주지 않으면서 변동이 일정함을 볼 수 있다. 즉, 어떤 유형을 따르지 않는데 모수가 작을 때는 붓스트랩 표본이 0근처에 있는 경우가 발생하므로 정규근사처럼 하향추정되는 경향을 피할 수 없어 보인다. 앞에서 살펴보았듯이 붓스트랩방법은 평균포함확률 측면에서 장점을 지니고 있다.

### 3.2 신뢰구간의 기대 폭

좋은 신뢰구간을 선택하는 하나의 방법으로 신뢰구간이 너무 넓지 않기를 바랄 것이다. 그러므로 3.1 절의 포함확률의 그림을 통한 좋은 신뢰구간을 선택하는 방법에 대한 하나의 대안으로 신뢰구간의 기대폭을 생각할 수 있다. [그림 3]의 (a)는 표본의 크기 1에서 모수  $\mu$ 에 따른 신뢰구간들의 기대폭을 보여 주고 있다. 여기서 모수  $\mu$ 에 따라 신뢰구간의 폭은 비례적으로 증가하며 11가지 방법이 하나의 폭에 좁은 띠를 형성하고 있음을 볼 수 있다. 이에  $\mu$ 가 20~21일 때 (a)를 확대한 것이 [그림 3]의 (b)이다. [그림 3]의 (b)는 각 신뢰구간들의 특징을 확실히 보여준다고 하겠다. 우선 신뢰구간의 폭은 Hald 방법(HA)이 가장 좁으며, 연속성 수정을 가한 Score 방법(S2)의 신뢰구간의 폭이 가장 넓음을 볼 수 있다. 11가지 방법들은 신뢰구간의 폭에 있어서 일정한 순서를 유지하고 있는데, 중간에 형성되는 신뢰구간은 Score 방법(SC)이며, 붓스트랩 방법 역시 정확신뢰구간과 Score 신뢰구간 사이에서 변동을 지니면서 존재하고 있으며, 가끔 정확신뢰구간의 폭과 일치함을 볼 수 있다. 특이한 것은 붓스트랩 신뢰구간의 기대 폭은 모수  $\mu$ 가 커질수록 변동의 변화가 아래로 내려가면서 Score 신뢰구간의 폭과 일치되는 경향이 나타난다는 점이다. 신뢰구간의 폭이 너무 좁으면 주어진 명목수준에 비해 하향 추정되는 경향이 있으므로 중간에 존재하는 신뢰구간을 다소 선호한다고 하겠다. 그런데, 모수가 상당히 크다면 (이는 표본의 크기와 비례한다) 신뢰구간의 폭이 좁은 방법을 선호하는 것도 좋은 방법이라고 하겠다. 그 이유는 포아송 신뢰구간을 형성하는 모든 방법들이 모수가 클수록 주어진 명목수준에 수렴하는 경향이 있기 때문이다.



