

모의실험에 의한 리커트형 설문분석 방법의 비교[†]

김현철¹ · 최승경² · 최동호³

¹군산대학교 통계컴퓨터과학과 · ²숙명여자대학교 통계학과 · ³성균관대학교 교육학과

접수 2016년 2월 22일, 수정 2016년 3월 22일, 게재확정 2016년 3월 25일

요 약

리커트형 데이터가 순서척도임에도 불구하고 많은 연구자들이 구간척도로 간주하여 모수 방법을 적용하고 있다. 본 연구에서는 리커트형 데이터를 어떻게 분석하는 것이 적절한지 모의 실험을 통하여 알아본다. 다양한 분포를 갖는 5점 리커트형 표본을 추출하여 위치를 비교하고, 순서척도의 경우 위치비교 보다 응답분포를 살펴보는 것이 더 타당하므로 응답분포 검정을 실시한다. 위치를 비교하는 방법으로는 구간척도로 생각하고 분석하는 모수 방법인 t -검정과 순서척도로 생각하는 비모수 방법인 만-휘트니검정 (M-W검정)을 적용하고, 응답분포를 검정하기 위해서는 카이제곱 검정과 콜모고로프 - 스미르노프검정 (K-S검정)을 실시한다. 네 가지 방법의 효율성을 비교하기 위하여 제 1종 오류 (Type I error)의 비율과 검정력 (power)을 구한다.

주요용어: 구간척도, 리커트 척도, 모수 방법, 비모수 방법, 순서척도.

1. 서론

일반적으로 데이터의 측정 척도는 스티븐스 (Stevens)의 이론에 따라 명목척도 (nominal scale) 순서척도 (ordinal scale), 구간척도 (interval scale) 그리고 비율척도 (ratio scale)로 나누어 이해한다. 이 4가지의 측정 척도에 따라 같은 목적의 통계분석을 하더라도 통계분석방법은 달라진다. 그러므로 데이터의 측정 척도에 따라 적절한 통계분석 방법을 선택하는 것은 매우 중요하다.

사회과학에서 조사연구는 종종 개인의 주관적인 감정, 태도를 측정하기 위해 실시된다. 이런 목적으로 개발된 측정방법의 하나로 널리 사용되는 것은 리커트에 의해 개발되어 리커트 척도라는 이름이 붙은 평가 척도가 있다. 5점 리커트 척도는 가장 일반적으로 사용되는 척도로서, 1=강한부정, 2=부정, 3=보통, 4=긍정, 5=강한긍정으로 순서화된 답지를 주고 응답자가 느끼는 감정이나 태도를 표시하게 한다. 리커트 평가 척도는 이런 유형의 설문을 여러 개 만들어서 이를 종합하여 응답자의 감정이나 태도를 측정하는 방법이다. Lee (2015)은 전통떡류에 대한 인지도, 활용도, 편의도 및 선호도를 측정하는데 리커트 척도를 사용하고, Lee와 Kwon (2015)도 행동통제력 지각, 예견된 후회, 음주운전태도, 주관적 규범, 흡연습관 및 운동습관에 대한 측정을 위해 리커트 척도를 사용하였다. 리커트 평가 척도의 응답지는 1에서 5까지 감정이나 태도의 정도를 점점 강하게 혹은 약하게 나타내도록 설계되므로 그 응답은 순서척도의 자료가 된다.

[†] 이 논문은 2014학년도 군산대학교 대학자체 학술공모과제 연구비 지원에 의하여 연구되었음.

¹ (54150) 전라북도 군산시 대학로 558, 군산대학교 통계컴퓨터과학과, 교수.

² 교신저자: (04310) 서울특별시 용산구 청파로 47길 100, 숙명여자대학교 통계학과, 박사.

E-mail: skchoi@sookmyung.ac.kr

³ (03063) 서울특별시 종로구 성균관로 25-2, 성균관대학교 교육학과, 박사과정.

p 점 리커트 척도의 경우, 비록 개별 문항의 응답은 순서척도로 1~ p 까지 p 개만 발생 가능 하지만, 여러 문항의 합계를 구하는 과정에서 문항의 수 p 에 비례해서 발생 가능한 응답의 수가 급속히 증가한다. 따라서 이를 구간형 척도로 보고 모수 방법을 사용하여 분석하는 것이 일반적이다. 여기에 대해서는 논란이 크지 않는 것으로 보인다. 그러나 p 점 리커트 척도의 단일 문항 (이를 Clason과 Dormody 1994는 ‘리커트형’으로 명명했다) 자료는 분명하게 순서척도이므로 리커트 평가척도의 분석방법을 그대로 적용하는 것은 이론적으로 문제가 있다.

