The Korean Communications
in Statistics Vol. 5, No. 1, 1998
pp. 85-97

완전확률화모형 및 랜덤화블럭모형하에서 순위변환을 이용한 다중비교의 시뮬레이션 분석 1)

최 영 훈 2)

요 약

완전확률화모형 및 랜덤화블럭모형하에서의 주요한 다중비교 분석기법들을 시뮬레이션을 이용하여 검토하고자 하였다. 시뮬레이션 결과는 순위변환과 최소유의차검정을 이용한 다중비교 분석기법이 모수적 ANOVA F 검정과 Fisher의 유의차검정, 비모수적 Kruskal-Wallis 검정과 최소유의차검정 및 Friedman 검정과 최소유의차검정을 이용한 분석기법보다 전체실험오차율, 전체실험검정력 및 개별쌍검정력 면에서 상대적으로 뛰어남을 보여준다. 즉 순위변환한 ANOVA F 검정의 전체실험오차율은 명목상의 유의수준을 잘 유지하고 있으며, 전체실험검정력 및 개별쌍검정력은 모수적 ANOVA F 검정과 Kruskal-Wallis 검정 및 Friedman 검정기법보다 전반적으로 우수함을 알 수 있다.

1. 서 론

일반적으로 두 모집단 이상간의 처리효과 차이유무를 분석하기 위해서는 전체모집단 처리효과간에 차이가 존재하는지를 알기 위한 목적뿐만 아니라 특정모집단 처리효과간에 차이가 존재하는지를 알기 위한 관심사항으로 구별할 수 있다. 즉 실험계획모형에 대한 분산분석을 수행하고자 할때 처리평균간에 차이가 존재한다면 어느 모집단의 처리평균간에 차이가 존재하는지를 알기 위한 비교과정을 다중비교(multiple comparisons)라 한다.

이를 위해 David, Lachenbruch와 Brandis(1972)는 정규분포하에서의 표준화 범위를 이용한 다중비교의 검정력 분석을, Carmer와 Swanson(1973)은 몬테칼로 방법에 의한 열가지 다중비교 분석기법의 차이점 분석을, Einot와 Gabriel(1975) 및 Ramsey (1978)는 모수적 다중비교 과정에 따른 검정력의 의태연구를 구체적으로 수행하였으며, Lehmann과 Shaffer(1977)는 다중비교의 기본적인 이론적 토대를 이룩하였다.

한편 Lin과 Haseman(1978)은 Mann-Whitney 및 Kruskal-Wallis의 비모수기법을 이용한 다중비교 연구를, Conover와 Iman(1979)은 순위변환(rank transformation)을 이용한 일원배치법하에서의 다중비교 기법을 처음으로 소개하였다. 추가로 David (1987) 및 Andrews와 David(1990)는 불균형 순위자료에 대한 비모수적 다중비교 과정을 연구하였다. 이외에도 Hochberg와 Varon(1984)

---

1) 본 논문은 1997년도 한신대학교 특별연구비의 지원에 의하여 연구되었음.
2) (447-791) 경기도 오산시 양산동 411, 한신대학교 통계학과 부교수.

및 Uusipaikka(1985)는 각각 공분산분석하에서의 다중비교 및 다중비교의 개념을 신뢰구간 등의 문제로 확대적용하였으며, 최근에 Hayter(1989) 및 Kunert(1990)는 기존의 연구된 다중비교에 관한 시뮬레이션 결과의 이론적 근거를 제시하는 등으로 다양화하였다.

그러나 모수적 및 비모수적 다중비교 방법들간의 체계적인 비교연구는 지금까지 본격적으로 수행되지 못하였으며, 특히 순위변환을 이용한 다중비교 기법과 모수적 혹은 다른 비모수적 다중비교 기법과의 차이점에 관한 연구는 아직 구체적으로 이루어지지 않았으므로 좀더 세심한 관찰 및 분석을 필요로 한다. 다시 말하여 최근까지 순위변환의 적용은 실험계획모형의 분석에 주로 치중되어 왔으며, Choi(1998)에 의한 2^3 요인계획법하에서의 검정력 연구로까지 발전하여 왔다.

따라서 본 연구의 주된 관심사는 이제까지 개별적으로 연구, 발전하여온 다중비교 기법을 순위변환을 이용한 다중비교 기법을 중심으로 연구하고자 하되 대표적인 모수적 및 비모수적 다중비교 분석기법과의 차이점 비교에 촛점을 맞추고자 한다. 이를 위한 비교대상의 모형으로는 임의로 추출한 실험단위에 처리를 배치하는 간단한 완전확률화모형(completely randomized design) 및 한 블럭에 모든 처리가 포함되고 블럭내 처리의 배치를 완전랜덤으로 하는 랜덤화블럭모형(randomized complete block design: RCBD)을 이용하고자 하며, 몬테칼로 시뮬레이션을 통한 제1종오류와 검정력면에서 비교분석하고자 한다.

