Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 2, pp. 465 ~ 472

Bootstrap Confidence Intervals for a One Parameter Model using Multinomial Sampling

Hyeong Chul Jeong [1] · Dae Hak Kim [2]

Abstract

We considered a bootstrap method for constructing confidenc intervals for a one parameter model using multinomial sampling. The convergence rates of the proposed bootstrap method are calculated for model-based maximum likelihood estimators(MLE) using multinomial sampling. Monte Carlo simulation was used to compare the performance of bootstrap methods with normal approximations in terms of the average coverage probability criterion.

Key Words and Phrases: Bootstrap, Model-based maximum likelihood estimator, Average Coverage Probability, One parameter model, Hardy-Weinberg proportions, multinomial sampling

1. Introduction

The bootstrap is a widely used method in many areas of statistics that can be used to approximate the sampling distribution of some statistical quantity of interest. We will consider the convergence rates of the bootstrap method for model-based maximum likelihood estimators(MLE) in the lattice case.

In the non-lattice case, Singh (1981) showed that the bootstrap estimator differs from the actual one by an order of magnitude smaller than $1/\sqrt{n}$ with probability one as $n \to \infty$. In the lattice case, Woodroofe and Jhun(1988) showed that in terms of average coverage probability, the bootstrap estimator differs from the very weak expansions by a term of order $1/\sqrt{n}$. However, it was also shown that the coefficient of the term is very small in many examples. The actual convergence rate of the bootstrap method is $o(1/\sqrt{n})$ in an average coverage probability sense. This

[1]Full time Instructor, Department of Computer Science & Statistics, Pyongtaek University, Pyongtaek, Kyungki, 450-701, Korea
[2]Associate Professor, Department of Statistics, Catholic University of Taegu-Hyosung, Kyoungsan, Kyungpuk, 713-702, Korea

is a meaningful result because loglinear models and other models using multinomial sampling also belong to the underlying lattice distribution.

In this paper, we extend the results of Woodroofe and Jhun(1989) to the one parameter model $\boldsymbol{\pi} = \boldsymbol{\pi}(\theta)$, where $\boldsymbol{\pi}$ are multinomial cell probabilities $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_K)'$ and θ denotes an unknown parameter. In section 2, we introduce the results of Woodroofe and Jhun(1989). In section 3, we consider the asymptotic behavior of the model-based estimator $\boldsymbol{\pi}$. However this study is limited to the one parameter model. The Hardy-Weinberg model is a good example of a one parameter model using multinomial sampling. We considered a simulation study that compares the bootstrap method with a normal approximation in section 4. The results demonstrate the efficiency of the bootstrap method using average coverage probability in the Hardy-Weinberg problem.

2. The Bootstrap Method in the Lattice Case

In this section, we introduce the results from Woodroofe and Jhun used later in section 3.

Let $X_1, X_2, \cdots, X_n$ denote i.i.d. integer valued random variables with common discrete density of the form

$$f_\omega(x) = h(x)\exp[\omega x - \psi(\omega)], \quad x \in \mathbf{I}, \omega \in \Omega$$

where $\mathbf{I}$ denotes the integers, h is a non-negative function and ω denotes an unknown parameter with values in an open interval Ω. Then, the mean and variance of X are

$$\theta = \psi'(\omega), \quad \text{and } \sigma^2(\theta) = \psi''(\omega).$$

Let $\bar{X}_n$ denote the sample mean, and S_n denote the sample sum. Now, consider the distribution of

$$Z_n = \frac{\sqrt{n}}{\sigma}(\bar{X}_n - \theta) = \frac{S_n - n\theta}{\sigma\sqrt{n}}, n \geq 1. \tag{2.1}$$

Let the distribution of Z_n denote $G_n(\omega, t) = P_\omega\{Z_n \leq t\}$, where $t \in \mathbf{R}$, $\omega \in \Omega$ and $n \geq 1$. Let $\hat{\omega}_n = \hat{\omega}_n(X_1, \cdots, X_n)$ denote the maximum likelihood estimator of ω. The parametric bootstrap estimator of the distribution of Z_n is represented by

$$\hat{G}_n(t) = G_n(\hat{\omega}_n, t), t \in \mathbf{R}, n \geq 1.$$

If $t \in \mathbf{R}$, denote its integer part by $[t]$ and denote the fractional part by $< t >= t - [t]$. By Edgeworth expansion, the distribution of Z_n could be expressed as follows

