

로짓모형의 비모수적 추론의 비교

정형철¹⁾ 김대학²⁾

요약

범주형 자료의 구조파악에 주로 이용되는 로짓모형에서 비모수적 방법을 이용한 모수의 신뢰구간추정과 가설검정 등의 통계적 추론에 대하여 살펴보았다. 모수에 대한 통계적 추론에서 정규분포에 근거한 모수적 방법(Wald 방법)보다는 붓스트랩 방법이나 임의순열을 활용한 비모수적 방법이 많이 활용되고 있다. 본 연구에서는 로짓모형의 모수에 대한 비모수적 추론방법으로 붓스트랩(bootstrap)과 임의순열(random permutation)의 두 방법을 고려하고 모의실험을 통하여 가설검정의 검정력과 신뢰구간 추정의 포함확률을 비교하였고 사례분석을 다루었다.

주요용어: 로짓모형, 붓스트랩, 임의순열, 신뢰구간추정, 포함확률, 제1종오류

1. 서론

Z 를 이항형 반응변수 ($Z = 0, 1$), X_1, \dots, X_p 를 p 개의 설명변수라 하자. 관측자료가

$$(z_1, x_{11}, \dots, x_{1p}), (z_2, x_{21}, \dots, x_{2p}), \dots, (z_n, x_{n1}, \dots, x_{np})$$

로 주어졌을 때, 성공($Z = 1$)의 조건부 확률을 $\pi(x) = P(Z = 1|x_1, \dots, x_p)$ 라고 놓자. 이때 로짓모형은

$$\text{logit}(\pi) = \log \frac{P(Z = 1|x_1, \dots, x_p)}{P(Z = 0|x_1, \dots, x_p)} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (1.1)$$

이다. 로짓모형에서 모수 $\beta_i, i = 1, \dots, p$ 의 가설검정과 신뢰구간 구축을 고려하자. 모수 β_i 에 대한 가설 $H_0 : \beta_i = 0$ 을 검정하는 방법으로 대표본 이론에 근거한 Wald 검정, 일반화 우도검정(generalized likelihood ratio test), 피셔의 점수검정(Fisher's score test) 등이 활용된다. Agresti(1995)는 Wald 검정법에서 $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$ 의 최우추정량 $\hat{\beta}$ 가 로그우도 함수의 모양에 영향을 받으며 $\hat{\beta}^2 / \text{Cov}(\hat{\beta})$ 의 형태로 카이제곱분포를 따름을 밝힌 바 있다.

본 연구에서는 이와같은 대표본 이론에 의존하지 않고 $\hat{\beta}$ 의 분포를 비모수적 방법으로 유도하여 가설 $H_0 : \beta_i = 0$ 에 대한 검정과 신뢰구간 구축을 다루고자한다. 대표적인 비모수적 방법으로 임의순열에 기초한 임의순열(random permutation)과 붓스트랩 방법을 들 수 있다. 임의순열 검정은 R.A Fisher(1935)가 제안한 이래로 Good(1994), Manly(1997) 등에

1) (450-701) 경기도 평택시 용이동 111, 평택대학교 정보통계학과, 조교수

E-mail: jhc@ptuniv.ac.kr

2) (712-702) 경북 경산시 하양읍 금락리 330, 대구가톨릭대학교 정보통계학과, 교수

E-mail: dhkim@cataegu.ac.kr

의해 많은 연구가 있어왔다. 특히 로짓모형에 대한 임의순열 검정방법으로 허명희(1998)는 Kennedy(1995)의 회귀분석에 대한 개념을 확장하여 제안한 바 있다. 또한, 임의순열 검정 방법과 경쟁적 입장에 있는 비모수적 방법은 붓스트랩 검정이라 할 수 있다. 로짓모형에 대한 붓스트랩 검정은 김대학, 정형철(2002)에 의해 다루어졌다. 이와같은 비모수적 방법은 자료특정적(data specific)이며 계산집중적(computer intensive)이다(허명희, 1998).

본 연구에서는 임의순열검정 방법과 붓스트랩 검정 방법들을 살펴보고 이들 비모수적 방법을 모의실험과 사례분석을 통하여 비교하고자 한다. 2절에서는 단순로짓모형에서 붓스트랩방법과 임의순열검정방법을 비교하고, 제 1종 오류(Type I error)와 검정력을 추정하고자 한다. 또한 모수의 신뢰구간을 비모수적으로 구축하는 방법을 소개하고, 신뢰구간의 포함확률(coverage probability)를 추정하고자 한다. 3절에서는 다중로짓모형의 경우로 확장하였다.

2. 단순로짓모형

2.1. 비모수적 유의확률 추정

단순로짓모형

$$\text{logit}(\pi) = \log \frac{P(Z = 1|x_1)}{P(Z = 0|x_1)} = \beta_0 + \beta_1 x_1$$

에서 x_1 의 계수 β_1 에 대한 가설 $H_0 : \beta_1 = 0$ 의 임의순열검정과 붓스트랩 검정을 생각하자. 로짓모형에서 임의순열검정은 허명희(1998)에 의해 제안되었는데 그 방법은 다음과 같다.