앞으로 2절에서는 완전확률화모형 및 랜덤화블럭모형하에서의 다중비교 분석을 위한 시뮬레이션 연구방법을 자세히 설명하고자 하며, 3절에서는 잘 알려지고 이해하기 쉬운 실험계획모형의 분산분석 기법을 적용한 모수적 F 검정, 순위변환을 이용한 F 검정 및 Kruskal-Wallis 검정, Friedman 검정기법에 의한 전체실험오차율을 각각 살펴보고자 하며, 이어서 4절에서는 전체모집단간의 모평균 차이가 존재하는지를 알기 위한 기법으로 특히 유용한 앞서 언급한 네가지 검정기법에 의한 전체실험검정력의 분석을 꾀하고자 하며, 5절에서는 전체모집단간의 모평균 차이가 존재할때 Fisher의 유의차검정 및 최소유의차 검정을 이용하여 각각의 쌍에 대한 개별쌍검정력 분석을 시도하고자 한다.

2. 연구방법

완전확률화모형하에서 확률변수 X_{ij}, $i=1,\ldots,k$, $j=1,\ldots,n_i$, 단 잠정적으로 $n_i=n$, 는 k 모집단으로부터 추출한 i 번째 처리수준의 j 번째 관측값을, 랜덤화블럭모형하에서 확률변수 X_{ij}, $i=1,\ldots,k$, $j=1,\ldots,b$, 는 k 모집단으로부터 b 블럭에 걸쳐 추출한 i 번째 처리수준 및 j 번째 블럭수준을 반영한 독립인 표본이라 정의하자. 그렇다면 관심대상의 모형은 아래와 같이 각각 표현할 수 있다.

$$X_{ij} = \mu + \tau_i + \varepsilon_{ij},$$

$$\text{혹은} \quad X_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, \tag{2.1}$$

단 μ 는 총평균, τ_i 는 i 번째 수준의 처리효과, β_j 는 j 번째 수준의 블럭효과, 오차항 ε_{ij} 는

각각 표준정규분포의 모집단 $N(0,1)$, 로그정규분포의 모집단 $LN(0,1)$, 지수분포의 모집단 $EXP(1)$ 및 균일분포의 모집단 $U(0,1)$ 으로부터 추출된 독립인 관측치를 나타낸다. 아울러 처리와 블럭은 고정효과를 갖는다고 가정하고, (2.1)의 블럭효과 β_j 는 가운데 블록효과의 크기를 -0.05 및 0.05 으로하되 모든 b 블록간에 걸쳐 처음블럭보다 다음블럭에 있어서 0.1 씩 증가하는 효과차이가 존재한다고 가정하자 (예를들어 $b=10$ 인 경우에는 블럭효과 $\beta_1=-0.45$, $\beta_2=-0.35$, $\cdots$, $\beta_5=-0.05$, $\beta_6=0.05$, $\cdots$, $\beta_9=0.35$, $\beta_{10}=0.45$ 를 가정한다).

이때 전체실험의 다중비교를 위한 주된 가설은

$$H_0 : \ \tau_1 = \tau_2 = \cdots = \tau_k \ , \tag{2.2}$$

$$H_1 : \ \tau_i \neq \tau_j \ , \text{ 적어도 하나의 쌍 } (i, j) \text{ 에 대하여} \tag{2.3}$$

와 같다. 한편 R_{ij}, $i=1,\ldots,k$, $j=1,\ldots,n$, (혹은 b), 는 X_{ij} 에 해당하는 순위라 정의하고, 순위할당은 앞으로 살펴볼 순위변환 검정기법에 대하여는 관측치의 원자료를 오름차순으로 1 부터 nk 혹은 bk 까지 할당하며, Friedman 검정기법의 순위는 모든 주어진 블럭 j 에 대하여 오름차순으로 1 부터 k 까지 할당한다.

시뮬레이션 연구에 있어서 전체실험오차율(experimentwise error rate)은 전체모집단간에 모평균 차이가 존재하지 않을때 적어도 한쌍의 모집단간에 모평균 차이가 존재한다고 잘못 판단한 비율인 제1종오류를 일컫는다. 반면에 전체실험검정력(experimentwise power)은 전체모집단간에 모평균 차이가 존재할때 적어도 한쌍의 모집단간에 모평균 차이가 존재한다고 올바르게 판단한 비율인 검정력으로 정의하며, 개별쌍검정력(pairwise power)은 전체실험검정을 위한 거짓의 귀무가설 (2.2)가 기각될때 추가적으로 발생하는 개별쌍의 다중비교에 따른 검정력이라 정의할 수 있다.

본 의태연구를 위하여 처리수준 $k=3,4,5,10$ 및 n, $b=6,10,20,30,50$ 인 경우를 고려하되 전체실험오차율, 전체실험검정력 및 개별쌍검정력을 위한 확률변량 생성은 C 프로그램의 서브루틴을 기본으로 활용하였다. 그러나 본 논문에서는 실제로 자주 이용되는 완전확률화모형 및 랜덤화블럭모형의 처리수준인 $k=5$ 및 전체실험검정력의 효율성을 쉽게 논의하기 위하여 n, $b=6, 10, 20$ 인 경우(한편 개별쌍검정력의 효율성을 보다 쉽게 논의하기 위하여는 n, $b=6, 20, 30$ 인 경우)의 시뮬레이션 결과만을 중심으로 검토하고자 한다.

