응용통계연구 제 19권 1호, 2006년, pp. 111-120

이단계 이수준 균형지분모형의 순위변환 기법연구*

최영훈[1)]

요 약

이단계 이수준 균형지분모형에서 주효과 및 지분효과를 검정하기 위한 모수적 검정과 순위변환을 이용한 검정은 전반적으로 제1종 오류율이 상당히 유사하며, 주효과 및 지분효과를 검정하기 위한 순위변환 통계량의 검정력은 모수적 통계량의 검정력보다 상대적으로 뛰어난 수준임을 보여준다. 한편 효과의 크기와 표본의 크기를 증가시킬수록 모수적 통계량과 순위변환 통계량의 검정력 증가량의 크기는 현저하게 향상되며, 특히 지수분포와 같은 비대칭분포하에서 모든 인자가 고정일때 순위변환 통계량의 검정력이 모수적 통계량의 검정력보다 월등히 높은 수준임을 나타낸다.

주요용어: 지분모형, 순위변환, 제1종 오류, 검정력, ANOVA F 통계량.

1. 서론

Conover and Iman(1981)이 순위변환을 일련의 관측된 원자료를 가장 작은 값부터 가장 큰 값까지 순위를 부여한 후 모수적 검정을 적용하는 과정으로 정의한 이래 Hora and Conover(1984), Hora and Iman(1988), Kepner and Robinson(1988)는 랜덤화 블록모형 및 이원배치법하에서의 주효과 검정을, Fabian(1991), Thompson(1991), Akritas and Arnold(1994) Gorman and Akritas(2001)는 이원배치법하에서의 상호작용 개념 등의 복잡성과 문제점을 제시하고 있으며, Choi(1998)는 삼원배치법하에서의 검정력을 분석연구하였다.

따라서 본 연구대상의 주된 관심사는 이제까지의 기존의 제한된 요인계획법 모형에서 벗어나서 현실적으로 자주 이용되는 지분모형의 순위변환 검정을 살펴보고자 한다. 구체적으로 정규, 지수, 이중지수, 균일분포 등의 다양한 모집단 분포하에서의 자료를 토대로 이단계 이수준 균형지분모형의 주인자 A와 이에 지분되어 있는 인자 B(A)의 효과유무의 검정을 위하여 모수적 ANOVA F 통계량과 순위 변환된 ANOVA FR 통계량 사이의 제1종 오류와 검정력을 시뮬레이션을 이용하여 비교분석해 보고자 한다. 특히 이단계 이수준 균형지분모형의 의미있는 모든 가능한 경우의 검정을 검토한 바, 첫째로 주인자와 지분인자가 모두 고정인 경우의 주효과의 제1종 오류율과 검정력, 둘째로 주인자와 지분인자가 모두 고정인 경우의 지분효과의 제1종 오류율과 검정력, 셋째로 주인자는 고정이나 지분인자는 랜덤인 경우의 주효과의 제1종 오류율과 검정력을 면밀히 분석하고자 한다.

* 본 연구는 2005년 한신대학교 학술연구비 지원에 의한 것임.

1) (447-791) 경기도 오산시 양산동 411, 한신대학교 정보통계학과, 교수
E-mail: choicyh@hanshin.ac.kr

연구대상인 이단계 이수준 균형지분모형은

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \epsilon_{k(ij)} \qquad i,j = 1,2,\ k = 1,2,\ldots,n$$

으로 표현할 수 있다. 이때 μ는 총평균을 나타내며 연구내용상 0 으로 간주하도록 한다. α_i는 주인자 A의 i 번째 수준의 주효과를 나타내며, $\beta_{j(i)}$ 는 주인자 A의 i번째 수준에 지분되어 있는 지분인자 B(A)의 j 번째 수준의 지분효과를 나타낸다. 오차항 $\epsilon_{k(ij)}$는 표준정규모집단 N(0,1), 지수모집단 EXP(1), 이중지수모집단 DOUBLE(0,1) 및 균일모집단 U(0,1) 으로부터 추출된 독립인 관측치를 나타낸다. 한편 주효과 및 지분효과 검정을 위한 귀무가설 $H_0 : \alpha_i = 0$ 및 $H_0 : \beta_{j(i)} = 0$, 단 $i,j = 1,2,$에 대하여 F 와 FR 은 각각 모수적 ANOVA 검정통계량과 원자료의 순위에 바탕을 둔 F 검정통계량으로 아래와 같이 정의한다.