리커트형 데이터를 분석하는 방법에는 크게 두 가지 흐름이 있다. 하나는 비록 순서척도이지만 구간척도로 간주하여 마치 리커트 평가척도의 데이터처럼 모수 방법으로 분석하는 것이고, 다른 하나는 순서척도로 인식하고 비모수 방법을 적용하는 것이다. 이론적으로는 앞서 언급한 바와 같이 리커트형 데이터는 순서척도이기 때문에, 모수 방법 보다는 비모수 방법을 사용하는 것이 적절하다고 할 수 있다. 그러나 많은 연구자들이 리커트형 자료를 그냥 리커트 평가척도의 한 유형인 것처럼 분석하여 리커트형 데이터 척도에 대한 오해를 불러올 수 있고 동시에 어떤 방법이 더 적절한지에 대한 논란이 계속되어 왔다. 대표적인 선행 연구들은 다음과 같다.

Mogey (1999)는 순서형 데이터를 기술하기 위해 중앙값과 최빈값을 사용하고, 그룹간의 차이를 비교하는데 비모수 방법을 사용할 것을 제안했다. 그러나 반대로 Traylor (1983)는 많은 경우에 순서형 데이터를 구간형 데이터로 처리해도 정확성에서 큰 손실이 없고 등간격으로 처리함으로써 해석관점에서 장점이 있음을 주장했다. Goldstein과 Hersen (1984)도 5점 리커트 척도의 각 간격이 등간격임을 가정하고 있다고 인정했다.

Clason과 Dormody (1994)는 Journal of Agricultural Education 27권에서 32권까지에 실린 논문들 중 리커트형 데이터를 사용한 95개 논문들을 보면서 리커트형 데이터를 분석하는 방법들에 대해 소개하였다. 이 논문들에서 사용된 분석방법은 54%는 기술통계 (평균, 표준편차, 범주별 비율), 13%는 두 집단을 비교하는 비모수 검정 (카이제곱 검정, 만- 휘트니 검정 (M-W 검정), 크루스칼- 윌리스 검정 (K-W검정)), 그리고 34%는 모수 검정 (t -검정, 분산분석)이었다.

Kromrey와 Hogarty (1998)는 모의실험을 통해, 순서를 갖는 범주형 자료의 집단 차이를 검정하는 네 가지 통계분석 방법을 비교하였다. 분석방법은 평균 차이를 비교하는 모수 방법인 t -검정, 비모수 방법으로 동질성을 검정하는 카이제곱 검정과 Agresti (1996)의 누적 로짓모델, 그리고 Cliff (1993)의 델타 통계량 (delta statistic)이다. 이들은 다양한 표본 크기와 분포 모양에 대해 모의실험을 통해 몬테칼로 방법으로 제 1종 오류가 잘 통제되는지와 검정력을 비교하였다. 이들의 모의실험 결과, t -검정과 누적 로짓검정은 제 1종 오류를 잘 통제하는 것으로 나타났으나, 검정력이 가장 좋은 것은 아니었다. 그리고 표본이 작거나 왜도가 심한 분포인 경우에는 제 1종 오류도 잘 통제되지 않았다. 반면 동질성에 대한 카이제곱 검정은 검정력이 가장 좋으면서 제 1종 오류의 통제능력도 가지고 있는 가장 좋은 방법이었다. De Winter와 Dodou (2010)도 다양한 분포모양에 대해 t -검정과 M-W 검정을 비교하였는데 두 방법에 거의 차이가 없음을 보고하였다.

Warachan (2011)은 다양한 시나리오에 따라 리커트형 데이터를 생성하여 t -검정, M-W검정 그리고 콜모고로프 - 스미르노프검정 (K-S검정)에 대한 로버스트성과 검정력을 비교하였다. 분석결과 t -검정과 M-W검정은 제 1종 오류를 잘 통제하면서 검정력도 좋았다. 그러나 K-S검정은 모든 경우에서 로버스트성이나 검정력이 좋지 않았다.