구체적으로 시뮬레이션의 첫번째 단계는 Box-Muller 기법을 이용하여 표준정규분포 $N(0,1)$ 의 유사확률변량을, 이어서 관계식 $exp[N(0,1)]$ 을 이용하여 로그정규분포 $LN(0,1)$ 의 유사확률변량을, 그리고 C 언어의 rand() 함수를 이용하여 균일분포 $U(0,1)$ 의 유사확률변량 U 를, 또한 $-\ln U$ 의 역변환기법을 이용하여 지수분포 $EXP(1)$ 의 유사확률변량을 생성하고자 한다. 두번째 단계는 완전확률화모형 및 랜덤화블럭모형의 ANOVA F 검정통계량, 순위변환한 ANOVA F 검정통계량 및 Kruskal-Wallis 검정통계량, Friedman 검정통계량을 각각 계산한 후 0.05의 유

의수준 및 귀무가설 (2.2)하에서의 전체실험오차율 및 전체실험검정력을 살펴보고자 한다. 세번째 단계는 만일 전체실험검정을 위한 귀무가설 (2.2)가 기각될 경우에 한하여 개별쌍검정력을 구하고자 하며, 이를 위하여 모평균이 서로 다른 모든 가능한 개별쌍을 대상으로 추가적인 이표본 검정기법을 적용하고자 한다. 이때 매경우의 실험에 있어서 각각 10000번씩의 반복을 수행하고자 한다.

3. 전체실험오차율의 결과분석

전체실험오차율의 분석을 위하여 표 3.1에 제시된 귀무가설 (2.2)하에서의 모집단 유형에 따른 순위변환한 F 검정기법을 모수적 ANOVA F 검정 및 Kruskal-Wallis 검정, Friedman 검정기법과 비교하도록 하자.

이때 표 3.1이 나타내는 바와 같이 모집단의 수 $k=5$ 인 완전확률화모형 및 랜덤화블럭모형을 전제로 한다. (사실상 표 3.1의 오차율 분석을 위한 모집단 조건은 앞으로 살펴볼 4절의 검정력 분석을 위한 모집단 조건들과 대조를 이루고 이해를 돕기 위한 목적으로 아래표와 같이 정리하였다)

표 3.1 : 전체실험오차율의 비교분석을 위한 모집단 조건

모집단	$k=5$				
	1 (첫번째)	2 (두번째)	3 (세번째)	4 (네번째)	5 (다섯번째)
정규분포	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$	$N(0,1)$
로그정규분포	$LN(0,1)$	$LN(0,1)$	$LN(0,1)$	$LN(0,1)$	$LN(0,1)$
지수분포	$EXP(1)$	$EXP(1)$	$EXP(1)$	$EXP(1)$	$EXP(1)$
균일분포	$U(0,1)$	$U(0,1)$	$U(0,1)$	$U(0,1)$	$U(0,1)$

아래의 표 3.2는 위의 표 3.1에서 제시된 상황하에서의 전체실험오차율의 결과를 나타낸다. 이때 표 3.2의 첫번째 열은 모집단 유형을, 두번째 및 네번째 열의 "통계량"은 세가지 통계적 검정기법을 나타낸다. 즉 완전확률화모형 및 랜덤화블럭모형하에서 "F_1 , F_2"는 각각의 모수적 ANOVA 검정통계량을, "FR_1 , FR_2"는 순위변환을 이용한 F 검정통계량을, "KW" 및 "FD"는 잘 알려진 대표적인 비모수기법인 Kruskal-Wallis 및 Friedman 검정통계량을 나타낸다.

구체적으로 완전확률화모형에서의 F_1 검정통계량 및 랜덤화블럭모형하에서의 F_2 검정통계량은 처리효과 검정을 위한 모수적 ANOVA F 통계량을 의미한다. 또한 완전확률화모형에서의 FR_1 검정통계량 및 랜덤화블럭모형하에서의 FR_2 검정통계량은 처리효과 검정을 위하여 순위로 변환된 ANOVA FR 통계량을 의미한다. 반면에 Kruskal-Wallis 검정통계량 KW 는 Kruskal and Wallis(1952)에 의하여 정의된 통계량을 의미하며, Friedman 검정통계량 FD 는 Friedman(1937)에

의하여 제시된 통계량으로 정의한다.

따라서 표 3.2는 F_1 , F_2 , FR_1 , FR_2 , KW 및 FD 검정통계량을 사용할때 참인 귀무가설 (2.2)를 기각하는 비율을 나타낸다.

표 3.2 : ANOVA F_1 , F_2 검정, 순위변환한 FR_1 , FR_2 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 에 따른 전체실험오차율

모집단	통계량	완전확률화모형			통계량	랜덤화블럭모형		
		$n=6$	$n=10$	$n=20$		$b=6$	$b=10$	$b=20$
정규분포	F_1	0.051	0.052	0.048	F_2	0.053	0.051	0.049
	FR_1	0.054	0.053	0.049	FR_2	0.055	0.053	0.051
	KW	0.041	0.045	0.046	FD	0.045	0.044	0.045
로그정규분포	F_1	0.034	0.037	0.036	F_2	0.033	0.035	0.037
	FR_1	0.054	0.053	0.049	FR_2	0.055	0.053	0.051
	KW	0.041	0.045	0.046	FD	0.045	0.044	0.045
지수분포	F_1	0.043	0.041	0.045	F_2	0.046	0.041	0.045
	FR_1	0.053	0.051	0.051	FR_2	0.051	0.050	0.048
	KW	0.040	0.044	0.048	FD	0.040	0.044	0.048
균일분포	F_1	0.055	0.052	0.050	F_2	0.052	0.050	0.051
	FR_1	0.053	0.051	0.051	FR_2	0.053	0.050	0.050
	KW	0.040	0.044	0.048	FD	0.040	0.044	0.048