가설	A 및 B(A):고정	A 및 B(A):고정	A:고정, B(A):랜덤
통계량	$H_0 : \alpha_i = 0$	$H_0 : \beta_{j(i)} = 0$	$H_0 : \alpha_i = 0$
F	$\frac{8n(n-1)\sum_{i=1}^{2}(\bar{y}_{i..}-\bar{y}_{...})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2}$	$\frac{2n(n-1)\sum_{i=1}^{2}\sum_{j=1}^{2}(\bar{y}_{ij.}-\bar{y}_{i..})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2}$	$\frac{4\sum_{i=1}^{2}(\bar{y}_{i..}-\bar{y}_{...})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}(\bar{y}_{ij.}-\bar{y}_{i..})^2}$
FR	$\frac{8n(n-1)\sum_{i=1}^{2}(\bar{R}_{i..}-\bar{R}_{...})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n}(R_{ijk}-\bar{R}_{ij.})^2}$	$\frac{2n(n-1)\sum_{i=1}^{2}\sum_{j=1}^{2}(\bar{R}_{ij.}-\bar{R}_{i..})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n}(R_{ijk}-\bar{R}_{ij.})^2}$	$\frac{4\sum_{i=1}^{2}(\bar{R}_{i..}-\bar{R}_{...})^2}{\sum_{i=1}^{2}\sum_{j=1}^{2}(\bar{R}_{ij.}-\bar{R}_{i..})^2}$

이때 $\bar{y}_{ij.} = \sum_{k=1}^{n} y_{ijk}/n$, $\bar{y}_{i..} = \sum_{j=1}^{2}\sum_{k=1}^{n} y_{ijk}/2n$, $\bar{y}_{...} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} y_{ijk}/4n$ 을 의미하며 순위 $R_{ijk} = R(y_{ijk})$, $\bar{R}_{ij.} = \sum_{k=1}^{n} R_{ijk}/n$, $\bar{R}_{i..} = \sum_{j=1}^{2}\sum_{k=1}^{n} R_{ijk}/2n$, $\bar{R}_{...} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n} R_{ijk}/4n$을 나타낸다.

시뮬레이션 과정은 정규, 지수, 이중지수, 균일분포를 따르는 변량을 생성시킨후, 주어진 효과를 발생시키기 위하여 각각의 효과크기에 상응하는 상수를 가산하며, 주효과 및 지분효과 검정을 위한 모수적 ANOVA F 통계량 및 순위로 변환된 ANOVA FR 통계량을 계산한다. 그리고 5% 유의수준하에서의 기각되는 비율을 각각 계산한다. 효과의 크기 c 는 0.25 부터 1.5 까지 0.25 씩 증가하였으며, 표본의 크기 $n = 4, 10, 20, 30, 50$ 을 고려하였다. 한편 이와같은 일련의 과정을 10000번이상 반복하였으며, 변량를 발생시키기 위하여 C^{++} 언어의 rand() 함수 등을 사용하였으며, 결과물의 비교를 위하여 EXCEL 을 이용하였다.

2. 제1종 오류의 분석

본 절에서는 주인자 A 및 지분인자 B(A)의 효과를 검정하기 위한 두 통계량 F 와 FR의 제1종 오류율을 다음과 같은 세가지 상황하에서 분석하고자 한다.

(2.1) A 와 B(A) 인자가 모두 고정인 경우의 주효과 α_i 의 제1종 오류율 : $\beta_{j(i)} = c$

(2.2) A 와 B(A) 인자가 모두 고정인 경우의 지분효과 $\beta_{j(i)}$ 의 제1종 오류율 : $\alpha_i = c$

(2.3) A 인자는 고정이나, B(A) 인자는 랜덤인 경우로서 오차항의 모집단 분포형태를 유지할 목적으로 오차항과 동일한 분포를 가정한 경우의 주효과 α_i 의 제1종 오류율.

아래의 표 2.1은 두인자가 모두 고정인 (2.1)의 경우로서 정규분포, 지수분포, 이중지수분포 및 균일분포 모집단으로부터 표본크기 $n = 4, 10, 30$일 때 유의수준 0.05 하에서의 주

효과의 제1종 오류율을 요약한 도표이다. 그리고 표 2.2는 두 인자가 모두 고정인 (2.2)의 경우로서 지분효과의 제1종 오류율을, 표 2.3은 혼합효과인 (2.3)의 경우로서 주효과의 제1종 오류율을 나타낸다.