[단계 1] 관측자료에 단순로짓모형을 적합시켜 β_0, β_1 의 MLE $\hat{\beta}_0, \hat{\beta}_1$ 을 계산한다.

[단계 2] z_1, z_2, \dots, z_n 의 임의순열 $z_1^*, z_2^*, \dots, z_n^*$ 를 만들어 순서대로 공변량 x_1 에 결합하여 임의순열자료를 생성하고 β_0, β_1 의 MLE $\hat{\beta}_0^*$ 와 $\hat{\beta}_1^*$ 를 계산한다.

[단계 3] [단계2]을 M번 반복하여 양측검정에 의한 유의확률을 다음의 상대적 비율로 추정한다.

$$p\text{-value} = P\{|\hat{\beta}_1^*| \geq |\hat{\beta}_1|\}$$

로짓모형에서 β_1 에 대한 붓스트랩검정과 임의순열검정의 차이는 [단계2]의 대표본 추출에 있다. 김대학, 정형철(2002)에 의하면, 붓스트랩 검정을 위해 붓스트랩 표본을 생성할 때 베르누이분포 $B(1, \hat{P})$ 에서 n 개의 모수적 붓스트랩 표본 $\{z_j^*; j = 1, \dots, n\}$ 을 발생하게 되어 있다. 여기서 확률 \hat{P} 는 $\hat{P}(Z = 1) = \sum z_i/n$ 로 $\text{logit}(\pi) = \beta_0$ 에서 얻은 값이다. 또한, 붓스트랩 검정시 붓스트랩 분포 $\hat{\beta}_1^*$ 를 사용할 수 있으나, 표준화에 기초한 $\hat{\beta}_1^*$ 의 주축통계량에 의한 붓스트랩 검정을 실시할 수 있다. 이를 우리는 표준화붓스트랩검정이라 하겠다. 즉, 붓스트랩 검정은 H_0 하에서 x_1 의 각 수준에서 이항분포를 따른다는 모수적 성질을 활용한 비복원 랜덤표본을 사용하고, 임의순열검정은 H_0 하에서 임의자료가 관찰자료와 확률적으로 동등하다는 성질을 사용한 복원 랜덤표본을 사용한다.

2.2. 비모수적 신뢰구간 구축

모수 β_1 의 신뢰구간 구축에 대해 살펴보자. 붓스트랩 신뢰구간 구축은 다음과 같다.

[단계 1] 관측자료에 $logit(\pi) = \beta_0 + \beta_1 x_1$ 의 모형을 적합시켜 β_0, β_1 을 계산한 후 적합한 모형으로부터 각 반응범주 수준에서 성공확률 $\{\hat{\pi}_j; j = 1, \dots, n\}$ 를 추정한다.

[단계 2] $\{\hat{\pi}_j; j = 1, \dots, n\}$ 로부터 붓스트랩 표본 $\{z_j^*; j = 1, \dots, n\}$ 을 발생하여 붓스트랩 최우추정치 $\hat{\beta}_0^*$ 와 $\hat{\beta}_1^*$ 을 계산한다.

[단계 3] [단계2]를 B 회 반복하여 다음을 계산한다.

$$B^*(b) = \frac{\sqrt{n}(\hat{\beta}_1^* - \hat{\beta}_1)}{ASE(\hat{\beta}_1^*)}, \quad b = 1, 2, \dots, B$$

[단계 4] 모수 β_1 의 $100(1 - \alpha)\%$ 붓스트랩 신뢰구간

$$[L_1^{BOOT} = \hat{\beta}_1 - B_{(1-\alpha/2)}SE(\hat{\beta}_1), U_1^{BOOT} = \hat{\beta}_1 - B_{(\alpha/2)}SE(\hat{\beta}_1)] \quad (2.1)$$

을 구한다. 이때 $B_{(\alpha)}$ 는 붓스트랩 분포 $B^*(\cdot)$ 에서 $100 \times \alpha\%$ 에 해당되는 분위점이다.

붓스트랩 신뢰구간은 귀무가설에 의존하지 않는다는 점에서 붓스트랩 검정과 차이가 있다. 즉, x_1 의 각 수준에서 성공확률을 추정할 때, 검정에서는 모형 $logit(\pi) = \beta_0$ 를 사용하고 신뢰구간에서는 $logit(\pi) = \beta_0 + \beta_1 x_1$ 를 사용한다. 이에 대한 내용은 김대학, 정형철(2002)을 참고 할 수 있다. 한편, 임의순열자료에 의한 신뢰구간 구축 방법은 Good(1994), Garthwaite(1996) 등의 방법이 있는데 본 연구에서는 붓스트랩 방법과 유사한 방법으로 신뢰구간을 구축하고자 한다. 2.1 절에서 H_0 의 가설검정을 위해 실시한 임의순열자료가 H_0 하에서 관찰자료와 확률적으로 동등하며, H_0 하에서 $\hat{\beta}_1$ 의 임의순열 분포를 유도한다는 사실을 사용하여 임의순열 분포의 표준화된 분포를 신뢰구간 구축에 사용하고자 한다. 즉 $\hat{\beta}_1^*/ASE(\hat{\beta}_1^*)$ 의 표준화된 임의순열 분포를 $Q^*(\cdot)$ 라 놓으면 $100(1 - \alpha)\%$ 임의순열 신뢰구간은 다음과 같다.

$$[L_1^{Pert} = \hat{\beta}_1 - Q_{(1-\alpha/2)}SE(\hat{\beta}_1), U_1^{Pert} = \hat{\beta}_1 - Q_{(\alpha/2)}SE(\hat{\beta}_1)]. \quad (2.2)$$

여기서, $Q_{(1-\alpha)}$ 는 표준화된 임의순열 분포 $Q^*(\cdot)$ 에서 $100(1 - \alpha)\%$ 에 해당되는 분위점이다.