한편 표 3.2의 결과를 살펴보면 전반적으로 전체실험오차율은 모든 ANOVA F_1 , F_2 검정통계량, 순위변환을 이용한 FR_1 , FR_2 검정통계량 및 Kruskal-Wallis 검정통계량 KW, Friedman 검정통계량 FD 에 걸쳐 블럭수가 증가할수록 매우 비슷한 경향을 보여주고 있다. 그중에서도 특히 Kruskal-Wallis 검정 및 Friedman 검정에 따른 전체실험오차율이 모집단의 유형에 상관없이 명목상의 유의수준 0.05 보다 다소 낮으며, 동시에 로그정규분포 및 지수분포와 같은 이상치가 존재할 수 있는 분포하에서의 F_1 , F_2 검정통계량의 전체실험오차율도 약간 낮은 수준임을 알 수 있다. 이에 비해 순위변환을 이용한 FR_1 , FR_2 검정통계량은 모집단의 유형에 상관없이 명목상의 유의수준과 상당히 가까운 고른 수준임을 알 수 있다.

4. 전체실험검정력의 결과분석

전체실험검정력의 분석을 위하여 표 4.1로 제시된 대립가설 (2.3)의 효과를 반영한 모집단 및 표 4.2로 제시된 또다른 대립가설 (2.3)의 효과를 반영한 모집단의 상황하에서 순위변환한 FR 검정을 모수적 ANOVA F 검정 및 Kruskal-Wallis 검정, Friedman 검정과 비교하도록 하자. 다시 말하

여 전체실험검정력의 계산을 위하여 표 4.1 및 표 4.2로 주어진 두가지의 다른 상황을 고려하였다.

우선 아래의 표 4.1은 모집단이 정규분포와 로그정규분포일 경우에는 첫번째 및 두번째 모집단과는 달리 세번째, 네번째 특히 다섯번째 모집단의 분포를 다르게, 지수분포일 경우에는 첫번째, 두번째 및 세번째 모집단과는 달리 네번째 특히 다섯번째 모집단의 분포를 다르게, 균일분포일 경우에는 다섯번째 모집단의 분포만을 유난히 다르게 가정함으로써 대립가설의 효과를 반영한 전체실험검정력 비교분석을 위한 모집단 조건이다.

표 4.1 : 전체실험검정력 비교분석을 위한 모집단 조건 I

모집단	$k=5$				
	1 (첫번째)	2 (두번째)	3 (세번째)	4 (네번째)	5 (다섯번째)
정규분포	$N(0,1)$	$N(0,1)$	$N(0.5,1)$	$N(1,1)$	$N(1.5,3^2)$
로그정규분포	$LN(0,1)$	$LN(0,1)$	$LN(0.5,1)$	$LN(1,1)$	$LN(1.2,1)$
지수분포	$EXP(1)$	$EXP(1)$	$EXP(1)$	$EXP(0.5)$	$EXP(3)$
균일분포	$U(0,1)$	$U(0,1)$	$U(0,1)$	$U(0,1)$	$U(0.5,3.5)$

이에 반해 다음의 표 4.2로 제시된 경우는 앞서 표 4.1로 제시된 경우보다 서로 다른 분포를 갖는 모집단의 수를 늘림으로써 전체적인 모집단간의 분포차이를 보다 명백히 가정한 경우이다. 즉 모집단이 정규분포와 로그정규분포일 경우에는 모든 모집단간의 분포를 다르게, 지수분포일 경우에는 첫번째 및 두번째 모집단과는 달리 세번째, 네번째 및 다섯번째 모집단의 분포를 다르게, 균일분포일 경우에는 첫번째, 두번째 및 세번째 모집단과는 달리 네번째 특히 다섯번째 모집단의 분포를 다르게 가정함으로써 또다른 대립가설의 효과를 반영한 전체실험검정력 비교분석을 위한 모집단 조건이다

따라서 전체실험검정력 비교분석을 위하여 표 4.2로 제시된 경우의 모집단 조건은 표 4.1로 제시된 경우의 모집단 조건에 비하여 전반적으로 높은 검정력을 갖음을 예견할수 있을 것이다.

표 4.2 : 전체실험검정력 비교분석을 위한 모집단 조건 II

모집단	$k=5$				
	1 (첫번째)	2 (두번째)	3 (세번째)	4 (네번째)	5 (다섯번째)
정규분포	$N(0,1)$	$N(0.5,1)$	$N(1,1)$	$N(1.5,1)$	$N(2,3^2)$
로그정규분포	$LN(0,1)$	$LN(0.5,1)$	$LN(1,1)$	$LN(1.5,1)$	$LN(2,1)$
지수분포	$EXP(1)$	$EXP(1)$	$EXP(0.5)$	$EXP(2)$	$EXP(3)$
균일분포	$U(0,1)$	$U(0,1)$	$U(0,1)$	$U(0.5,1.5)$	$U(0.5,3.5)$

아래의 표 4.3 및 표 4.4의 결과는 표 4.1 및 표 4.2에서 고려한 다양한 경우에 있어서의 결과를 요약한 것이다. 즉 표 4.3 및 표 4.4는 각각 표 4.1 및 표 4.2로 제공된 대립가설 효과를 위한 모집단 조건들하에서의 전체실험검정력을 보여주고 있다. 이때 전체실험검정력은 모든 모집단간의 차이가 존재하지 않는다는 거짓의 귀무가설 (2.2)를 기각하는 비율로서 정의한다. 물론 표 4.3 및 표 4.4의 통계량 "F_1 , F_2", "FR_1 , FR_2" 및 "KW, FD"는 각각 완전확률화모형 및 랜덤화블럭모형 하에서의 ANOVA F 검정통계량, 순위변환을 이용한 FR 검정통계량 및 Kruskal-Wallis 검정통계량, Friedman 검정통계량을 의미한다.