표 2.1: A 와 B(A) 인자가 모두 고정인 경우의 주효과 α_i 의 제1종 오류, 단 $\beta_{j(i)} = c$

		정규			지수			이중 지수			균일		
c	통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
0.25	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.050	0.052	0.051	0.054	0.054	0.051	0.053	0.054	0.053	0.053	0.053	0.051
0.50	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.049	0.053	0.051	0.053	0.054	0.052	0.050	0.054	0.053	0.057	0.054	0.050
0.75	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.048	0.052	0.051	0.050	0.053	0.052	0.050	0.053	0.053	0.057	0.054	0.050
1.00	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.048	0.052	0.051	0.050	0.053	0.051	0.049	0.053	0.053	0.057	0.054	0.050
1.25	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.048	0.052	0.051	0.052	0.053	0.051	0.049	0.053	0.053	0.057	0.054	0.050
1.50	F	0.049	0.052	0.052	0.047	0.050	0.051	0.047	0.053	0.052	0.056	0.054	0.051
	FR	0.050	0.052	0.051	0.053	0.053	0.051	0.050	0.052	0.052	0.057	0.054	0.050

표 2.2: A 와 B(A) 인자가 모두 고정인 경우의 주효과 $\beta_{j(i)}$ 의 제1종 오류, 단 $\alpha_i = c$

		정규			지수			이중 지수			균일		
c	통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
0.25	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.052	0.052	0.050	0.053	0.054	0.053	0.052	0.051	0.051	0.048	0.055	0.052
0.50	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.050	0.050	0.050	0.050	0.053	0.054	0.048	0.052	0.052	0.049	0.053	0.051
0.75	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.047	0.050	0.051	0.045	0.052	0.053	0.043	0.053	0.052	0.049	0.053	0.051
1.00	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.045	0.048	0.050	0.044	0.050	0.050	0.045	0.051	0.052	0.049	0.053	0.051
1.25	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.046	0.049	0.049	0.045	0.047	0.050	0.045	0.051	0.051	0.049	0.053	0.051
1.50	F	0.049	0.050	0.051	0.044	0.046	0.049	0.044	0.049	0.049	0.053	0.055	0.052
	FR	0.047	0.049	0.050	0.046	0.048	0.049	0.045	0.051	0.051	0.049	0.053	0.051

종합적으로 표 2.1, 표 2.2 및 표 2.3 에서 알 수 있는 바와 같이 제1종 오류율은 모수적 F 통계량과 순위 변환 FR통계량 모두 0.05에서 크게 벗어나지 않으며, 전반적으로 효과의 크

표 2.3: A 인자는 고정이나 B(A) 인자는 랜덤인 경우의 주효과 α_i 의 제1종 오류

	정규			지수			이중 지수			균일		
통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
F	0.042	0.041	0.041	0.049	0.051	0.054	0.049	0.052	0.052	0.053	0.053	0.054
FR	0.041	0.043	0.042	0.054	0.052	0.054	0.052	0.054	0.052	0.055	0.052	0.054

기나 표본의 크기 혹은 통계량의 유형과 상관없이 명목상의 유의수준과 비슷함을 알 수 있다. 결론적으로 모집단 분포를 달리하거나, 효과의 크기, 표본의 크기를 달리하는 거의 모든 상황에 걸쳐 다소간의 차이는 있지만 전반적으로 F검정 및 FR검정의 제1종 오류율이 상당히 유사함을 발견할 수 있다.

3. 검정력 분석

본 절에서는 주인자 A 및 지분인자 B(A)의 효과를 검정하기 위한 두 통계량 F 와 FR의 검정력을 다음과 같은 세가지 상황하에서 분석하고자 한다.

(3.1) A 와 B(A) 인자가 모두 고정인 경우의 주효과 α_i 의 검정력 :
- (1) B(A)의 효과를 A 효과의 0.5배로 가정한 경우 : $\alpha_i = c$, $\beta_{j(i)} = 0.5c$
- (2) B(A)의 효과를 A 효과의 1.5배로 가정한 경우 : $\alpha_i = c$, $\beta_{j(i)} = 1.5c$

(3.2) A 와 B(A) 인자가 모두 고정인 경우의 지분효과 $\beta_{j(i)}$의 검정력 :
- (1) A의 효과를 B(A) 효과의 0.5배로 가정한 경우 : $\alpha_i = 0.5c$, $\beta_{j(i)} = c$
- (2) A의 효과를 B(A) 효과의 1.5배로 가정한 경우 : $\alpha_i = 1.5c$, $\beta_{j(i)} = c$

(3.3) A 인자는 고정이나, B(A) 인자는 랜덤인 경우로서 오차항과 동일한 분포를 가정한 경우의 주효과 α_i의 검정력 : $\alpha_i = c$.