(a)  $\mu$ 에 따른 11가지 방법의 기대 폭

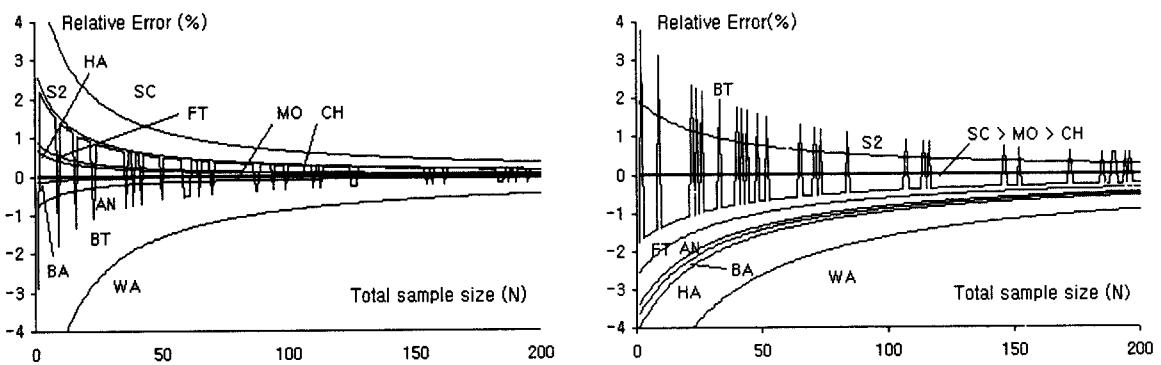


(b)  $\mu$ 가 20-21일 때 (a)의 확대

[그림 3] 표본의 크기가 1일 때 모수에 따른 95% 신뢰구간의 기대폭

### 3.3 상대오차

자료를 분석하다보면 주어진 신뢰구간이 정확신뢰구간에 얼마나 근사한가에 관심을 둘 수 있다(Vollset, 1993). 본 연구에서는 근사이론에 기초한 신뢰구간과 정확신뢰구간을 비교하기 위해 상대오차를 계산하였다. 상대오차를 측정하기 위하여 2가지 상황을 설정하였는데, 하나는 표본의 합이 일정하게 주어질 때 표본의 크기가 증가되는 경우이다. 이때는 표본의 크기가 커짐에 따라 모수  $\mu$ 에 대한 추정은 0으로 근사하게 된다. 또 다른 하나는 모수에 대한 추정값  $\hat{\mu}$ 가 2로 일정하게 주어질 때 표본의 크기가 증가되는 경우이다.



(a) 하한에 대한 상대오차

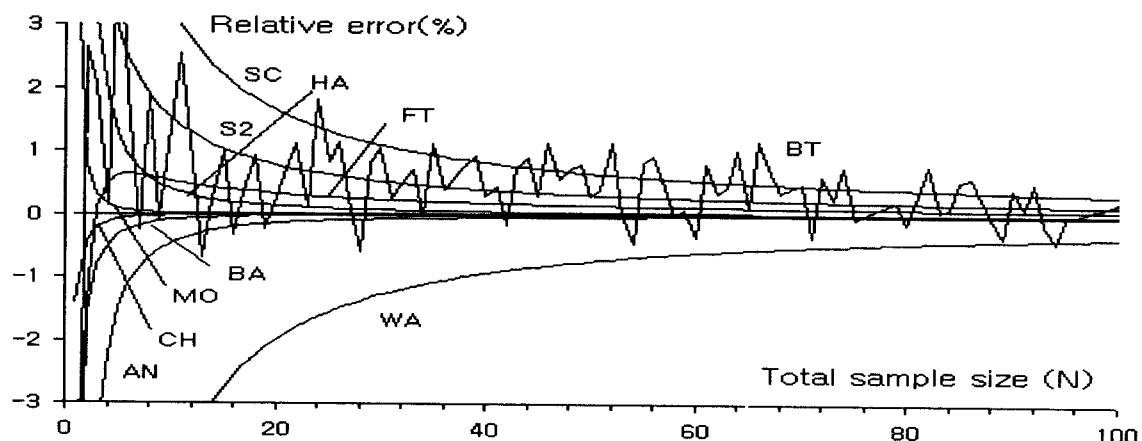
(b) 상한에 대한 상대오차

[그림 4] 표본의 합  $X$ 가 20일 때 표본의 크기에 따른 정확방법과 비교한 상대오차

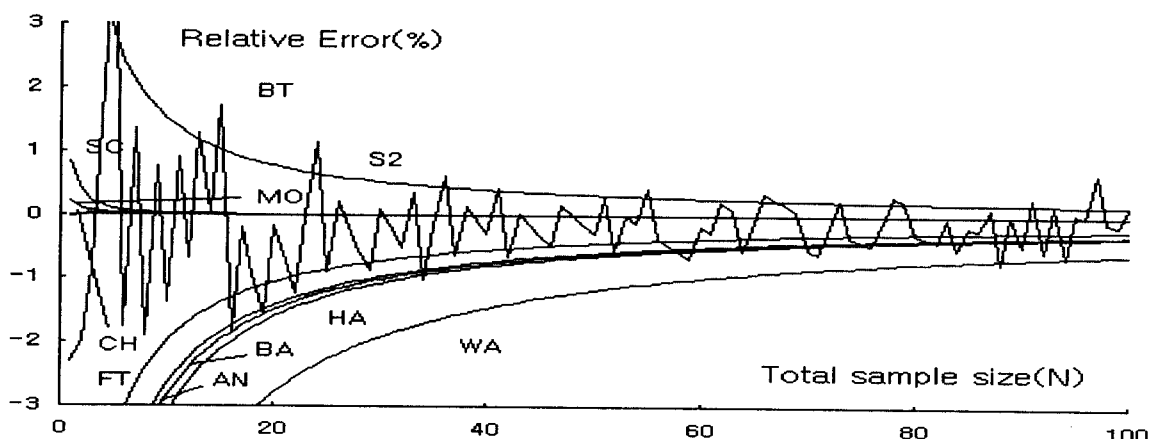
[그림 4]는 표본의 합  $X$ 가 20일 때, 하한과 상한에 대한 상대오차를 보여주고 있다. 여기서, 상대오차가 양의 값을 지니며 위에 형성되면 신뢰구간의 상한이나 하한이 정확신뢰구간보다 위에 있으며, 음의 값을 취할 때는 신뢰구간의 상한이나 하한이 정확신뢰구간보다 아래에 있음을 나타낸다고 하겠다. 우선적으로 표본의 크기가 커짐에 따라 모든 신뢰구간은 상대오차가 거의 0%에 수렴하게 된다.