표 4.3 : 표 4.1 에 제시된 모집단 조건 I 하에서의 ANOVA F_1 , F_2 검정, 순위변환한 FR_1 , FR_2 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 에 따른 전체실험검정력

모집단	통계량	완전확률화모형			통계량	랜덤블럭화모형		
		$n=6$	$n=10$	$n=20$		$b=6$	$b=10$	$b=20$
정규분포	F_1	0.286	0.411	0.720	F_2	0.276	0.403	0.718
	FR_1	0.299	0.499	0.853	FR_2	0.288	0.487	0.824
	KW	0.254	0.474	0.845	FD	0.216	0.394	1.000
로그정규분포	F_1	0.263	0.478	0.844	F_2	0.250	0.464	0.839
	FR_1	0.481	0.750	0.979	FR_2	0.462	0.725	0.972
	KW	0.423	0.728	0.978	FD	0.364	0.634	0.944
지수분포	F_1	0.345	0.645	0.979	F_2	0.332	0.631	0.977
	FR_1	0.464	0.748	0.984	FR_2	0.434	0.720	0.984
	KW	0.401	0.724	0.982	FD	0.333	0.619	0.954
균일분포	F_1	0.997	1.000	1.000	F_2	0.996	1.000	1.000
	FR_1	0.958	1.000	1.000	FR_2	0.963	1.000	1.000
	KW	0.940	1.000	1.000	FD	0.868	0.992	1.000

표 4.4 : 표 4.2 에 제시된 모집단 조건 II 하에서의 ANOVA F_1 , F_2 검정, 순위변환한 FR_1 , FR_2 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 에 따른 전체실험검정력

모집단	통계량	완전확률화모형			통계량	랜덤블럭화모형		
		$n=6$	$n=10$	$n=20$		$b=6$	$b=10$	$b=20$
정규분포	F_1	0.378	0.575	0.911	F_2	0.366	0.567	0.906
	FR_1	0.467	0.738	0.982	FR_2	0.447	0.720	0.974
	KW	0.409	0.714	0.981	FD	0.347	0.612	1.000
로그정규분포	F_1	0.472	0.765	0.970	F_2	0.448	0.749	0.969
	FR_1	0.805	0.972	1.000	FR_2	0.784	0.968	1.000
	KW	0.762	0.967	1.000	FD	0.672	0.928	0.999
지수분포	F_1	0.481	0.808	0.998	F_2	0.460	0.793	0.998
	FR_1	0.540	0.815	0.992	FR_2	0.520	0.814	0.994
	KW	0.483	0.794	0.991	FD	0.410	0.695	0.969
균일분포	F_1	0.998	1.000	1.000	F_2	0.997	1.000	1.000
	FR_1	0.989	1.000	1.000	FR_2	0.990	1.000	1.000
	KW	0.984	1.000	1.000	FD	0.954	0.999	1.000

표 4.3 및 표 4.4의 결과는 주어진 모집단의 모든 유형에 대하여 순위변환한 FR 검정은 모수적 F 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 보다 일반적으로 높은 전체실험검정력을 나타냄을 보여주고 있다. 아울러 전체적인 모집단간의 분포차이를 보다 명백히 가정한 표 4.2로 제시된 모집단 조건하에서의 검정력 결과인 표 4.4가 표 4.1로 제시된 모집단 조건하에서의 검정력 결과인 표 4.3보다 상대적으로 높은 전체실험검정력을 보인다. 또한 모수적 F 검정, 순위변환한 FR 검정 및 Kruskal-Wallis 검정, Friedman 검정은 블럭의 크기를 증가할수록 빠르게 동일한 전체실험검정력으로 점근함을 알 수 있다.

5. 개별쌍검정력의 결과분석

만일 전체모집단간의 차이가 존재하지 않는다는 귀무가설 (2.2)를 기각한다면, 비로서 표 4.1 및 표 4.2로 가정한 모집단 조건을 이용하여 모평균간에 차이가 존재하는 모든 두 모집단간의 개별쌍에 대하여 개별쌍검정력을 구할수 있을 것이다. 근본적으로 이러한 개별쌍검정력의 계산을 위하여 본 논문에서는 최소유의차(LSD : least significant difference)검정을 적용하였다.

즉 3절에서 정의한 모수적 ANOVA F 검정, 순위변환한 FR 검정 및 Kruskal-Wallis 검정, Friedman 검정기법을 이용하여 만일 모든 k 모집단간의 모평균이 동일하다는 귀무가설을 기각하는 경우에 한하여, 추가적으로 모평균이 다른 모든 두 모집단간의 개별쌍들에 0.05의 유의수준하

에서 최소유의차검정을 응용한 아래의 식 (5.1), (5.2) 및 (5.3), (5.4)를 각각 적용함으로써 개별쌍 검정력을 얻을수 있을 것이다.