본 연구에서는 모의실험을 통해 리커트형 데이터를 어떤 척도로 사용하는 것이 좋은지 알아 보고자 한다. 먼저 다양한 모집단 분포로부터 5점 리커트형 표본자료를 추출하여, 위치비교를 위해 구간척도로 간주하는 모수 방법인 t -검정과 순서척도로 생각하는 비모수 방법인 M-W검정의 효율성을 비교해본다. 두 번째는 순서형 자료를 분석할 때는 위치비교 보다는 응답분포를 살펴보는 것이 더 타당하므로 응답 분포에 대한 검정으로 카이제곱 검정과 K-S검정을 실시한다. 네가지 방법 중 가장 좋은 방법을 결정하

기 위해 제 1종 오류의 크기가 잘 유지되는지에 대한 것과 검정력 (power)을 비교한다. 이런 모의실험은 통계소프트웨어인 R로 구현하였다.

본 연구가 기존의 연구들과 다른 점은 검정력 비교를 위해 서로 다른 분포로부터 얻은 표본을 비교하였다는 점이다. Warachan과 Kromrey 등은 검정력 비교를 위해 하나의 표본을 얻은 후, 비교 표본은 각 응답에 어떤 특정한 값을 더해주는 방법을 사용하므로써 5점 척도의 범주를 벗어나는 값을 사용하는 오류를 범하고 있다. 즉 각 범주에 속하는 모집단의 비율이 20:20:20:20:20인 경우 비교 표본은 각 응답에 효과로 0.1, 0.3, 0.5를 각각 더해주고 비교함으로써 현실적으로 발생할 수 없는 응답을 만들어 비교하고 있다. 본 연구에서는 검정력 비교를 위해 서로 다른 분포로부터 얻은 표본을 비교하였다.

본 논문의 구성은 다음과 같다. 2절에서 연구방법을 기술하고, 3절에서는 연구결과를 제시하며, 마지막으로 4절에서는 결론을 제시하고 있다.

2. 연구방법

모의실험에서 5점 리커트형 질문의 응답 분포로 균등분포 (uniform type), 종모양 (bell type), 한쪽으로 치우친 정도에 따라 구분한 편향분포 두 종류 (skewed type1, 2), 그리고 양봉분포 (bimodal type) 이렇게 다섯가지 응답분포를 고려하였다. 균등분포의 경우에는 1부터 5까지 각각의 응답률을 0.2로 균일하게 구성하고, 나머지 분포는 각 응답을 선택하는 횟수의 분포가 이항분포를 따른다고 가정하고 각 분포모양에 적합하도록 성공확률 p 를 결정하여 분포를 구성하였다. 종모양의 경우 $p = 0.5$, 첫 번째 편향분포 $p = 0.75$, 두 번째 편향분포 $p = 0.85$ 그리고 양봉분포의 경우에는 $p = 0.25$ 인 경우와 $p = 0.75$ 인 경우를 고려하여 각 응답이 나올 확률을 더한 후 2로 나누었다. 예를 들어, 1의 응답이 나올 비율은 $[dbinom(0,4,0.25) + dbinom(0,4,0.75)]/2$ 이다. 여기서 $dbinom(0,4,0.25)$ 는 시행의 횟수가 4이고 성공확률이 0.25인 이항분포에서 성공의 횟수가 0일 확률을 의미한다. 각 모집단에서 5점 척도의 응답 분포 구성비는 Table 2.1과 같고, Figure 2.1은 각 모집단 대한 히스토그램이다. 오른쪽으로 긴꼬리를 갖는 $p = 0.25$ 또는 $p = 0.15$ 인 편향분포는 본 논문에서 고려한 편향분포와 분석결과가 같다. 그래서 본 논문에서는 왼쪽으로 긴꼬리를 갖는 편향분포만 사용한다.

Table 2.1 Five population types the 5-point Likert scale

5-point scale	uniform type	bell type B(4,0.5)	skewed type1 B(4,0.75)	skewed type2 B(4,0.85)	bimodal type
1	0.2	0.0625	0.0039	0.0005	0.1601
2	0.2	0.25	0.0469	0.0115	0.2344
3	0.2	0.375	0.2109	0.0975	0.2109
4	0.2	0.25	0.4219	0.3685	0.2344
5	0.2	0.0625	0.3164	0.522	0.1601