[그림 4]를 통해서 붓스트랩 신뢰구간을 제외한 신뢰구간들이 정확신뢰구간과 비교하여 특정한 형태를 따른다는 점을 발견할 수 있다. 우선 Score 방법(SC), 연속성 수정을 한 Score 방법(S2), Freeman and Tukey 방법(FT) 그리고 Hald 방법(HA)은 정확방법에 비해 신뢰구간의 하한이 정확신뢰구간보다 위에 형성되며, Anscombe 방법(AN)과 Wald 방법(WA)은 정확신뢰구간보다 신뢰구간의 하한이 더 아래에 있음을 볼 수 있다. 특히 Wald 방법(WA)은 상한도 정확신뢰구간보다 아래에 있으므로 전체적으로 정확신뢰구간보다 신뢰구간이 아래에서 형성되고 있음을 볼 수 있다. 이에 반해 연속성 수정을 한 Score 방법(S2)은 정확신뢰구간보다 신뢰구간이 위에 형성되고 있음을 볼 수 있다. 이는 정확신뢰구간이 가장 표준적인 신뢰구간을 형성했다고 가정하면, Wald 신뢰구간은 아래 부분에서, 연속성 수정한 Score 방법(S2)은 위 부분에 신뢰구간이 형성되어 왜도 현상이 있다는 것을 의미하고 있다. 신뢰구간이 정확신뢰구간에 항상 포함되는 신뢰구간으로는 Freeman and Tukey 신뢰구간(FT)을 들 수 있다. 이는 Freeman and Tukey 방법(FT)의 상대오차는 하한은 위에, 상한은 아래에 형성되는 모습을 취하므로 이와같은 성질을 관찰할 수 있겠다.

또한, 상한에서 Score 방법(SC)과 Molenaar 방법(MO) 그리고 카이제곱분포를 정규근사 시킨 방법(CH)은 상한이 정확신뢰구간과 거의 일치하는 현상을 보여주고 있다. 부스트랩 신뢰구간은 틱낄과 같은 모양을 취하면서 양과 음의 값을 이동하는 현상이 발생함을 알 수 있다.



(a) 하한에 대한 상대오차



(b) 상한에 대한 상대오차

[그림 5] 표본의 크기에 따라  $\hat{\mu}$  가 2일 때, 정확방법과 비교한 상대오차

[그림 5]는 모수에 대한 추정값  $\hat{\mu}$ 가 2로 고정되었을 때 표본의 크기에 따른 상대오차를 나타내고 있다. 이 역시 [그림 4]와 신뢰구간이 형성되는 위치에 대해 별 차이를 보이지 않는다. 다만 부스트랩 신뢰구간의 변동이 좀더 활동적으로 보일 뿐이며, 표본의 크기가 커짐에 따라 상대오차가 0으로 수렴하는 속도가 [그림 4]보다 더 빠름을 볼 수 있다. 앞 절에서 관찰된 현상과 비슷한 결과이지만 정확신뢰구간과 비교하여 상한과 하한 모두에서 Molenaar 방법(MO)과 카이제곱분포를

정규근사 시킨 방법(CH)은 정확신뢰구간과 거의 같으며, 상한에서는 Score 방법(SC)이 정확방법과 일치한다는 사실을 발견 할 수 있다.