① : 첫번째로 모수적 ANOVA F 검정에 이어서, 만일 아래의 Fisher 유의차검정식 (5.1)을 만족한다면 i 번째와 j 번째 모집단간에는 모평균 차이가 존재한다고 말할 수 있다. 다시 말하여 Fisher 유의차검정(FSD : Fisher's significant difference)은 단지 일차적인 모수적 ANOVA F 검정에 의하여 귀무가설 (2.2)를 기각하는 경우에만 추가적인 개별쌍 비교를 위하여 최소유의차검정을 이용하는 기법으로 정의한다.

$$| \overline{X}_{i.} - \overline{X}_{j.} | > t_{1-\frac{\alpha}{2},\,\mathrm{df}} \cdot \sqrt{\frac{2\mathrm{MSE}}{n(\text{혹은 } b)}} \tag{5.1}$$

이때 식 (5.1)의 $\overline{X}_{i.}$ 및 $\overline{X}_{j.}$ 는 각각 i 번째와 j 번째 모집단의 표본평균을 나타내며, $t_{1-\alpha/2,\,\mathrm{df}}$ 는 α 의 유의수준하에서 자유도 df 로서 완전확률화모형일 경우 $k(n-1)$ 이며 랜덤화블럭모형일 경우 $(b-1)(k-1)$ 인 t 분포의 $1-\alpha/2$ 분위수를, 그리고 MSE는 ANOVA F 검정통계량의 분모부분인 평균오차제곱합을 의미한다.

② : 두번째로 순위변환한 FR 검정에 이어서, 만일 아래의 순위변환을 이용한 최소유의차검정식 (5.2)를 만족한다면 i 번째와 j 번째 모집단간에는 모평균 차이가 존재한다고 말할 수 있다. 즉 만일 순위변환 자료를 이용한 일차적인 FR 검정에 의하여 귀무가설 (2.2)를 기각하는 경우에만 추가적인 개별쌍 비교를 위하여 순위변환 자료에 대한 최소유의차검정을 적용하며, 이는 간단히 순위변환을 이용한 Fisher 유의차검정기법으로 정의할 수 있다.

$$| \overline{R}_{i.} - \overline{R}_{j.} | > t_{1-\frac{\alpha}{2},\,\mathrm{df}} \cdot \sqrt{\frac{2\mathrm{MSE}}{n(\text{혹은 } b)}} \tag{5.2}$$

단 식 (5.2)의 $\overline{R}_{i.}$ 및 $\overline{R}_{j.}$ 는 각각 i 번째와 j 번째 모집단으로부터 추출한 표본의 순위합을 표본크기로 나눈 평균순위로 정의한다. 이때의 순위는 처리 또는 블럭에 상관없이 결합된 원자료를 가장 작은 값부터 가장 큰 값까지 크기순서의 오름차순으로 배정한 결합순위임을 유의하자. 아울러 MSE는 순위변환을 이용한 FR 검정통계량의 분모부분인 평균오차제곱합을 의미한다.

③ : 세번째로 대표적인 비모수 검정기법인 Kruskal-Wallis 검정, Friedman 검정에 이어서, 만일 아래의 Kruskal-Wallis 순위, Friedman 순위(주어진 블럭내에서 오름차순으로 배정한 순위)를 이용한 최소유의차검정식 (5.3)을 만족한다면 i 번째와 j 번째 모집단간에는 모평균 차이가 존재한다고 말할 수 있다.

즉 만일 Kruskal-Wallis 순위를 이용한 일차적인 Kruskal-Wallis 검정 KW 및 Friedman 순위

를 이용한 일차적인 Friedman 검정 FD 에 의하여 귀무가설 (2.2)를 기각하는 경우에만 추가적인 개별쌍 비교를 위하여 Kruskal-Wallis 순위 및 Friedman 순위에 대한 최소유의차검정을 적용하며, 이는 간단히 Kruskal-Wallis 순위 및 Friedman 순위를 이용한 Fisher 유의차검정기법으로 정의할 수 있다. 한편 완전확률화모형의 개별쌍 검정을 위한 아래식 (5.3)의 KW 는 Kruskal and Wallis(1952)에 의하여 정의된 통계량을 의미하며, 랜덤화블럭모형 개별쌍 검정을 위한 (5.4)는 평균오차제곱합 MSE의 결과를 (5.2)에 치환함으로써 구할 수 있다.

한편 비모수적 다중비교에서는 정확한 표본분포의 유도가 어려우므로 정규근사화를 인용하는 방법이 일반적이나, Conover(1980, p231 및 p300)에 의하여 제안된 바와 같이 아래식 (5.3) 및 (5.4)는 t 분포를 사용함을 유념하도록 하자.

$$|R_{i.} - R_{j.}| > t_{1-\frac{\alpha}{2},\, k(n-1)} \cdot \sqrt{2n \cdot S^2 \cdot \frac{nk-1-\mathrm{KW}}{k(n-1)}} \tag{5.3}$$

$$|R_{i.} - R_{j.}| > t_{1-\frac{\alpha}{2},\, (b-1)(k-1)} \cdot \sqrt{\frac{2b \cdot [\sum_{i=1}^{k}\sum_{j=1}^{b} R_{ij}^2 - \sum_{i=1}^{k} R_{i.}^2 / b]}{(b-1)(k-1)}} \tag{5.4}$$

$$\text{단} \quad S^2 = \frac{1}{nk-1}\left[\sum_{i=1}^{k}\sum_{j=1}^{n} R_{ij}^2 - \frac{nk(nk+1)^2}{4}\right]$$

표 5.1 및 표 5.2는 각각 표 4.1 및 표 4.2에서 제시한 모집단 조건들에 기초한 개별쌍비교의 검정력 결과를 나타낸다. 표 5.1과 표 5.2의 두번째 및 네번째 열 "통계량"은 앞서 살펴본 개별쌍비교를 위한 세가지 검정유형을 나타낸다 : "FSD"는 완전확률화모형 및 랜덤화블럭모형하에서 모수적 ANOVA F 검정에 이은 최소유의차검정인 Fisher 유의차검정을 의미하며, "RT"는 순위변환을 이용한 최소유의차검정을 의미하고, "KW-LSDR"은 Kruskal-Wallis 검정에 이은 최소유의차검정을 또한 "FD-LSDR"은 Friedman 검정에 이은 최소유의차검정을 의미한다.