Wald 신뢰구간은 정규분포를 사용한다는 점에서 비모수적 신뢰구간과 차이가 있다. Wald의 신뢰구간은 다음과 같다.

$$[L_1^{Wald} = \hat{\beta}_1 - Z_{(1-\alpha/2)}SE(\hat{\beta}_1), U_1^{Wald} = \hat{\beta}_1 - Z_{(\alpha/2)}SE(\hat{\beta}_1)]. \quad (2.3)$$

여기서, $Z_{(1-\alpha)}$ 는 표준정규분포에서 $100(1 - \alpha)\%$ 에 해당되는 분위점이다.

표 2.1: Lee(1974)의 Cancer Remission 자료

관측번호	1	2	3	4	5	6	7	8	9
Remiss	1	1	0	0	1	0	1	0	0
Blast	1.100	0.740	0.176	1.053	0.519	1.230	1.354	0.322	0.000
Temp	0.996	0.992	0.982	0.986	0.980	0.982	0.992	1.020	0.999
관측번호	10	11	12	13	14	15	16	17	18
Remiss	0	0	0	0	0	0	1	0	0
Blast	0.000	0.279	0.146	0.380	0.114	1.037	2.064	0.114	1.322
Temp	1.038	0.988	0.982	1.006	0.990	0.990	1.020	1.014	1.004
관측번호	19	20	21	22	23	24	25	26	27
Remiss	0	1	0	0	1	0	1	1	0
Blast	0.114	1.072	0.176	1.591	0.531	0.886	0.964	0.398	0.398
Temp	0.990	0.986	1.010	1.020	1.002	0.998	0.990	0.986	0.986

2.3. 단순로짓모형의 사례분석

허명희(1998)의 임의순열 검정에서 사용된 cancer remission 자료를 이용하여 Wald 방법과 붓스트랩 그리고 임의순열방법을 비교하고자 한다. 표 2.1은 허명희(1998)에 의해 인용된 Lee(1974)의 cancer remission 자료이다. 이 자료는 각 공변량의 수준에서 반복이 1회 인 비그룹화 자료라 할 수 있다.

위 자료에서 변수 Remiss를 이항형 반응변수, Blast를 설명변수로 고려한 단순로짓모형을 생각해 보자. 자료를 적합하면, β_0 와 β_1 의 MLE는 $\hat{\beta}_0 = -1.826$, $\hat{\beta}_1 = 1.526$ 이며, β_1 에 대한 Wald의 양측 p -값은 0.0791이다. 그리고, 10,000번의 반복실험에 의한 임의순열검정 p -값은 0.0880, 붓스트랩 p -값은 0.1005(표준화 붓스트랩 p -값은 0.0901)으로 추정된다. 또한 Wald의 95% 신뢰구간은 (-0.1757, 3.2279), 임의순열 95% 신뢰구간은 (-0.1751, 3.0635) 그리고 붓스트랩 95% 신뢰구간은 (-0.1459, 3.0878)로 추정된다. 유의확률에 있어서는 붓스트랩 검정이 상대적으로 다른 두 검정에 비해 보수적인 결과를 보여준다. 또한 신뢰구간에서는 임의순열이나 붓스트랩 신뢰구간의 상한값이 Wald의 상한신뢰구간에 비해 낮은 수준에서 형성되고 있음을 볼 수 있다. 이는 $\hat{\beta}_1$ 의 표준화분포를 그린 그림 2.1에서 상한의 꼬리부분이 약간의 차이가 있음을 보여주는 결과라 하겠다. 그림 2.1에서는 세가지 분포간에 꼬리부분과 중앙에서 약간의 차이가 있음을 보여주고 있다. 신뢰구간의 폭에 있어서는 붓스트랩신뢰구간과 임의순열신뢰구간에 큰 차이가 발견되지 않으며, 상대적으로 Wald 신뢰구간보다는 좁게 형성됨을 볼 수 있다.