3.1. 주인자 A 효과의 검정력

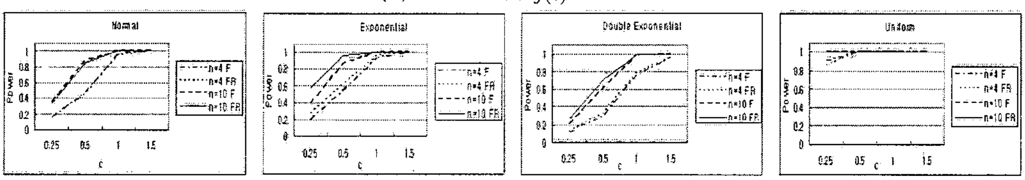
첫번째로 A 와 B(A) 인자가 모두 고정인 경우의 주효과 α_i의 검정력을 분석하고자 한다. 아래의 표 3.1 및 그림 3.1은 두인자가 모두 고정인 (3.1)의 경우로서 [(1): B(A)의 효과를 A 효과의 0.5배로 가정한 경우인 $\alpha_i = c$, $\beta_{j(i)} = 0.5c$ 및 (2): B(A)의 효과를 A 효과의 1.5배로 가정한 경우인 $\alpha_i = c$, $\beta_{j(i)} = 1.5c$] 정규분포, 지수분포, 이중지수분포 및 균일분포 모집단으로부터 표본크기 $n = 4, 10, 30$ 일때 유의수준 0.05 하에서의 주효과의 검정력을 종합적으로 요약한 도표 및 그림이다.

표 3.1 및 그림 3.1에서 알 수 있는 바와 같이 전반적으로 FR 통계량의 검정력은 F 통계량의 검정력보다 뛰어난 수준임을 알 수 있다. 구체적으로 모집단이 정규성을 가정할 때 우수한 검정력으로 알려진 F통계량과 FR통계량의 검정력은 거의 일치하고 있다. 그러나 대표적 비대칭분포인 지수분포 및 양쪽 끝이 길게 늘어진 이중지수분포하에서는 FR 통계량의 검정력이 F 통계량의 검정력보다 월등히 높음에 주목할 필요가 있다. 한편 효과의 크

표 3.1: A와 B(A) 인자가 모두 고정인 경우의 주효과 α_i의 검정력, 단 (1): $\alpha_i = c$, $\beta_{j(i)} = 0.5c$, (2): $\alpha_i = c$, $\beta_{j(i)} = 1.5c$

		정규			지수			이중 지수			균일		
c	통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
0.25	F	0.161	0.345	0.781	0.193	0.382	0.782	0.112	0.211	0.501	0.900	0.999	1.000
	FR(1)	0.159	0.333	0.759	0.269	0.589	0.965	0.133	0.265	0.641	0.853	0.999	1.000
	FR(2)	0.154	0.331	0.757	0.233	0.507	0.932	0.126	0.250	0.616	0.825	0.999	1.000
0.50	F	0.464	0.875	0.999	0.537	0.865	0.998	0.299	0.610	0.968	1.000	1.000	1.000
	FR(1)	0.447	0.859	0.999	0.627	0.961	1.000	0.337	0.711	0.993	1.000	1.000	1.000
	FR(2)	0.432	0.853	0.999	0.550	0.914	1.000	0.314	0.669	0.987	1.000	1.000	1.000
0.75	F	0.790	0.997	1.000	0.807	0.988	1.000	0.548	0.902	1.000	1.000	1.000	1.000
	FR(1)	0.769	0.996	1.000	0.856	0.998	1.000	0.583	0.948	1.000	1.000	1.000	1.000
	FR(2)	0.745	0.995	1.000	0.793	0.993	1.000	0.545	0.921	1.000	0.000	1.000	1.000
1.00	F	0.963	1.000	1.000	0.933	0.999	1.000	0.759	0.986	1.000	1.000	1.000	1.000
	FR(1)	0.948	1.000	1.000	0.954	1.000	1.000	0.784	0.995	1.000	1.000	1.000	1.000
	FR(2)	0.938	1.000	1.000	0.917	1.000	1.000	0.748	0.990	1.000	1.000	1.000	1.000
1.25	F	0.997	1.000	1.000	0.978	1.000	1.000	0.894	0.998	1.000	1.000	1.000	1.000
	FR(1)	0.994	1.000	1.000	0.987	1.000	1.000	0.904	1.000	1.000	1.000	1.000	1.000
	FR(2)	0.993	1.000	1.000	0.969	1.000	1.000	0.877	0.979	1.000	1.000	1.000	1.000
1.50	F	0.999	1.000	1.000	0.993	1.000	1.000	0.958	1.000	1.000	1.000	1.000	1.000
	FR(1)	0.999	1.000	1.000	0.996	1.000	1.000	0.963	1.000	1.000	1.000	1.000	1.000
	FR(2)	0.999	1.000	1.000	0.989	1.000	1.000	0.946	1.000	1.000	1.000	1.000	1.000