결국 표 5.1 및 표 5.2는 각각 표 4.1 및 표 4.2의 모집단 조건하에서 일차적으로 귀무가설 (2.2)가 기각될때 이루어지는 세가지 개별쌍비교 기법에 의한 검정력을 보여주고 있다. 이때 표 5.1 및 표 5.2의 개별쌍검정력은 매상황마다 고려된 모든 개별쌍의 검정력들을 고려된 모든 쌍의 수로 나눈 평균값의 개념이며, 개별쌍검정력은 전체실험검정력보다 관측값의 증가 또는 블럭수준의 증가에 따른 검정력의 증가속도가 빠르지 못하므로 검정력의 효율성을 쉽게 비교하기 위하여 $b = 6, 20, 30$ 인 경우의 시뮬레이션 결과를 중심으로 검토하고자 한다.

표 5.1 : 표 4.1 에 제시된 모집단 조건 I 하에서의 ANOVA F_1 , F_2 검정, 순위변환한 FR_1 , FR_2 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 에 이은 최소유의차검정에 의한 개별쌍검정력

모집단	통계량	완전확률화모형			통계량	랜덤화블럭모형		
		$n=6$	$n=20$	$n=30$		$b=6$	$b=20$	$b=30$
정규분포	FSD	0.294	0.294	0.408	FSD	0.296	0.293	0.408
	RT	0.403	0.498	0.576	RT	0.412	0.493	0.558
	KW-LSDR	0.428	0.500	0.576	FD-LSDR	0.283	0.385	0.367
로그정규분포	FSD	0.343	0.372	0.434	FSD	0.346	0.372	0.433
	RT	0.417	0.596	0.688	RT	0.423	0.582	0.665
	KW-LSDR	0.439	0.597	0.688	FD-LSDR	0.287	0.530	0.506
지수분포	FSD	0.285	0.366	0.367	FSD	0.290	0.365	0.367
	RT	0.463	0.687	0.807	RT	0.462	0.675	0.818
	KW-LSDR	0.486	0.687	0.807	FD-LSDR	0.326	0.603	0.582
균일분포	FSD	0.913	0.993	0.999	FSD	0.913	0.993	0.999
	RT	0.957	1.000	1.000	RT	0.970	1.000	1.000
	KW-LSDR	0.962	1.000	1.000	FD-LSDR	0.806	1.000	1.000

표 5.2 : 표 4.2 에 제시된 모집단 조건 II 하에서의 ANOVA F_1 , F_2 검정, 순위변환한 FR_1 , FR_2 검정 및 Kruskal-Wallis 검정 KW, Friedman 검정 FD 에 이은 최소유의차검정에 의한 개별쌍검정력

모집단	통계량	완전확률화모형			통계량	랜덤화블럭모형		
		$n=6$	$n=20$	$n=30$		$b=6$	$b=20$	$b=30$
정규분포	FSD	0.261	0.296	0.452	FSD	0.263	0.294	0.451
	RT	0.382	0.524	0.613	RT	0.388	0.513	0.592
	KW-LSDR	0.404	0.525	0.613	FD-LSDR	0.261	0.444	0.429
로그정규분포	FSD	0.351	0.439	0.500	FSD	0.355	0.438	0.500
	RT	0.437	0.690	0.774	RT	0.438	0.678	0.758
	KW-LSDR	0.449	0.690	0.774	FD-LSDR	0.286	0.579	0.569
지수분포	FSD	0.238	0.328	0.332	FSD	0.245	0.328	0.332
	RT	0.416	0.616	0.726	RT	0.417	0.605	0.714
	KW-LSDR	0.436	0.616	0.726	FD-LSDR	0.287	0.530	0.506
균일분포	FSD	0.456	0.496	0.499	FSD	0.457	0.496	0.499
	RT	0.866	1.000	1.000	RT	0.854	0.861	1.000
	KW-LSDR	0.869	1.000	1.000	FD-LSDR	0.627	0.987	0.996

표 5.1 및 표 5.2의 결과는 순위변환(RT)을 이용한 개별쌍검정력은 모수적 Fisher 유의차검정(FSD) 및 Friedman 순위를 이용한 최소유의차검정(FD-LSDR)의 개별쌍검정력보다 뛰어나며, Kruskal-Wallis 순위를 이용한 최소유의차검정(KW-LSDR)과 상당히 비슷한 수준임을 알 수 있다. 게다가 관측값의 크기 또는 블럭크기를 증가할수록 순위변환 기법의 명백한 검정력 우위를 알 수 있으며, 특히 로그정규분포, 지수분포 및 균일분포의 비정규분포하에서 뛰어남을 알 수 있다. 그런데 Friedman 순위를 이용한 최소유의차검정의 개별쌍검정력이 블럭의 크기가 증가하여도 상대적으로 빠르게 증가하지 않는 이유는 처리수준이 적은 상태에서 주어진 블럭내에서 할당되는 Friedman 순위배정의 특성으로 말미암아 기인함을 예상할 수 있다. 이에 반해 전체 관측치에 대하여 오름차순으로 결합순위를 할당하는 순위변환 기법을 이용할 때 발생하는 높은 개별쌍검정력은 순위변환의 편리함 및 쉬운 이해도와 함께 순위변환의 큰 장점을 대변한다고 할 수 있다.