2.4. 단순로짓모형의 모의실험

본 절에서는 모의실험을 통하여 임의순열방법, 붓스트랩방법 그리고 표준화 붓스트랩방법의 검정력과 신뢰구간을 살펴보겠다. 모의실험을 위하여

모형1: $\text{logit}(\pi) = 0 + 0x_1$ 모형2: $\text{logit}(\pi) = -2.5 + 0.5x_1$ 모형3: $\text{logit}(\pi) = -5.0 + 1.0x_1$ 의 세가지 로짓모형을 사용하였다. 설명변수 x_1 의 수준은 1부터 10까지 정수의 값을 취하고

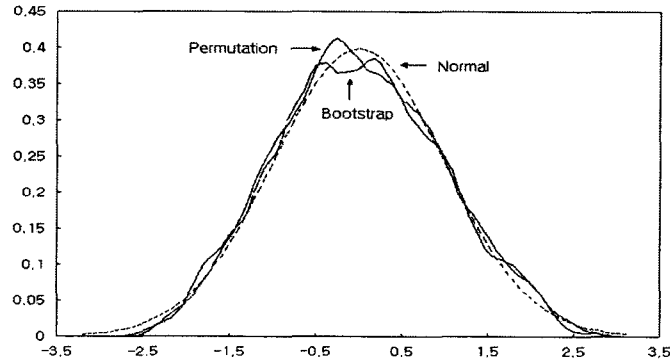


그림 2.1: 커널밀도함수 추정을 이용한 β_1 의 분포

표 2.2: 각 모형의 설명변수의 수준에 따른 성공의 기대확률

x_1	1	2	3	4	5	6	7	8	9	10
모형1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
모형2	0.119	0.182	0.269	0.378	0.500	0.622	0.731	0.818	0.881	0.924
모형3	0.018	0.047	0.119	0.269	0.500	0.731	0.881	0.953	0.982	0.993

록 하였다. 각 모형에서 검정력과 신뢰구간을 구하기 위해, 명목수준 $1-\alpha = 0.90, 0.95, 0.99$ 에서 표본의 크기는 $n = 10, 20, 30, 40, 50$ 으로 설정하였다. 이는 표본의 크기가 10인 경우를 제외하면, 주어진 x_i 의 수준에서 각각 2회, 3회, 4회, 5회의 반복 실험을 한 그룹화자료임을 의미한다.

1000번의 반복을 통한 경험적 검정력 결과는 표 2.3, 신뢰구간 결과는 표 2.4에 주어졌다. 표 2.3에서 Bootstrap1은 비표준화 붓스트랩 Bootstrap2는 표준화 붓스트랩을 나타내며 붓스트랩모형 1에 대한 검정력은 모수 β_1 에 대한 1종 오류를 추정한 것이다. 유의수준 0.05에서 1종 오류를 살펴보자. 여기서 Wald의 검정이 다소 보수적이며, 비표준화 붓스트랩 검정이 그 다음으로 보수적임을 알 수 있다. 그런데 임의순열검정은 다소 개방적(표본의 크기가 10일 때, 유의수준 0.05에서 0.051이며, 표본의 크기가 40, 50 일 때 유의수준 0.01에서 모두 0.012, 0.015로 유의수준을 넘게 됨)이며, 표준화 붓스트랩 방법이 가장 개방적인 입장임을 볼 수 있다. 또한 표본의 크기가 40이하에서 대체적으로 표준화 붓스트랩을 제외한 세 방법 모두 보수적 입장을 취하고 있음을 볼 수 있다. 이와같은 크기의 순은 다른 유의수준에서도 계속 유지되고 있다. 여기서 표본의 크기가 작을 때 ($n \leq 40$) 1종 오류를 조절하는 능력은 임의순열 검정이 정규근사 방법이나 비표준화붓스트랩 보다는 뛰어난을 볼 수 있다. 그리고 임의순열 검정과 표준화붓스트랩 간에는 큰 차이가 보이지 않는다. 특이한 점은 표준화 붓스트랩 검정이 비표준화 붓스트랩 검정보다 1종 오류를 조절하는 능력이 대부분의 유의수준에서 뛰어나다는 사실이다. 그런데, 이들 방법 역시 표본의 크기가

표 2.3: 경험적 검정력의 비교

모형 1에서의 β_1 에 대한 1종 오류(Type I error rates)												
n	$\alpha =$	Wald			Permutation			Bootstrap1			Bootstrap2	
		0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05
10	.024	.024	.057	.009	.051	.092	.014	.035	.076	.014	.051	.108
20	.000	.029	.081	.008	.046	.088	.006	.036	.084	.011	.056	.093
30	.001	.040	.086	.009	.049	.087	.009	.043	.084	.010	.051	.094
40	.006	.038	.092	.012	.044	.096	.006	.043	.089	.009	.048	.093
50	.009	.050	.104	.015	.050	.102	.011	.051	.099	.014	.053	.104
모형 2에서의 β_1 에 대한 경험적 검정력(power)												
n	$\alpha =$	Wald			Permutation			Bootstrap1			Bootstrap2	
		0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05
10	.191	.191	.352	.117	.413	.571	.168	.340	.484	.168	.345	.587
20	.017	.700	.858	.514	.767	.865	.427	.740	.854	.498	.779	.872
30	.623	.906	.963	.757	.919	.968	.733	.912	.965	.772	.925	.969
40	.864	.979	.991	.899	.980	.991	.896	.978	.992	.905	.981	.992
50	.956	.993	.999	.967	.996	.999	.959	.993	.999	.969	.995	.999
모형 3에서의 β_1 에 대한 경험적 검정력(power)												
n	$\alpha =$	Wald			Permutation			Bootstrap1			Bootstrap2	
		0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05
10	.525	.525	.681	.396	.846	.937	.492	.777	.899	.493	.671	.908
20	.173	.895	1.000	.976	1.000	1.000	.949	.999	1.000	.803	1.000	1.000
30	.735	1.000	1.000	.999	1.000	1.000	.999	1.000	1.000	.998	1.000	1.000
40	.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

50이 되면서 모두 1종 오류 수준에 근사함을 볼 수 있다.