(1): $\alpha_i = c$, $\beta_{j(i)} = 0.5c$



(2) : $\alpha_i = c$, $\beta_{j(i)} = 1.5c$

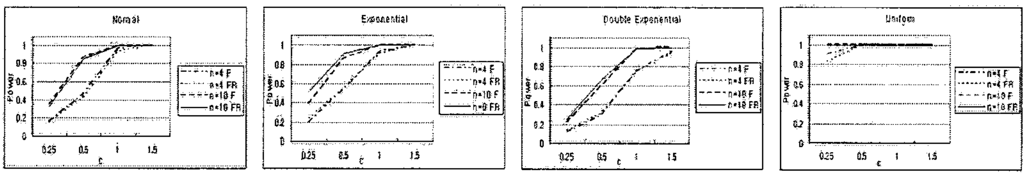


그림 3.1: A와 B(A) 인자가 모두 고정인 경우의 주효과 α_i의 검정력, 단 (1): $\alpha_i = c$, $\beta_{j(i)} = 0.5c$, (2): $\alpha_i = c$, $\beta_{j(i)} = 1.5c$

기 c 와 표본의 크기 n 을 증가시킬수록 F 통계량과 FR 통계량의 검정력 증가량은 현저하게 커짐을 쉽게 발견할 수 있을 것이다. 특히 지수분포하에서는 효과의 크기가 작고 표본의 크기가 커질수록 FR 통계량의 검정력은 F 통계량의 검정력보다 큰 격차의 우위를 보이고 있다. 균일분포하에서는 대부분의 경우에 효과의 크기, 표본의 크기 및 F 와 FR 통계량에 무관하게 검정력이 1 로 나타나고 있다.

그리고 B(A)의 지분효과를 A의 주효과의 0.5배로 작게 가정한 경우가 1.5배로 크게 가정한 경우보다 FR 통계량의 검정력이 보다 증가하는 경향을 보이며 상대적으로 FR 통계량의 검정력은 F 통계량의 검정력보다 우수함을 발견할 수 있다.

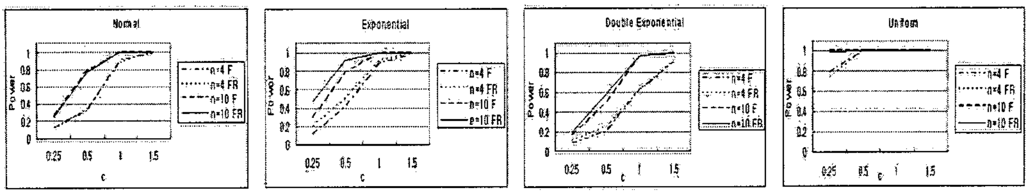
3.2. 지분인자 B(A) 효과의 검정력

두번째로 A와 B(A) 인자가 모두 고정인 경우의 지분효과 $\beta_{j(i)}$의 검정력을 분석하고자 한다. 아래의 표 3.2 및 그림 3.2는 두인자가 모두 고정인 (3.2)의 경우로서 [(1): A의 효과를 B(A) 효과의 0.5배로 가정한 경우인 $\alpha_i = 0.5c$, $\beta_{j(i)} = c$ 및 (2): A의 효과를 B(A) 효과의 1.5배로 가정한 경우인 $\alpha_i = 1.5c$, $\beta_{j(i)} = c$] 정규분포, 지수분포, 이중지수분포 및 균일분포 모집단으로부터 표본크기 $n = 4, 10, 30$일 때 유의수준 0.05하에서의 지분효과의 검정력을 종합적으로 요약한 도표 및 그림이다.