한편 실례로서 SAS 통계패키지를 이용한다면 일차적으로 PROC RANK 프로시져를 사용하여 완전확률화모형 및 랜덤화블럭모형의 순위변환을 이용한 순위할당을 배정한 후 PROC GLM 혹은 PROC ANOVA 프로시져의 옵션 LSD를 사용하여 간단히 다중비교를 수행할 수 있을 것이다. 물론 모집단간의 개별쌍에 대한 다중비교시 순위변환을 이용한 최소유의차검정의 분석기법이 본 논문에서 인용하지 않은 다른 방법들보다도 절대적으로 우수하므로 일반적으로 권장할 방법인지에 대해서는 앞으로 지속적인 연구가 있어야 할 것이다.

6. 결 론

본 시뮬레이션 연구결과는 완전확률화모형 및 랜덤화블럭모형에 대한 다중비교에 있어서 순위변환과 최소유의차검정을 이용한 분석기법이 모수적 ANOVA F 검정과 Fisher의 유의차검정을 이용한 기법 및 비모수적 Kruskal-Wallis 검정, Friedman 검정과 최소유의차검정을 이용한 분석기법보다 전체실험오차율에 있어서 명목상의 유의수준을 상당히 잘 유지하고 있으며, 전체실험검정력 및 개별쌍검정력 면에 있어서도 상당히 뛰어난 기법임을 말해주고 있다.

이와같은 시뮬레이션 결과는 순위변환을 이용한 기법의 간편성과 적용하기 쉬운 장점으로 인하여 앞으로 다양한 실험계획모형하에서의 순위변환을 이용한 다중비교의 활용을 증대시킬 수 있을 것이다.

참 고 문 헌

[1] Andrews, D. M. and David, H. A. (1990). Nonparametric Analysis of Unbalancd Paired Comparison or Ranked Data, *Journal of the American Statistical Association*, Vol. 85, 1140-1146.

[2] Carmer, S. G. and Swanson, M. R. (1973). An Evaluation of Ten Pairwise Multiple Comparison Procedures by Monte Carlo Methods, *Journal of the American Statistical Association*. Vol. 68. 66-74.

[3] Choi, Y. H. (1998). A Study of the Power of the Rank Transform Test in a 2^3 Factorial Experiment, *Communications in Statistics - Simulation and Computation*, Vol. 27, No 1, In Press.

[4] Conover, W. J. (1980). *Practical Nonparametric Statistics*, John Wiley & Sons Inc., New York.

[5] Conover, W. J. and Iman, R. L. (1979). On Multiple Comparison Procedures, *Technical Report* LA-7677-MS, Los Alamos Scientific Laboratory.

[6] David, H. A. (1987). Ranking From Unbalanced Paired-Comparison Data, *Biometrika*, Vol. 74, 432-436.

[7] David, H. A., Lachenbruch, P. A. and Brandis, H. P. (1972). The Power Function of Range and Studentized Range Test in Normal Samples, *Biometrika*, Vol. 59, 161-168.

[8] Einot, I. and Gabriel, K. R. (1975). A Study of the Powers of Several Methods of Multiple Comparisons, *Journal of the American Statistical Association*, Vol. 70, 574-583.

[9] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journal of the American Statistical Association*, Vol. 32, 675-701.

[10] Hayter, A. J. (1989). Pairwise Comparisons of Generally Correlated Means, *Journal of the American Statistical Association*, Vol. 84, 208-213.

[11] Hochberg, Y. and Varon, Y. (1984). On Simultaneous Pairwise Comparisons in Analysis of Covariance, *Journal of the American Statistical Association*, Vol. 79, 863-866.

[12] Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks on one-criterion variance analysis, *Journal of the American Statistical Association*, Vol. 47, 583-621.

[13] Kunert, J. (1990). On the Power of Tests for Multiple Comparisons of Three Normal Means, *Journal of the American Statistical Association*, Vol. 85, 808-812.

[14] Lehmann, E. L. and Shaffer, J. P. (1977). On a Fundamental Theorem in Multiple Comparisons, *Journal of the American Statistical Association*, Vol. 72, 576-578.

[15] Lin, F. A. and Haseman, J. K. (1978). An Evaluation of Some Nonparametric Multiple Comparison Procedures by Monte Carlo Methods, *Communications in Statistics - Simulation and Computation*, Vol. 7, 117-128.

[16] Ramsey, P. H. (1978). Power Differences Between Pairwise Multiple Comparisons, *Journal of the American Statistical Association*, Vol. 73, 479-487.

[17] Uusipaikka, E. (1985). Exact Simultaneous Confidence Intervals for Multiple Comparisons Among Three or Four Mean Values, *Journal of the American Statistical Association*. Vol. 80. 196-201.