$$G_n(\omega, t) = \Phi(t) + \frac{1}{\sqrt{n}}\varphi(t)\left[\frac{\rho}{6}(1 - t^2) + R_n(\omega, t)\right] + o\left(\frac{1}{\sqrt{n}}\right), \tag{2.2}$$

where Φ is the standard normal distribution function, φ denotes the standard normal density, $\rho = \rho(\omega) = \psi'''(\omega)/\sigma^3$ and $R_n(\omega, t) = \sigma^{-1}\left\{\frac{1}{2} - < n\theta + \sqrt{n}\sigma t >\right\}$
By Edgeworth expansion, the parametric bootstrap distribution $\hat{G}_n(t)$ is

$$G_n(\hat{\omega}_n, t) = \Phi(t) + \frac{1}{\sqrt{n}}\varphi(t)\left[\frac{\hat{\rho}_n}{6}(1-t^2) + R_n(\hat{\omega}_n, t)\right] + o\left(\frac{1}{\sqrt{n}}\right), \qquad (2.3)$$

where $\hat{\rho}_n = \rho(\hat{\omega}_n)$, $R_n(\hat{\omega}_n, t) = \hat{\sigma}_n^{-1}\left\{\frac{1}{2} - < n\hat{\theta}_n + \sqrt{n}\hat{\sigma}_n t >\right\}$, $\hat{\theta}_n = \psi'(\hat{\omega}_n)$, and $\hat{\sigma}_n = \sigma(\hat{\theta}_n) = \sqrt{\psi''(\hat{\omega}_n)}$. Singh's(1981) theorem shows that $\hat{G}_n$ and $G(\omega, \cdot)$ differ by $O(1/\sqrt{n})$ with probability 1 in the non-lattice case.

Theorem (Woodroofe and Jhun; 1989)
For $t \in \mathbf{R}$, $\omega \in \Omega$, $n \geq 1$, $r > 0$ and $0 \leq \mathrm{m} < 1$ and all densities ξ with compact support in Ω, let $G_n^*(\omega, t) = \Phi(t) + \frac{\rho}{6\sqrt{n}}(1-t^2)\varphi(t)$, and $e(\mathrm{m}, r) = \int_{\mathbf{R}}[\frac{1}{2} - < m + rs >]\varphi(s)ds$. Then

(1) $\lim_n \int_\Omega \sqrt{n}\,[G_n(\omega, t) - G_n^*(\omega, t)]\,\xi(\omega)d(\omega) = 0$
(2) $\lim_n \sqrt{n}\left[\hat{G}_n(t) - G_n^*(\omega, t)\right] = \sigma^{-1}\varphi(t)\left\{\frac{1}{2} - << t\sqrt{n}\sigma > + \frac{1}{2}\rho\sigma t Z_n >\right\} + o_p(1)$
(3) $E_\omega\left\{\hat{G}_n(t) - G_n^*(\omega, t)\right\} = \frac{1}{\sigma\sqrt{n}}\varphi(t)e\left(< t\sqrt{n}\sigma >, \frac{1}{2}\rho\sigma t\right) + o(\frac{1}{\sqrt{n}})$
Proof. See details in Woodroofe and Jhun(1989). □

Woodroofe and Jhun(1989) argued that if $G(\omega, t)$ is regarded as the coverage probability of a confidence set at ω, then $\int_\Omega G_n(\omega, t)\xi(\omega)d\omega$ may be regarded as the long run relative frequency of coverage in many independent replications of the experiment, when ω is drawn from the density ξ. Therefore, $\int_\Omega G_n(\omega, t)\xi(\omega)d\omega$ can be called the average coverage probability at ξ. Agresti and Brent(1998) compared several interval estimation methods for binomial proportions using average coverage probability, which they call mean coverage probability.

The function e decreases quickly as r increases. For example $\sup_m |e(\mathrm{m}, \frac{1}{2})| <$ 0.01, $\sup_m |e(\mathrm{m}, 1)| < 10^{-9}$. In the cases where $\sup_m |e(\mathrm{m}, \rho\sigma/2)|$ is small, the expected value of the bootstrap estimator should be close to $o(1/\sqrt{n})$. This implies that the bootstrap method is more efficient than the approximation of the limiting normal distribution in the lattice case.