t -검정과 M-W검정, K-S검정과 카이제곱검정의 효율성을 비교하기 위해 구현된 모집단으로부터 두 표본을 추출한다. 첫 번째 표본 추출방법은 동일 모집단으로부터 다양한 표본크기에 따라 표본 세트 (s_1, s_2)를 1,000번 반복하여 추출한다. 이 때 고려되는 표본크기들은 $n = m = 30, 50, 100$ 으로 동일한 경우와 $(n=30, m=50)$, $(n=30, m=100)$, $(n=50, m=100)$ 인 총 6가지 표본크기이다. 동일한 분포에서 얻은 이 표본세트는 유의수준 (제1종 오류)이 잘 통제되는지를 평가하기 위한 것이다. 즉 1,000번의 반복에서 같은 모집단에서 얻은 두 표본이 같지 않다고 의사결정을 하는 경우의 비율이 유의수준과 얼마나 일치하는가를 평가한다.

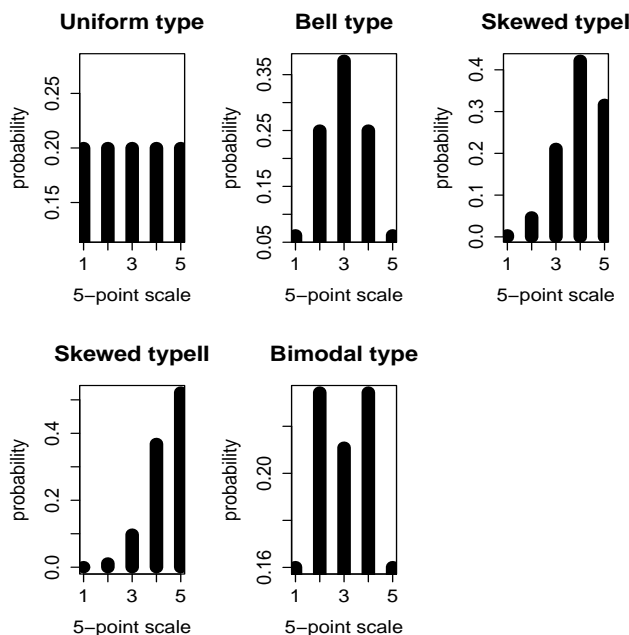


Figure 2.1 Population distributions

두 번째 표본 추출방법은 서로 다른 두 모집단으로부터 각각 표본을 추출하여 표본세트를 이루는 경우이다. 이 때 사용되는 표본크기 및 반복 횟수는 첫 번째 방법과 동일하다. 이 표본세트는 서로 다른 모집단에서 얻은 표본으로 실제로 다르다고 의사결정을 하는 비율, 즉 검정력을 비교하기 위한 것이다.

첫 번째 방법으로 추출된 표본세트를 가지고 “ H_0 : 두 모집단의 위치모수는 같다.”라는 가설검정을 위해 모수 방법인 t -검정과 비모수 방법인 M-W검정을 실시한다. 그리고 “ H_0 : 두 모집단의 분포는 같다.”라는 가설검정을 위해서는 K-S검정과 카이제곱 검정을 한다. 이 때, H_0 를 기각한다는 검정결과는 제 1종 오류를 범하는 것으로 1,000개의 유의확률 (p -값) 중 0.05보다 작은 값의 비율을 통해 제 1종 오류의 비율을 구한다. 이 비율은 검정방법들이 유의수준 5%를 얼마나 잘 유지하고 있는지 나타내는 값이다. 두 번째 방법으로 추출된 표본세트를 가지고 두 가설을 검정한 결과 귀무가설을 기각한다는 것은 옳은 결정이므로, 1,000개의 p -값 중 0.05보다 작은 값의 비율을 통해 검정방법들의 검정력을 구할 수 있다.

3. 분석결과

Table 3.1는 다섯 가지 모집단으로부터 동일한 분포를 갖는 두 표본을 추출하여 t -검정과 M-W검정, K-S검정과 카이제곱검정을 실시 후, 제 1종 오류의 비율을 표본 크기에 따라 정리한 표이다. 제 1종 오류의 비율을 비교해 보면 전반적으로 t -검정, M-W검정 그리고 카이제곱검정은 제 1종 오류의 비율이 0.05근처 값을 갖고 있다. 그러므로 세 검정방법은 유의수준 5%를 잘 유지하고 있는 검정방법이라고 할 수 있다. K-S검정은 제 1종 오류의 비율이 0.05보다 아주 작은 값들을 가지고 있으므로 유의수준 5%보다 낮은 유의수준을 가지고 있다고 할 수 있다. t -검정은 표본크기가 작거나, 표본크기 차이가 큰 경우 M-W검정에 비해 유의수준 5%에서 더 어긋나는 값을 갖는 경향을 보였다. 즉, 이런 경우에는 모수 방법인 t -검정보다는 비모수 방법인 M-W검정이 더 효율적이라고 할 수 있다.