표 3.2: A와 B(A) 인자가 모두 고정인 경우의 지분효과 $\beta_{j(i)}$의 검정력, 단 (1) : $\alpha_i = 0.5c$, $\beta_{j(i)} = c$, (2) : $\alpha_i = 1.5c$, $\beta_{j(i)} = c$

		정규			지수			이중 지수			균일		
c	통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
0.25	F	0.113	0.256	0.676	0.125	0.291	0.686	0.078	0.160	0.401	0.780	0.999	1.000
	FR(1)	0.116	0.250	0.656	0.203	0.472	0.935	0.104	0.198	0.535	0.741	0.997	1.000
	FR(2)	0.110	0.247	0.650	0.167	0.395	0.997	0.096	0.189	0.507	0.679	0.997	1.000
0.50	F	0.323	0.783	0.999	0.406	0.788	0.997	0.206	0.495	0.936	1.000	1.000	1.000
	FR(1)	0.326	0.762	0.999	0.507	0.920	1.000	0.250	0.595	0.981	1.000	1.000	1.000
	FR(2)	0.293	0.751	0.999	0.413	0.853	1.000	0.219	0.548	0.967	1.000	1.000	1.000
0.75	F	0.652	0.987	1.000	0.701	0.975	1.000	0.410	0.824	0.999	1.000	1.000	1.000
	FR(1)	0.641	0.984	1.000	0.765	0.994	1.000	0.461	0.892	1.000	1.000	1.000	1.000
	FR(2)	0.586	0.981	1.000	0.676	0.984	1.000	0.411	0.853	1.000	0.000	1.000	1.000
1.00	F	0.891	1.000	1.000	0.873	0.998	1.000	0.630	0.967	1.000	1.000	1.000	1.000
	FR(1)	0.873	1.000	1.000	0.904	1.000	1.000	0.668	0.985	1.000	1.000	1.000	1.000
	FR(2)	0.846	1.000	1.000	0.841	0.999	1.000	0.609	0.972	1.000	1.000	1.000	1.000
1.25	F	0.980	1.000	1.000	0.957	1.000	1.000	0.804	0.997	1.000	1.000	1.000	1.000
	FR(1)	0.976	1.000	1.000	0.968	1.000	1.000	0.826	0.999	1.000	1.000	1.000	1.000
	FR(2)	0.966	1.000	1.000	0.929	1.000	1.000	0.776	0.997	1.000	1.000	1.000	1.000
1.50	F	0.998	1.000	1.000	0.984	1.000	1.000	0.912	1.000	1.000	1.000	1.000	1.000
	FR(1)	0.996	1.000	1.000	0.989	1.000	1.000	0.917	1.000	1.000	1.000	1.000	1.000
	FR(2)	0.995	1.000	1.000	0.967	1.000	1.000	0.884	1.000	1.000	1.000	1.000	1.000

(1): $\alpha_i = 0.5c$, $\beta_{j(i)} = c$



(2) : $\alpha_i = 1.5c$, $\beta_{j(i)} = c$

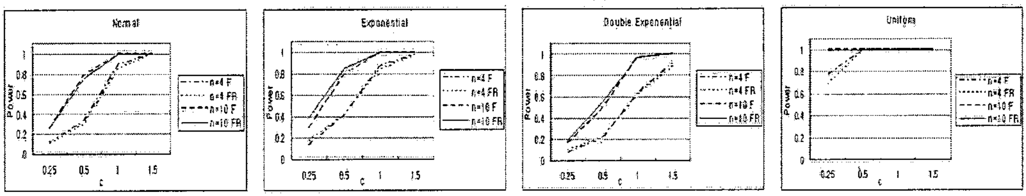


그림 3.2: A와 B(A) 인자가 모두 고정인 경우의 주효과 α_i 의 검정력, 단 (1): $\alpha_i = 0.5c$, $\beta_{j(i)} = c$, (2): $\alpha_i = 1.5c$, $\beta_{j(i)} = c$

표 3.2 및 그림 3.2에서 알 수 있는 바와 같이 전반적으로 FR 통계량의 검정력은 F 통계량의 검정력보다 뛰어난 수준임을 알 수 있다. 앞서 살펴본 주효과의 검정력에서의 분석결과가 동일하게 지분효과의 검정력에도 적용됨을 알 수 있다. 그리고 A의 주효과를 B(A)의 지분효과의 0.5배로 작게 가정한 경우가 1.5배로 크게 가정한 경우보다 FR 통계량의 검정력이 보다 증가하는 경향을 보이며 상대적으로 FR 통계량의 검정력은 F 통계량의 검정력보다 우수함을 발견할 수 있다.