3. The Bootstrap Method for the One Parameter Model

In this section we investigate the asymptotic behavior of the parametric bootstrap estimator of the model-based estimator $\hat{\boldsymbol{\pi}}$, which is a function of θ. We express it as $\boldsymbol{\pi} = \boldsymbol{\pi}(\theta)$, where $\boldsymbol{\pi}$ are cell probabilities $\boldsymbol{\pi} = (\pi_1, \cdots, \pi_{\mathrm{K}})'$ and θ denotes an unknown parameter with values in the parameter space Θ. We use θ and $\boldsymbol{\pi}$ to denote generic parameter and probability values and, θ_0 and $\boldsymbol{\pi}_0 = (\pi_{i0}, \cdots, \pi_{K0})$ to denote true values for a particular application.

Let $(n_1, n_2, \cdots, n_K)$ be a random sample from a multinomial distribution with cell probabilities $\boldsymbol{\pi}(\theta)$. Let $(n_1^*, n_2^*, \cdots, n_K^*)$ be a parametric bootstrap random sample from multinomial distribution with cell probabilities $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi}(\hat{\theta})$, where $\hat{\theta}$ is the maximum likelihood estimator of θ. $\hat{\theta}^*$ denotes the maximum likelihood estimator of $\hat{\theta}$ from a bootstrap sample. Let $\hat{\boldsymbol{\pi}}^*$ denote maximum likelihood estimator of $\hat{\boldsymbol{\pi}}$ from a bootstrap sample, p the sample proportion from raw counts and p^* the bootstrap sample proportion.

Then the likelihood equation of θ can be

$$L(\theta) = \log \prod_{i=1}^{K} \pi_i(\theta)^{n_i} = n \sum_{i=1}^{K} p_i \log \pi_i(\theta),$$

$$\frac{\partial L(\theta)}{\partial \theta} = n \sum_i \frac{p_i}{\pi_i(\theta)} \left(\frac{\partial \pi_i(\theta)}{\partial \theta} \right) = 0, \tag{3.1}$$

$$\sum_i \frac{\partial \pi_i(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left[\sum_i \pi_i(\theta) \right] = \frac{\partial}{\partial \theta}(1) = 0. \tag{3.2}$$

Subtracting a common term from (3.1) and (3.2), we obtain

$$\sum_i \frac{n(p_i - \pi_{i0})}{\hat{\pi}_i} \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right) = \sum_i \frac{n(\hat{\pi}_i - \pi_{i0})}{\hat{\pi}_i} \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right). \tag{3.3}$$

We use the Taylor-series expansion

$$\hat{\pi}_i - \pi_{i0} = (\hat{\theta} - \theta) \left(\frac{\partial \pi_i}{\partial \bar{\theta}} \right) = (\hat{\theta} - \theta) \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right) + o_p \left(\frac{1}{\sqrt{n}} \right)$$

where some point $\bar{\theta}$ falls between $\hat{\theta}$ and θ_0. The size of the remainder term follows from $(\hat{\theta} - \theta_0) = O_p(n^{-1/2})$. Therefore, from (3.3)

$$\sum_i \frac{\sqrt{n}(p_i - \pi_{i0})}{\hat{\pi}_i} \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right) = \sqrt{n}(\hat{\theta} - \theta) \left\{ \sum_i \frac{1}{\hat{\pi}_i} \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right) \left(\frac{\partial \pi_i}{\partial \bar{\theta}} \right) \right\}$$

$$= \sqrt{n}(\hat{\theta} - \theta) \sum_i \frac{1}{\hat{\pi}_i} \left(\frac{\partial \pi_i}{\partial \hat{\theta}} \right)^2 + o_p(1). \tag{3.4}$$

Let $\mathbf{A}$ denote the vector having elements $a_i = \pi_{i0}^{-1/2}(\partial \pi_i / \partial \theta_0)$. Then the K $\times$ 1 matrix $\mathbf{A}$ be $\mathbf{Diag}(\boldsymbol{\pi}_0)^{-1/2}(\partial \boldsymbol{\pi} / \partial \theta_0)$, where $(\partial \boldsymbol{\pi} / \partial \theta_0)$ denotes the Jacobian $(\partial \boldsymbol{\pi} / \partial \theta)$ evaluated at θ_0. As $\hat{\theta} \to \theta_0$, (3.4) has form

$$\sqrt{n}(\hat{\theta} - \theta) = (\mathbf{A}'\mathbf{A})^{-1} \sum_{i=1}^{K} \frac{\sqrt{n}(p_i - \pi_{i0})}{\pi_{i0}} \left(\frac{\partial \pi_i}{\partial \theta_0} \right) + o_p(1).$$