표 2.3에서 모형2와 모형3의 결과는 경험적 검정력을 나타내고 있다. 모형2를 살펴보면 유의수준 0.05에서 네 방법간 검정력은 차이가 있음을 볼 수 있다. 특히, 표본의 크기가 10일 때 임의순열검정에 의한 검정력은 0.413으로 가장 높으며 다음으로 표준화부스트랩 검정력은 0.345, 비표준화부스트랩 검정력은 0.340임을 볼 수 있다. 이에 반해 Wald의 검정력은 0.191로 비모수적 방법에 비해 표본의 크기가 작을 때 현저히 검정력이 떨어짐을 볼 수 있다. 그리고 표준화 부스트랩의 검정력은 비표준화 부스트랩의 검정력에 비해 항상 높은 수준을 유지하고 있다. 또한 모형2의 경우, 표준화 부스트랩의 검정력이 대부분의 유의수준에서 다른 방법의 검정력에 비해 높음을 볼 수 있다. 그러나 모형1과 마찬가지로 표준화 부스트랩과 임의순열검정과 는 검정력에서 유의한 차이를 보여주지는 못한다고 하겠다. 결국 모형2에서도 표본의 크기에 커지면서 각 방법들의 검정력간에는 거의 차이가 없어지나, 표준화 부스트랩, 임의순열검정, 비표준화 부스트랩, Wald 방법 순으로 검정력이 유지되고 있다고 하겠다. 이제, 모형3을 살펴보자. 여기서 유의수준 0.01을 제외하면 모형3에서의 검정력 역시 모형2의 경우와 비슷하다고 하겠다. 이상에서 보면 표준화 부스트랩과 임의순열검정이 검정력이 높으며, 1종 오류를 조절하는 능력 역시 우수하다고 할 수 있으나,

표 2.4: 모형 1에서 신뢰구간의 포함확률

n	방법	0.90	폭	SE	0.95	폭	SE	0.99	폭	SE
10	Permutation	.888	1.460	.128	.927	1.605	.132	.971	3.580	.192
	Bootstrap	.885	1.129	.058	.941	1.997	.165	.978	4.378	.238
20	Permutation	.911	.556	.010	.951	.638	.010	.992	.776	.010
	Bootstrap	.910	.547	.003	.947	.634	.011	.989	.771	.018
30	Permutation	.903	.435	.001	.949	.507	.001	.991	.640	.001
	Bootstrap	.901	.433	.001	.948	.506	.001	.989	.637	.002
40	Permutation	.899	.373	.001	.953	.437	.001	.987	.560	.001
	Bootstrap	.904	.372	.001	.950	.436	.001	.993	.560	.001
50	Permutation	.889	.330	.000	.941	.389	.001	.987	.504	.001
	Bootstrap	.891	.331	.000	.947	.389	.001	.989	.504	.001

두 방법 간에는 거의 차이가 없다고 하겠다. 물론 붓스트랩 검정 내에서는 표준화 붓스트랩이 비표준화 붓스트랩보다는 우수하다고 하겠다.

이제, 신뢰구간에 대해 살펴보자. 신뢰구간에서는 주축통계량을 사용하기 때문에 검정에서처럼 서로 다른 붓스트랩방법이 존재하지 않는다. 김대학, 정형철(2002)에 의해 Wald 신뢰구간이 붓스트랩 신뢰구간에 비해 다소 보수적 즉, Wald의 포함확률이 붓스트랩의 포함확률에 비해 높은 수준임을 볼 수 있었다. 그러므로 표 2.4에서는 2.2절에서 언급한 (2.1)의 붓스트랩 포함확률과 (2.2)의 임의순열 포함확률을 제시하기로 한다. 표 2.4의 결과는 검정과 비슷하게 두 비모수적 방법간에 유의한 차이가 없음을 보여주고 있다. 또한 표본의 크기가 20이상부터 포함확률이 주어진 명목수준에 상당히 근사함으로 비모수적 방법이 상당히 잘 적용되고 있음도 볼 수 있다. 이는 두 방법 모두 정규분포로 분포수렴하는 성질이 있기 때문이다. 모형 2와 모형 3의 신뢰구간에서도 모형 1과 비슷한 현상이 발생하므로 여기서는 제시하지 않았다.

3. 다중로짓모형

3.1. 다중로짓모형에서 붓스트랩과 임의순열방법

본 장에서는 p 개의 설명변수가 있는 다중로짓모형에서 x_p 의 계수 β_p 에 대한 비모수적 검정과 신뢰구간 구축을 살펴보자. 허명희(1998)는 가설 $H_0 : \beta_p = 0$ 를 검정하기 위한 다음의 임의순열 검정을 제안하였다.