한편 상대적으로 앞서 제시한 표 3.1 및 그림 3.1이 나타내는 주인자 A 효과의 F 와 FR 검정력보다 표 3.2 및 그림 3.2이 나타내는 지분인자 B(A) 효과의 F와 FR 검정력이 다소 떨어지는 것을 알 수 있으며, 이는 B(A) 인자가 A 인자에 지분되어 있는 지분모형의 제약적 특성에 기인함을 유추할 수 있다.

3.3. 혼합효과 존재시 주인자 A 효과의 검정력

세번째로 A 인자는 고정이나 B(A) 인자는 랜덤인 경우로서 오차항과 동일한 분포를 가정한 경우의 주효과 α_i의 검정력을 분석하고자 한다. 아래의 표 3.3 및 그림 3.3은 혼합효과가 존재하는 $\alpha_i = c$인 (3.3)의 경우로서 정규분포, 지수분포, 이중지수분포 및 균일분포 모집단으로부터 표본크기 $n = 4, 10, 30$일 때 유의수준 0.05하에서의 주효과의 검정력을 종합적으로 요약한 도표 및 그림이다.

표 3.3 및 그림 3.3도 전반적으로 FR 통계량의 검정력이 F 통계량의 검정력보다 뛰어난 수준임을 보여준다. 특히 고정효과와 랜덤효과가 동시에 존재하는 혼합효과의 경우에는 두 인자가 모두 고정효과인 경우보다는 F 검정 및 FR 검정의 검정력이 낮게 나타나는 결과를 보였다. 이와같은 현상은 랜덤효과가 오차항의 분산을 증가시키는 결과를 야기시키며, 이로 말미암아 나머지 고정효과를 검정하는 검정력을 다소 떨어뜨림에 기인한다 볼 수 있다.

그러나 혼합효과가 존재하는 지분모형의 경우에도 모집단 분포, 효과의 크기, 표본의 크기 등에 상관없이 순위변환 FR 통계량의 검정력이 모수적 F 통계량의 검정력보다 상대적으로 우월함을 보이며, 고정효과만이 존재하는 경우와 마찬가지로 동일한 결과해석이 유지됨을 알 수 있다.

표 3.3: A 인자는 고정이나 B(A) 인자는 랜덤인 경우의 주효과 α_i의 검정력, 단 $\alpha_i = c$

		정규			지수			이중 지수			균일		
c	통계량	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30	n=4	n=10	n=30
0.25	F	0.063	0.089	0.177	0.110	0.211	0.500	0.077	0.128	0.287	0.968	1.000	1.000
	FR	0.061	0.087	0.174	0.128	0.264	0.643	0.084	0.139	0.320	0.949	1.000	1.000
0.50	F	0.121	0.214	0.476	0.295	0.610	0.968	0.167	0.357	0.786	1.000	1.000	1.000
	FR	0.119	0.221	0.476	0.342	0.718	0.994	0.177	0.389	0.837	1.000	1.000	1.000
0.75	F	0.204	0.397	0.753	0.544	0.899	1.000	0.309	0.651	0.983	1.000	1.000	1.000
	FR	0.208	0.400	0.761	0.595	0.950	1.000	0.324	0.686	0.991	1.000	1.000	1.000
1.00	F	0.312	0.578	0.917	0.760	0.986	1.000	0.484	0.871	1.000	1.000	1.000	1.000
	FR	0.322	0.593	0.925	0.795	0.996	1.000	0.499	0.900	1.000	1.000	1.000	1.000
1.25	F	0.431	0.731	0.980	0.896	0.999	1.000	0.657	0.968	1.000	1.000	1.000	1.000
	FR	0.447	0.758	0.985	0.908	1.000	1.000	0.670	0.979	1.000	1.000	1.000	1.000
1.50	F	0.541	0.847	0.996	0.959	1.000	1.000	0.796	0.995	1.000	1.000	1.000	1.000
	FR	0.563	0.873	0.998	0.963	1.000	1.000	0.801	0.997	1.000	1.000	1.000	1.000