Therefore we have

$$\begin{aligned}\frac{\sqrt{n}(\hat{\theta}-\theta)}{\sigma(\hat{\theta})} &= (\mathbf{A}'\mathbf{A})^{-1}\sum_{i=1}^{K}\frac{\sqrt{n}(p_i-\pi_i)}{\sqrt{\pi_i(1-\pi_i)}}\frac{\sqrt{\pi_i(1-\pi_i)}}{\sigma(\hat{\theta})}\left(\frac{\partial\pi_i}{\partial\theta_0}\right)\frac{1}{\pi_i}+o_p(1)\\ &= (\mathbf{A}'\mathbf{A})^{-1}\sum_{i=1}^{K}\frac{\sqrt{n}(p_i-\pi_i)}{\sqrt{\pi_i(1-\pi_i)}}C_i+o_p(1)\end{aligned} \tag{3.5}$$

where $C_i=\sqrt{(1-\pi_i)/\pi_i}(1/\sigma(\hat{\theta}))(\partial\pi_i/\partial\theta_0)$. For sufficiently large n, we have

$$\begin{aligned}P\left[\frac{\sqrt{n}(\hat{\theta}-\theta)}{\sigma(\hat{\theta})}\le t\right] &= P\left[(\mathbf{A}'\mathbf{A})^{-1}\sum_{i=1}^{K}\frac{\sqrt{n}(p_i-\pi_i)}{\sqrt{\pi_i(1-\pi_i)}}C_i\le t\right]\\ &= P\left[\sum_{i=1}^{K}\frac{\sqrt{n}(p_i-\pi_i)}{\sqrt{\pi_i(1-\pi_i)}}C_i\le t^*\right]\end{aligned} \tag{3.6}$$

where $t^*=(\mathbf{A}'\mathbf{A})t$. Since $\sqrt{n}(\mathbf{p}-\boldsymbol{\pi})\xrightarrow{d}\mathrm{N}(\mathbf{0},\Sigma)$ in distribution, we also have

$$B=\left[\frac{\sqrt{n}(p_1-\pi_1)}{\sqrt{\pi_1(1-\pi_1)}},\frac{\sqrt{n}(p_2-\pi_2)}{\sqrt{\pi_2(1-\pi_2)}},\cdots,\frac{\sqrt{n}(p_K-\pi_K)}{\sqrt{\pi_K(1-\pi_K)}}\right]\xrightarrow{d}\mathrm{N}(\mathbf{0},\Sigma^*)$$

where Σ^* is the variance-covariance matrix as

$$\begin{pmatrix}1 & -\sqrt{\frac{\pi_1\pi_2}{(1-\pi_1)(1-\pi_2)}} & \cdots & -\sqrt{\frac{\pi_1\pi_K}{(1-\pi_1)(1-\pi_K)}}\\ & 1 & \cdots & -\sqrt{\frac{\pi_2\pi_K}{(1-\pi_2)(1-\pi_K)}}\\ & & \cdots & \\ & & & 1\end{pmatrix}.$$

Therefore, we have

$$C'B=\sum_{i=1}^{K}C_i\frac{\sqrt{n}(p_i-\pi_i)}{\sqrt{\pi_i(1-\pi_i)}}\xrightarrow{d}\mathrm{N}(\mathbf{0},C'\Sigma^*C)\text{ in distribution.}$$

Let $B_i=\sqrt{n}(p_i-\pi_i)/\sqrt{\pi_i(1-\pi_i)}$, and let $S^2=C'\Sigma^*C$. By the property of exponential families, the equation (3.6) is

$$\begin{aligned}P\left[\sum_{i=1}^{K}B_iC_i\le t^*\right] &= P\left[\sum_{i=1}^{K}\frac{B_iC_i}{S}\le\frac{t^*}{S}\right]=P\left[\sum_{i=1}^{K}\frac{B_iC_i}{S}\le t^{**}\right]\\ &= \Phi(t^{**})+\frac{1}{\sqrt{n}}\varphi(t^{**})\left[\frac{\rho}{6}(1-t^{2^{**}})+R_n(\omega,t^{**})\right]+o\left(\frac{1}{\sqrt{n}}\right),\end{aligned} \tag{3.7}$$

where $t^{**} = t^*/S$, ρ denotes the third moment of $\sum B_i C_i$ and ω denotes an unknown parameter with values in $\sum B_i C_i$.