[단계 1] 관측자료에 다중로짓모형을 적합하여 $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ 를 계산한다. 또한 $H_0 : \beta_p = 0$ 하에서 다중로짓모형을 적합시킨다.

[단계 2] z_1, z_2, \dots, z_n 의 임의배열 $z_1^*, z_2^*, \dots, z_n^*$ 를 만들어 순서대로 반응변수 x_1, \dots, x_p 에 결합한다. H_0 하에서 임의자료의 실현확률은

$$\prod_{i=1}^n \pi(x_i)^{z_i^*} [1 - \pi(x_i)]^{1-z_i^*} \tag{3.1}$$

에 비례한다. 여기서 $\hat{\pi}(x_i)$ 는 $H_0: \beta_p = 0$ 하에서 확률의 추정값이다.

[단계 3] 임의배열자료를 다중로짓모형에 적합하여 MLE $\tilde{\beta}_0^*, \dots, \tilde{\beta}_p^*$ 를 계산한다.

[단계 4] 단계2와 단계3을 M번 반복하여 $|\tilde{\beta}_p^*| \geq |\hat{\beta}_p|$ 의 상대적 비율로 유의확률을 추정하되 (3.1)을 가중치 $\omega_i (i = 1, \dots, M)$ 로 고려하여 다음의 확률로 유의확률을 추정한다.

$$P_p = \left\{ \sum_{i=1}^M \omega_i I(|\tilde{\beta}_p^*| \geq |\hat{\beta}_p|) \right\} / \sum_{i=1}^M \omega_i$$

이와같은 임의순열검정방법은 H_0 하에서 임의자료의 발생확률을 고려한다는 점에서 단순로짓모형과 차이가 있다. 그러나, 단순로짓모형에서는 가중치가 상수로 결정되므로 다중로짓모형의 특수한 경우에 해당한다고 하겠다(허명희, 1998).

이제, H_0 하에서 김대학, 정형철(2002)의 붓스트랩 검정에 대해 살펴보자. 다중로짓모형에서의 붓스트랩 검정은 단순로짓모형에서의 붓스트랩 검정과 큰 차이가 없다. 즉, $\beta_p = 0$ 인 가정하에서, 추정된 $\text{logit}(\pi)$ 로부터 각각의 설명변수 x_1, x_2, \dots, x_{p-1} 수준에서 확률 $\pi(x_i)$ 를 $\hat{\pi}(x_i)$ 로 추정한다. 그리고 베르누이분포 $B(1, \hat{\pi}(x_i))$ 에서 n 개의 모수적 붓스트랩 표본 z_i^* 을 발생하여 붓스트랩 최우추정치 $\hat{\beta}_0^*, \dots, \hat{\beta}_p^*$ 를 계산하고 다음의 상대적 비율로 유의확률을 추정한다.

$$p\text{-value} = P\{|\hat{\beta}_p^*| \geq |\hat{\beta}_p|\}$$

상대적으로 붓스트랩 검정이 임의순열검정보다 가중치를 따로 계산할 필요가 없다는 점에서 장점을 지닌다고 하겠다. 붓스트랩검정에서 가중치를 고려할 필요가 없는 이유는 붓스트랩 표본추출 시 귀무가설에 따라 각 반응범주의 확률분포를 따르는 붓스트랩 표본이 추출되기 때문이다. 즉 임의순열검정에서 최종적으로 사용되는 가중치가 붓스트랩표본 추출 시 미리 반영되기 때문이다. 물론 붓스트랩 검정 시 표준화에 의한 주축통계량을 사용하는 표준화붓스트랩검정을 실시 할 수 있다. 다중로짓모형에서 신뢰구간 구축 방법은 앞의 2.2절에서 소개한 단순로짓모형의 신뢰구간 구축방법과 차이가 없다.

3.2. 다중로짓모형의 사례분석

2.3절에서 사용한 cancer remission data를 이용하여 다중로짓모형에서 Wald 방법과 붓스트랩 그리고 임의순열방법으로 유의확률과 신뢰구간을 구축하여 보자. 이항변수 Remiss를 Y 로 Blast와 Temp를 X_1, X_2 로 하는 다중로짓모형을 생각하자. 자료를 적합하면, $\beta_0, \beta_1, \beta_2$ 의 MLE는 각각 $\hat{\beta}_0 = 76.88, \hat{\beta}_1 = -2.613, \hat{\beta}_2 = -81.005$ 이며, β_2 에 대한 Wald의 양측 p -값은 0.126이다. 그리고 3000회의 반복실험에 의한 임의순열검정에 의한 p -값은 0.055, 비표준화붓스트랩 p -값은 0.087, 표준화붓스트랩 p -값은 0.118로 추정되었다. 또한 Wald의 95% 신뢰구간은 (-182.294, 22.283), 임의순열 95% 신뢰구간은 (-164.939, 1.818) 그리고 붓스트랩 95% 신뢰구간은 (-163.703, 2.099)으로 추정된다. 여기서 임의순열 검정 결과가 다소 유의하므로 다른 두 검정과 비교된다고 할 수 있다. 신뢰구간에서도 같은 결과를 확인할 수 있다. 신뢰구간은 유의확률보다 보다 많은 정보를 제공하여 주는데 자료분석