Table 3.1 Type I error rate of the five population types

	test method	sample size (n, m)					
		30,30	30,50	30,100	50,50	50,100	100,100
uniform	t -test	0.060	0.053	0.059	0.055	0.038	0.060
	M-W	0.059	0.051	0.053	0.054	0.039	0.061
	K-S	0.016	0.009	0.008	0.009	0.003	0.015
	chisq	0.049	0.055	0.049	0.057	0.035	0.047
bell	t -test	0.050	0.044	0.050	0.050	0.052	0.048
	M-W	0.048	0.046	0.051	0.051	0.049	0.042
	K-S	0.007	0.002	0.009	0.013	0.003	0.003
	chisq	0.060	0.041	0.057	0.049	0.040	0.041
skewed1	t -test	0.062	0.051	0.053	0.048	0.043	0.053
	M-W	0.062	0.047	0.045	0.050	0.044	0.047
	K-S	0.007	0.002	0.004	0.003	0.005	0.003
	chisq	0.053	0.055	0.051	0.044	0.041	0.062
skewed2	t -test	0.048	0.045	0.059	0.056	0.045	0.052
	M-W	0.052	0.050	0.046	0.046	0.044	0.060
	K-S	0.010	0.006	0.005	0.007	0.005	0.002
	chisq	0.060	0.041	0.046	0.052	0.034	0.057
bimodal	t -test	0.049	0.051	0.052	0.054	0.048	0.038
	M-W	0.042	0.050	0.046	0.051	0.048	0.039
	K-S	0.008	0.010	0.004	0.005	0.009	0.008
	chisq	0.038	0.050	0.045	0.042	0.051	0.050

Table 3.2와 Table 3.3은 각각 다른 분포를 갖는 표본을 추출하여 네 가지 검정방법을 실시 후, 검정력을 각각의 분포와 표본크기에 따라 정리한 표들이다.

Table 3.2 Statistical power of the uniform type versus four population types

sample size	test method	unifrom			
		bell	skewed1	skewed2	bimodal
30, 30	t -test	0.061	1.000	1.000	0.054
	M-W	0.065	0.999	1.000	0.049
	K-S	0.296	0.999	1.000	0.014
	chisq	0.984	1.000	1.000	0.126
30, 50	t -test	0.060	1.000	1.000	0.042
	M-W	0.070	1.000	1.000	0.041
	K-S	0.263	0.999	1.000	0.011
	chisq	0.971	1.000	1.000	0.114
30, 100	t -test	0.053	1.000	1.000	0.048
	M-W	0.057	0.999	1.000	0.047
	K-S	0.278	1.000	1.000	0.011
	chisq	0.976	1.000	1.000	0.121
50, 50	t -test	0.046	1.000	1.000	0.050
	M-W	0.046	1.000	1.000	0.049
	K-S	0.265	1.000	1.000	0.010
	chisq	0.975	1.000	1.000	0.128
50, 100	t -test	0.042	1.000	1.000	0.043
	M-W	0.050	0.999	1.000	0.045
	K-S	0.270	1.000	1.000	0.010
	chisq	0.978	1.000	1.000	0.114
100, 100	t -test	0.047	1.000	1.000	0.039
	M-W	0.05	0.999	1.000	0.044
	K-S	0.268	0.999	1.000	0.013
	chisq	0.976	1.000	1.000	0.122