Corollary 3.1 For all $t \in \mathbf{R}$, all ω defined on the parameter space of $\sum B_i C_i$, $n \geq 1$, and all $\theta \in \Theta$,

$$\begin{aligned}&\lim_n \sqrt{n}\left[G_n(\hat{\theta}, t) - G_n^*(\theta, t)\right] \\ &= S^{-1}\varphi(t^{**})\left\{\frac{1}{2} - << t^{**}\sqrt{n}S > + \frac{1}{2}\rho S t^{**} Z_n >\right\} + o_p(1)\end{aligned}$$

in P_ω -probability, where $t^{**} = (\mathbf{A}'\mathbf{A}/S)t$, $S^2 = C'\Sigma^* C$ and

$$G_n^*(\theta, t) = \Phi(t^{**}) + \frac{\rho}{6\sqrt{n}}(1 - t^{2^{**}})\varphi(t^{**}).$$

Proof. For sufficiently large n, we have

$$\begin{aligned} G_n(\hat{\theta}, t) &= P\left[\frac{\sqrt{n}(\hat{\theta}^* - \hat{\theta})}{\sigma(\hat{\theta})} \leq t\right] \\ &= P\left[\sum_{i=1}^{k} \frac{\sqrt{n}(p_i^* - \pi_i(\hat{\theta}))}{\sqrt{\pi_i(1 - \pi_i)}} C_i^* \leq t^*\right] \\ &= \Phi(t^{**}) + \frac{1}{\sqrt{n}}\varphi(t^{**})\left[\frac{\hat{\rho}_n(1 - t^{2^{**}})}{6} + R_n(\hat{\omega}, t^{**})\right] + o\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Then

$$\sqrt{n}\left[G_n(\hat{\theta}, t) - G_n^*(\theta, t)\right] = \frac{1}{6}(\hat{\rho}_n - \rho)(1 - t^{2^{**}})\phi(t^{**}) + \phi(t) R_n(\hat{\omega}_n, t^{**})] + o(1).$$

Since $\lim_n P\left(|\hat{\rho}_n - \rho| > \epsilon\right) = 0$, $\lim_n \left|\frac{1}{6}(\hat{\rho}_n - \rho)(1 - t^{2^{**}})\phi(t^{**})\right| = 0$.
From Theorem A(2), we have

$$R_n(\hat{\omega}_n, t^{**}) = \frac{1}{S}\left\{\frac{1}{2} - << t^{**}\sqrt{n}S > + \frac{1}{2}\rho S t^{**} Z_n >\right\} + o_p(1). \quad \square$$

Corollary 3.2 For $r > 0$ and $0 \leq \mathrm{m} < 1$,

$$\mathrm{E}_\theta\{\hat{G}_n(t) - G_n^*(\theta, t)\} = \frac{1}{S\sqrt{n}}\varphi(t^{**})\mathrm{e}(< t^{**}\sqrt{n}S),\ \frac{1}{2}\rho S t^{**}) + o(\frac{1}{\sqrt{n}})$$

as $n \to \infty$ for all $t \in \mathbf{R}$, all ω defined on the parameter space of $\sum B_i C_i$, $n \geq 1$, and all densities of $\theta \in \Theta$.
Proof. The proof follows from Corollary 3.1. $\square$

This implies that the average coverage probability can be applied to model-based estimators. Examples of model-based estimators can be found in the Hardy-Weinberg problem or the loglinear model. In these models, the bootstrap distribution for model-based parameters is closer to the true distribution than the approximation from the normal distribution.

4. Simulation Study

In this section the simulation results from the Hardy-Weinberg problem are presented. Let the random sample $\{n_i : i = 1, 2, 3\}$ be a random variable from the multinomial distribution with cell probabilities

$$\pi_1 = \theta^2, \ \pi_2 = 2\theta(1-\theta), \ \pi_3 = (1-\theta)^2, \ 0 < \theta < 1.$$

The cell probabilities are called Hardy-Weinberg proportions.