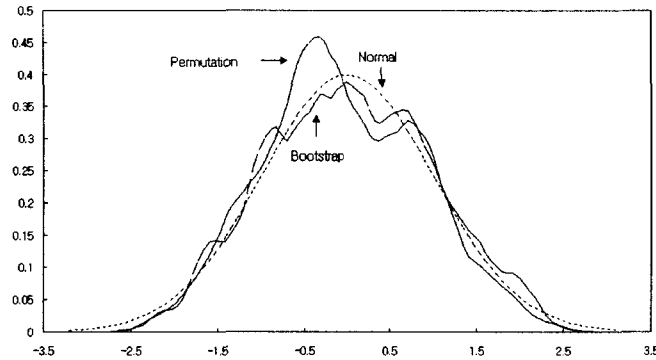


그림 3.1: 커널밀도함수 추정을 이용한 β_2 의 분포
(임의순열분포는 가중치를 고려하였음)

표 3.1: β_2 에 대한 제 1종 오류

n	α	Wald			Permutation			Bootstrap1			Bootstrap2		
		0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10	0.01	0.05	0.10
20	.000	.028	.090	.004	.024	.053	.001	.019	.057	.003	.035	.084	
30	.000	.028	.091	.008	.049	.099	.002	.027	.083	.003	.038	.099	
40	.004	.039	.090	.011	.061	.110	.007	.043	.089	.007	.045	.095	
50	.004	.052	.100	.014	.055	.111	.009	.050	.095	.008	.055	.101	

결과 붓스트랩 신뢰구간과 임의순열 신뢰구간의 상한값이 상대적으로 0에 가까움을 확인할 수 있겠다. 그림 3.1는 세 가지 방법에 대한 $\hat{\beta}_2$ 의 표준화 분포를 보여주고 있다. 여기서 임의순열의 분포는 가중치를 고려하여 제 계산한 분포이다. 우리는 여기서 임의순열의 가중분포가 정규분포나 붓스트랩 분포에 비해 약간 다른 형태를 취하고 있음을 확인할 수 있다. 특히, 0.4에서 확률밀도함수가 약간 하향하는 경향이 있으며 붓스트랩 분포는 임의순열 분포와 정규분포를 절충하는 형태임을 볼 수 있다.

3.3. 다중로짓모형의 모의실험

본 절에서는 이변량 다중로짓에 있어서 Wald 방법, 임의순열방법 그리고 붓스트랩 방법의 제 1종 오류와 신뢰구간을 살펴보고자 한다. 모의실험의 모형으로 $\beta_0 = \beta_1 = \beta_2 = 0$ 을 고려하였고 설명변수 x_1 의 수준은 단순로짓모형의 경우와 같이 1부터 10까지 정수의 값을 취하였고 설명변수 x_2 의 수준은 1,1,1,2,2,2,2,3,3,3으로 두었다. 설명변수 x_1 과 x_2 의 각 수준에서 반응범주의 확률값이 모두 0.5인 모형을 사용하였다. 또한 각 설명변수의 수준에서 각각 2회, 3회, 4회, 5회의 반복 실험을 한 그룹데이터를 모의자료로 사용하였다.

1000번 반복을 통한 모의실험의 결과가 표 3.1에 나타나 있다. 표본의 크기 40이하에서

표 3.2: 주어진 모형에서 신뢰구간들의 포함확률

n	방법	0.90	길이	SE	0.95	길이	SE	0.99	길이	SE
20	Pert	.909	6.892	.587	.955	9.052	1.400	.993	18.372	2.679
	Boot	.916	7.100	.106	.958	8.621	1.135	.992	13.385	2.176
30	Pert	.900	4.985	.011	.959	5.792	.013	.997	7.264	.015
	Boot	.898	4.990	.014	.959	5.810	.018	.995	7.420	.039
40	Pert	.906	4.251	.007	.957	4.977	.009	.993	6.374	.012
	Boot	.902	4.251	.007	.952	4.976	.008	.993	6.363	.012
50	Pert	.896	3.767	.005	.947	4.428	.006	.994	5.744	.009
	Boot	.899	3.772	.005	.945	4.434	.006	.992	5.730	.009

임의순열검정을 제외한 나머지 방법들은 주어진 유의수준보다 1종 오류가 낮게 형성되는 보수적 성향이 있음을 발견할 수 있다. 또한, 단순로짓의 경우와 같이 비표준화 붓스트랩검정은 표준화 붓스트랩검정에 비해 보수적 성향이 강함을 재확인 할 수 있다. 전반적인 모의 실험 결과 제 1종오류를 조절하는 능력은 Wald 검정, 비표준화 붓스트랩검정, 표준화 붓스트랩 검정, 임의순열 검정의 순으로 좋아지고 있음을 보여준다. 그런데, 단순로짓의 경우와 마찬가지로 임의순열검정과 표준화 붓스트랩 검정과는 큰 차이가 없음을 확인할 수 있다.