Table 3.3 Statistical power of each pair in bell type, skewed1 type, skewed2 type and bimodal type

sample size	test method	bell		skewed1		skewed2	
		skewed1	skewed2	bimodal	skewed2	bimodal	bimodal
30, 30	<i>t</i> -test	1.000	1.000	0.052	0.932	1.000	1.000
	M-W	1.000	1.000	0.056	0.921	1.000	1.000
	K-S	1.000	1.000	0.119	0.630	0.998	1.000
	chisq	1.000	1.000	0.852	0.853	1.000	1.000
30, 50	<i>t</i> -test	1.000	1.000	0.046	0.955	1.000	1.000
	M-W	1.000	1.000	0.051	0.940	1.000	1.000
	K-S	0.999	1.000	0.110	0.629	0.999	1.000
	chisq	1.000	1.000	0.867	0.870	1.000	1.000
30, 100	<i>t</i> -test	1.000	1.000	0.050	0.938	1.000	1.000
	M-W	1.000	1.000	0.051	0.929	0.999	1.000
	K-S	1.000	1.000	0.102	0.623	0.999	1.000
	chisq	1.000	1.000	0.859	0.858	1.000	1.000
50, 50	<i>t</i> -test	1.000	1.000	0.049	0.959	1.000	1.000
	M-W	1.000	1.000	0.052	0.953	1.000	1.000
	K-S	1.000	1.000	0.119	0.654	0.999	1.000
	chisq	1.000	1.000	0.855	0.880	1.000	1.000
50, 100	<i>t</i> -test	1.000	1.000	0.048	0.948	1.000	1.000
	M-W	1.000	1.000	0.056	0.939	1.000	1.000
	K-S	1.000	1.000	0.102	0.642	1.000	1.000
	chisq	1.000	1.000	0.852	0.863	1.000	1.000
100, 100	<i>t</i> -test	1.000	1.000	0.040	0.937	1.000	1.000
	M-W	1.000	1.000	0.045	0.928	1.000	1.000
	K-S	1.000	1.000	0.099	0.631	1.000	1.000
	chisq	1.000	1.000	0.860	0.852	1.000	1.000

Table 3.2의 결과를 보면 균일분포와 모양이 많이 다른 편향분포 1과 편향분포 2에서는 네 가지 검정 방법 모두 표본크기에 상관없이 검정력이 좋았다. 그러나 균일분포와 같은 위치모수를 갖는 양봉모형의 경우에는 전반적으로 검정력이 좋지 않았다. 그리고 종모양인 경우에는 *t*-검정과 M-W검정은 검정력이 낮고, K-S검정 보다 카이제곱검정의 검정력이 높은 것으로 나타났다. Table 3.3의 결과도 Table 3.2 과 같이 분포의 모양이 많이 다른 표본에서는 네 가지 검정방법이 좋은 검정력을 갖으나 위치모수가 같은 대칭형 모양의 분포인 경우에는 *t*-검정과 M-W 검정 그리고 K-S검정의 검정력은 현저히 나빠지는 것을 확인할 수 있었다. 그러나 카이제곱 검정은 모든 경우에 검정력이 좋았다. 네 가지 검정법을 살펴본 결과 카이제곱 검정이 가장 유의수준을 잘 유지하고, 검정력 또한 좋았다.

데이터 척도가 순서형인 경우에는 위치비교보다는 응답별 분포에 대한 검정을 하는 것이 타당한데, 연구 결과를 보면 리커트형 데이터에 대해 카이제곱 검정이 좋은 것으로 확인 되므로 리커트형 데이터는 순서형척도로 간주하고 분포를 비교하는 카이제곱 검정을 사용하는 것이 바람직하다.

4. 결론

5점 척도 리커트형 데이터를 순서형이 아닌 구간형으로 분석을 할 때 발생하는 문제점을 알아보기 위해 위치검정을 위한 모수 방법 *t*-검정과, 비모수 방법 M-W검정을 고려하고, 분포의 동질성에 대한 검정을 위해 K-S검정과 카이제곱검정을 실시하였다.

먼저 같은 분포를 갖는 두 표본을 가지고 네 검정방법을 실시하여 검정방법들이 유의수준을 잘 유지하고 있는지를 제1종 오류의 비율을 가지고 확인 결과 *t*-검정, M-W검정 그리고 카이제곱검정은 유의수준 5%를 잘 유지하고 있다. 그러나 (1) 표본의 크기가 작거나, (2) 분포의 왜도가 심하거나 (3) 표본의 크기가 많이 차이가 나는 경우에는 *t*-검정보다는 M-W검정이 더 효과적이라고 확인 되었다. 이 결과는 기

존의 연구결과 같았다. 즉, 대칭분포이면서 표본의 크기가 큰 경우에는 모수 방법이나 비모수 방법은 차이가 없으나, 그 외의 경우에는 결과에 차이가 있으므로 비모수 방법을 사용하는 것이 정확한 결과를 얻을 수 있다고 할 수 있다.