Let $\hat{\mathbf{A}}$ denote $\mathbf{Diag}(\boldsymbol{\pi}_0)^{-1/2}(\partial\boldsymbol{\pi}/\partial\theta_0)$ evaluated at the ML estimate $\hat{\theta}$. The maximum likelihood estimate of θ and the estimated variance are

$$\hat{\theta} = \frac{2n_1 + n_2}{2n}, \quad \hat{Cov}(\hat{\theta}) = \frac{(\hat{\mathbf{A}}'\hat{\mathbf{A}})^{-1}}{n}.$$

To compute the average coverage probability, we consider two prior distributions of θ as

$$1 : \theta \sim \text{ Uniform } (\frac{1}{5}, \frac{4}{5}) \quad 2 : \theta \sim \text{ Beta } (5, 5).$$

For comparison of the bootstrap method and the normal method, we compute the confidence interval by using an approximation of the normal theorem

$$\left[\hat{\theta} - z_{(1-\alpha/2)}\sqrt{\hat{\mathbf{A}}'\hat{\mathbf{A}}/n}, \ \hat{\theta} - z_{(\alpha/2)}\sqrt{\hat{\mathbf{A}}'\hat{\mathbf{A}}/n}\right] \tag{4.1}$$

and the bootstrap confidence interval

$$\left[\hat{\theta} - Q_{(1-\alpha/2)}\sqrt{\hat{\mathbf{A}}'\hat{\mathbf{A}}/n}, \ \hat{\theta} - Q_{(\alpha/2)}\sqrt{\hat{\mathbf{A}}'\hat{\mathbf{A}}/n}\right]. \tag{4.2}$$

Here $Q_{(1-\alpha/2)}$ and $Q_{(\alpha/2)}$ are, respectively the $100(1-\alpha/2)$th and $100(\alpha/2)$th percentile points of a bootstrap distribution computed by

$$Z^*(b) = \frac{\sqrt{n}(\hat{\theta}^* - \hat{\theta})}{\sqrt{(\hat{\mathbf{A}}^{*\prime}\hat{\mathbf{A}}^*)^{-1}}}, \quad b = 1, 2, \cdots, B$$

where $(\hat{\mathbf{A}}^{*\prime}\hat{\mathbf{A}}^*)^{-1}/n$ denotes the estimated variance of bootstrap samples.

For the simulation study, consider sample sizes $n = 10, 20, 30, 40, 50$, and nominal coverages $\alpha = 0.8, 0.9, 0.95, 0.99$, and bootstrap replication B=1000 times, and

Table 4.1 Estimated Average Coverage Probabilities for the Confidence interval of θ. Normal approximation (CLT) and Bootstrap refer to (4.1) and (4.2), respectively.

		θ : Uniform(0.2,0.8)				θ : Beta(5,5)			
n	$1-\alpha$	0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
10	CLT	0.768	0.864	0.915	0.954	0.773	0.867	0.919	0.966
	BOOT	0.845	0.919	0.957	0.973	0.849	0.915	0.958	0.975
20	CLT	0.771	0.882	0.925	0.967	0.796	0.886	0.933	0.978
	BOOT	0.822	0.928	0.968	0.989	0.831	0.915	0.962	0.993
30	CLT	0.795	0.885	0.942	0.986	0.767	0.873	0.929	0.981
	BOOT	0.826	0.912	0.959	0.994	0.795	0.885	0.954	0.995
40	CLT	0.790	0.875	0.928	0.976	0.792	0.900	0.952	0.982
	BOOT	0.811	0.900	0.952	0.991	0.812	0.909	0.959	0.991
50	CLT	0.766	0.868	0.934	0.992	0.797	0.883	0.935	0.984
	BOOT	0.776	0.889	0.951	0.996	0.802	0.896	0.949	0.994

simulation replications M=1000 times. The simulation results are present in Tables 4.1. The results of Table 4.1 show the expected values of the bootstrap method for model-based ML estimators approximate the nominal coverages well. When the prior distribution is uniform, the average coverage probabilities of the normal approximation method are generally below the nominal coverage. When the sample size is 10, the expected coverage probability of the bootstrap method reaches the nominal coverage. But the normal approximation method does not. When the prior distribution is a Beta distribution, the same result can be obtained.

In conclusion, Woodroofe and Jhun's results are extended to the one parameter model using multinomial sampling. The expected value of the bootstrap estimator of the parameter of multinimial proportions is closer to true distribution than the approximation from normal theory.

References

1. Agresti, A. and Brent, A. C. (1998). Approximation is better than "exact" for interval estimation of binomial proportions, *The American statistician*, 52, 2,2 119-126.

2. Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap, *The Annals of Statistics*, 9, 1187-1195.

3. Woodroofe, M. and Jhun, M. (1989). Singh's theorem in the lattice case, *Statistics Probability Letters*, 7, 201-205.