#### 4. 토의와 결론

이항분포에서 Vollset(1993), Agresti and Coull(1998)등에 의해 Wald 방법이 지니는 문제점이 많이 지적되어 왔으며 그 대안으로 Wilson(1927)에 의한 Score 방법들이 주로 추천되어 왔다. 또 Mehta and Walsh(1992)등이 고려한 오즈비(odds ratio)의 비교나 이항분포의 두 표본의 차에 대해 연구한 Newcombe(1998)등을 찾아볼 수 있다. 그러나 포아송분포에 있어서 Agresti and Coull(1998)가 이항분포와 비슷한 결과를 유도하리라는 언급을 제외하고 구체적인 포아송 신뢰구간에 대한 연구가 많지 않은 실정이다. 이에 본 연구에서는 Sahai and Khurshid(1993)의 연구에 기초하여 여러 포아송 신뢰구간들의 성질을 모의실험으로 비교하였다. 특히 모수에 대한 사전분포를 가정하여 붓스트랩 신뢰구간이 지니는 특징을 평균포함확률측면에서 살펴보았다. 물론 본 연구에서 다루지 않은 신뢰구간도 존재한다(Casella and Robert ;1987). 본 연구를 통해 이항분포에서 처럼 정확신뢰구간이 지니는 보수성을 확인 할 수 있었고, Wald 신뢰구간이 지니는 하향추정의 문제점도 확인 할 수 있었다. 물론, 이산형 자료에서 표본의 크기가 작은 경우 정확한 추론은 매우 중요한 역할을 한다는 점을 간과할 수 없다. 그런데 표본의 크기가 적당히 크면 근사이론에 기초한 통계적 추론이 경우에 따라 정확하게 작용한다는 점이다. 또한 Woodrooffe and Jhun(1989)의 결과에 의해, 평균포함확률 측면에서 붓스트랩 신뢰구간이 이산형 분포에서도 훌륭히 활용될 수 있다는 점을 주목하여야 한다. 본 연구를 통해 다음의 몇 가지 결과와 제안을 제시하고자 한다.

첫째, Wald 방법, Bartlett 방법 그리고 Hald 방법은 하향추정하는 경향이 있다. 둘째, 정확방법, 연속성 수정된 Score 방법, Molenaar 방법, 식 (2.2)의 정규근사방법(CH)은 상향추정을 하며 정확신뢰구간과 거의 일치한다. 그러므로 특별히 Score 방법에 연속성 수정을 가할 필요는 없다고 할 수 있다. 셋째, 모수가 5이상 일 때, Score 방법이나 Sahai and Khurshid(1993)의 Freeman and Tukey방법을 사용하길 권한다. 넷째, 붓스트랩 방법은 Score 방법이 지니는 이론적 어려움을 피하여 주며, Wald 방법에 대한 훌륭한 대안이 될 수 있다. 특히 표본의 크기에 따라 적절하게 Score 방법과 정확방법의 신뢰구간의 유형을 따라주며, 평균포함확률이 주어진 명목수준에 매우 일치한다는 점에서 모수에 대한 어떤 정보가 주어지지 않은 경우에 적절히 활용하길 권한다.

#### 참고문헌

- [1] Agresti, A. and Coull, B.(1998) Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician*, Vol. 52, 119-126.
- [2] Casella, G. and Robert, C.(1987) Refining Poisson Confidence Intervals. *Canadian Journal of Statistics*, Vol. 17, 45-57.
- [3] Efron, B.(1979) Bootstrap Methods : Another Look at the Jackknife. *The annals of Statistics*, Vol. 7, 1-26.

- [4] Mehta, C.R., and Walsh, S.J.(1992) Comparison of Exact, Mid-p and Mantel-Haenszel Confidence Intervals for the Common Odds Ratio Across Several  $2 \times 2$  Contingency Tables. *The American Statistician*, Vol. 46, 146-150.
- [5] Newcombe, R.G.(1998) Interval Estimation for the Difference Between Independent Proportions: Comparison of Eleven Methods. *Statistics in Medicine* Vol 17, 873-890.
- [6] Sahai, H. and Khurshid, A.(1993) Confidence Intervals for the Mean of a Poisson Distribution: A Review. *Biometrical Journal* vol 7. 857-67
- [7] Singh, K.(1981) On the asymptotic accuracy of Efron's bootstrap, *The Annals of Statistics*, Vol. 9, 1187-1195.
- [8] Vollset, S.E.(1993) Confidence Intervals for a Binomial Proportions, *Statistics in Medicine*, Vol 12, 809-824.
- [9] Wilson, E.B.(1927) Probable inference, the law of succession, and statistical inference, *Journal of the American Statistical Association*, Vol. 22, 209-212.
- [10] Woodroffe, M. and Jhun M.(1989) Singh's theorem in the lattice case. *Statistics Probability Letters*, Vol. 7, 201-205.