두 번째는 다른 분포의 모양을 가지는 두 표본에 대한 네 검정방법의 검정력을 확인해 본 결과 분포의 모양이 차이가 많이 날수록 네 검정법의 검정력은 매우 좋았으나, 비슷한 모양을 갖는 경우에는 t -검정과 M-W검정의 검정력이 좋지 않았고, M-W검정의 검정력이 t -검정보다 좀 더 낮다는 것을 확인 할 수 있었다. 그리고 동일분포인지에 대한 검정에서는 카이제곱검정이 K-S검정 보다 좋은 검정력을 가지고 있다고 할 수 있다.

마지막으로 유의수준이 통제되는 수준이나 검정력을 보면 카이제곱 검정법이 가장 좋은 결과를 보이고 있다. 이 결과는 순서척도의 경우 위치비교 문제보다는 응답별 분포를 비교하는 것이 더 타당하다는 주장을 지지한다. 그러므로 5점 척도 리커트형 데이터를 분석할 때는 위치비교보다는 응답별 분포에 대한 분석이 더 타당하고, 만약 위치비교를 꼭 해야 한다면 순서형을 구간형으로 간주하고 분석하는 것은 신중해야 한다.

References

- Agresti, A. (1996). *An introduction to categorical data analysis*, Wiley, New York.
- Clason, D. L. and Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, **35**, 31-35.
- Cliff, N. (1993). Dominance statistics: Ordinal analyses to answer ordinal questions. *Psychological Bulletin*, **114**, 494-509.
- De Winter, J. C. F. and Dodou, D. (2010). Five-point Likert items: T test versus Mann-Whitney-Wilcoxon. *Practical Assessment, Research & Evaluation*, **11**, 1-16
- Goldstein, G. and Hersen, M. (1984). *Handbook of psychological assessment*, Pergamon Press, New York.
- Kromrey, J. D. and Hogarty, K. Y. (1998). Analysis options for group difference on ordered categorical variables: An empirical investigation of type I error control and statistical power. *Multiple Linear Regression Viewpoints*, **25**, 70-82.
- Lee, K. H. and Kwon, Y. M. (2015). Influence of sociopsychological aspects, smoking habit, exercise habit on the intentions of drink-driving. *Journal of the Korean Data & Information Science Society*, **26**, 21-29.
- Lee, M. O. (2015). A study on differences by gender in influence of recognition, usage, convenience and preference on purchase intention for traditional korean rice cake. *Journal of the Korean Data & Information Science Society*, **26**, 875-883.
- Mogey, N. (1999). *So you want to use a Likert scale?* http://www.icbl.hw.ac.uk/ltldi/cookbook/info_likert_scale.
- Traylor, M. (1983). Ordinal and interval scaling. *Journal of Market Research Society*, **25**, 297-303.
- Warachan, B. (2011). Appropriate statistical analysis for two independent groups of likert-type data. Ph. D. Thesis, American University, Washington D.C..

A simulation comparison on the analysing methods of Likert type data[†]

Hyun Chul Kim¹ · Seung Kyoung Choi² · Dong Ho Choi³

¹Department of Statistics and Computer Science, Kunsan National University

²Department of Statistics, Sookmyung Women's University

³Department of Education, Sungkyunkwan University

Received 22 February 2016, revised 22 March 2016, accepted 25 March 2016

Abstract

Even though Likert type data is ordinal scale, many researchers who regard Likert type data as interval scale adapt as parametric methods. In this research, simulations have been used to find out a proper analysis of Likert type data. The locations and response distributions of five point Likert type data samples having diverse distribution have been evaluated. In estimating samples' locations, we considered parametric method and non-parametric method, which are t-test and Mann-Whitney test respectively. In addition, to test response distribution, we employed Chi-squared test and Kolmogorov-Smirnov test. In this study, we assessed the performance of the four aforementioned methods by comparing Type I error ratio and statistical power.

Keywords: Interval scale, Likert-type scale, nonparametric method, ordinal scale, parametric method.

[†] This paper was supported by research funds of Kunsan National University 2014.

¹ Professor, Department of Statistics and Computer Science, Kunsan National University, Kunsan 54150, Korea.

² Corresponding author: Doctor, Department of Statistics, Sookmyung Women's University, Seoul 04310, Korea. E-mail: skchoi@sookmyung.ac.kr

³ Doctoral course student, Department of Education, Sungkyunkwan University, Seoul 03063, Korea.