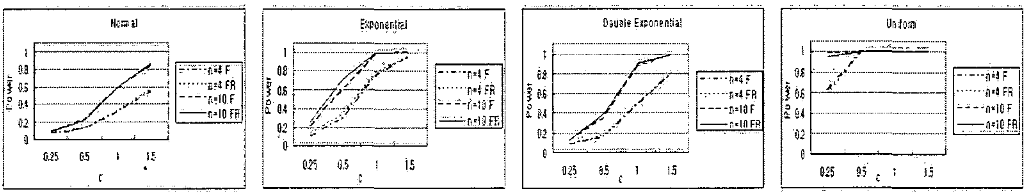


그림 3.3: A 인자는 고정이나 B(A) 인자는 랜덤인 경우의 주효과 α_i의 검정력, 단 $\alpha_i = c$

4. 결론

이단계 이수준 균형지분모형에서 주효과 및 지분효과를 검정하기 위한 F 검정과 순위변환을 이용한 FR 검정은 모집단 분포, 효과의 크기, 표본의 크기와 상관없이 거의 모든 상황에 걸쳐 전반적으로 F 검정과 FR 검정의 제1종 오류율이 상당히 유사하며, 그리고 일반적으로 주효과 및 지분효과를 검정하기 위한 FR 통계량의 검정력은 F 통계량의 검정력보다 뛰어난 수준임을 나타낸다.

특히 지수분포와 같은 비대칭분포하에서 모든인자가 고정일 때 FR 통계량의 검정력이 F 통계량의 검정력보다 월등히 높음에 주목할 필요가 있고, 효과의 크기와 표본의 크기를

증가시킬수록 F 통계량과 FR 통계량의 검정력 증가량의 크기는 현저하게 향상됨을 보여준다.

참고문헌

Akritas, M. G. and Arnold, S. F. (1994). Fully nonparametric hypotheses for factorial designs, *Journal of the American Statistical Association*, **89**, 336-343.

Choi, Y. H. (1998). A study of the power of the rank transform test in a 2(3) factorial experiment, *Communications in Statistics*, **27**, 251-266.

Conover, W. J. and Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics, *The American Statistician*, **35**, 124-128.

Fabian, V. (1991). On the problem of interactions in the analysis of variance, *Journal of the American Statistical Association*, **86**, 362-374.

Gorman, J. O. and Akritas, M. G. (2001). Nonparametric models and methods for designs with dependent censored data, *Biometrics*, **57**, 88-95.

Hora, S. C. and Conover, W. J. (1984). The F statistic in the two-way layout with rank score transformed data, *Journal of the American Statistical Association*, **79**, 668-673.

Hora, S. C. and Iman, R. L. (1988). Asymptotic relative efficiencies of the rank transformation procedure in randomized complete block design, *Journal of the American Statistical Association*, **83**, 462-470.

Kepner, J. L. and Robinson, D. H. (1988). Nonparametric methods for detecting treatment effects in repeated-measures designs, *Journal of the American Statistical Association*, **83**, 456-461.

Thompson, G, L. (1991), A note on the rank transform for interactions, *Biometrika*, **78**, 697-701.

[2005년 1월 접수, 2005년 9월 채택]

Rank Transformation Technique in a Two-stage Two-level Balanced Nested Design*

Young Hun Choi[1)]

Abstract

In a two-stage two-level balanced nested design, type I error rates for the parametric tests and the rank transformed tests for the main effects and the nested effects are in overall similar to each other. Furthermore, powers for the rank transformed statistic for the main effects and the nested effects in a two-stage two-level balanced nested design are generally superior to powers for the parametric statistic. When the effect size and the sample size are increased, we can find that powers increase for the parametric statistic and the rank transformed statistic are dramatically improved. Especially for the case of the fixed effects in the asymmetric distributions such as an exponential distribution, powers for the rank transformed tests are quite high rather than powers for the parametric tests.

Keywords: Nested design, Rank transformation, Type I error, Power, ANOVA F statistic

* This research was supported by Hanshin University Research Grant 2005.

1) Professor, Department of Information and Statistics, Hanshin University, 411 Yangsandong, Osan, Kyunggido, 447-791, Korea.
E-mail: choicyh@hanshin.ac.kr