주어진 명목수준에 대해 비모수적 방법에 대한 신뢰구간의 포함확률은 표 3.2에 제시하였다. Wald의 신뢰구간은 주어진 명목수준을 모두다 넘어서는 경향이 나타난다. 이는 Wald의 신뢰구간은 과대적합하며, 상대적으로 보수적인 방법임을 의미한다고 하겠다. 그런데 비모수적 방법인 붓스트랩방법과 임의순열방법의 포함확률간에는 거의 차이가 없음을 표 3.2는 보여준다. 특히 이들 방법은 표본의 크기가 크지 않더라도 주어진 명목수준에 상당히 근사하고 있음을 볼 수 있다.

4. 결론

본 연구에서는 로짓모형에 대한 기존의 정규근사 이론과 허명희(1998)의 임의순열에 기초한 방법과 김대학, 정형철(2002)의 붓스트랩 방법의 비교를 살펴보았다. 허명희(1998)는 회귀분석에서 선형회귀에서의 임의순열검정을 로짓모형에 적용 가능하도록 확장하였다. 본 연구에서 붓스트랩 신뢰구간과 비교를 위해서 임의순열분포를 이용하여 신뢰구간을 구축하였는데 이에 대해서는 다소 논의가 있을 수 있다.

모의실험을 통하여 우리는 임의순열검정과 표준화붓스트랩검정이 다소 개방적인 성향이 있으나, 검정력이 뛰어남을 볼 수 있었다. 검정에 있어서는 비표준화붓스트랩 검정보다는 표준화붓스트랩검정을 사용하는 것이 우수함을 볼 수 있었다. 일반적으로 임의순열검정이 개념적으로 붓스트랩 방법보다 단순하다고 할 수 있다(허명희, 1998). 그런데 임의순열검정을 다중로짓모형에서 실시하려면 우도함수의 가중치를 고려한 후 재계산해야 하는 번거로움이 있으나, 붓스트랩 방법은 표본 추출과정에서 미리 가중치가 부여되기에 임의순열 검정보다 계산에 있어 다소 간편하다고 하겠다. 신뢰구간의 포함확률 추정 결과는 검정의 결과와 유사하다. 신뢰구간에서도 Wald의 신뢰구간이 다소 보수적인 성향을 지니고 있

으며, 비모수적 방법이 주어진 명목수준에 근사하며, 비모수적 방법들간에는 뚜렷한 차이점을 발견할 수 없었다. 회귀분석에서 이와 비슷한 결과는 Huh and Jhun(2001)을 참고할 수 있다. 결론적으로 붓스트랩 방법과 임의순열방법은 로짓모형에서 일반적으로 사용되는 정규근사의 추론에 대한 비모수적 대안이라고 하겠다.

참고문헌

- [1] 김대학, 정형철 (2002). Bootstrapping logit model, <한국통계학회논문집>, 제9권 제1호, 281-289.
- [2] 허명희 (1998) 로지스틱 회귀에서의 임의순열검증, <응용통계>, 제 13권, 1-11.
- [3] Agresti A. (1995) *An Introduction to Categorical Data analysis*, New York : John Wiley & Sons.
- [4] Fisher, R.A. (1935) *The Design of Experiments*, Edinburgh : Oliver & Boyd.
- [5] Garthwaite P.H. (1996) Confidence intervals from randomization tests, *Biometrics*, Vol. 52, 1387-1393.
- [6] Good, P. (1994) *Permutation Tests : A Practical Guide to Resampling Methods for Testing Hypotheses*, Springer-Verlag, New York.
- [7] Kennedy, P.E. (1995) Randomization Tests in Econometrics, *Journal of Business and Economic Statistics*, Vol. 13, 85-94.
- [8] Lee, E.T. (1974) A computer program for linear logistics regression analysis, *Computer Programs in Biomedicine*, 80-92.
- [9] Manly, B.F.J. (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology*, Chapman and Hall, London.
- [10] Huh, M. and Jhun, M. (2001) Random permutation testing in multiple linear regression, *Communication in Statistics - Theory and Methods*, 30(10), 2023-2032.

[2002년 4월 접수, 2002년 7월 채택]

Comparison of Some Nonparametric Statistical Inference for Logit Model

Hyeong Chul Jeong ¹⁾ Daehak Kim ²⁾

ABSTRACT

Nonparametric statistical inference for the parameter of logit model were examined. Usually nonparametric approach is milder than parametric approach based on normal theory assumption. We compared the two nonparametric methods for logit model, the bootstrap and random permutation in the sense of coverage probability. Monte Carlo simulation is conducted for small sample cases. Empirical power of hypothesis test and coverage probability for confidence interval estimation were presented for simple and multiple logit model respectively. An example were also introduced.

Keywords: Logit Model, Bootstrap, Random Permutation, Confidence interval, Coverage probability, Type 1 Error

1) Assistant Professor, Department of information Statistics, Pyongtaek University.

E-mail: jhc@ptuniv.ac.kr

2) Professor, Department of Statistical Information, Catholic University of Daegu.

E-mail: dhkim@cataegu.ac.kr