【참고】 일원분산분석의 사후비교분석

① 일원분산분석에서 귀무가설이 기각되었다.
= 비교집단들의 모집단의 평균에 차이가 있다.
= 여러 모집단 평균들과의 많은 대비 중 최소한 하나(at lest one comparison) 이상의 대비(contrast)가 통계적으로 0이 아니다.
- H_0 : 모든 대비는 0이다.
- H_1 : 최소한의 하나 이상의 대비가 0이 아니다.

[예] (설명식, 시청각, 개인) 교수법에 대한 분산분석에서 6개의 대비 중 최소한 하나의 대비라도 0이 아니다.
㉠ 설명식 교수법 vs 시청각 교수법
㉡ 설명식 교수법 vs 개인 교수법
㉢ 시청각 교수법 vs 개인 교수법
㉣ 설명식 교수법 vs (시청각 교수법+개인 교수법)
㉤ 시청각 교수법 vs (설명식 교수법+개인 교수법)
㉥ 개인 교수법 vs (설명식 교수법+시청각 교수법)

② 그러나 어느 집단들 간에 차이가 있는지를 알 수는 없다. 이 문제는 Scheffé (1959)에 의하여 해답이 풀리기 시작했다.
③ 만약 귀무가설이 기각되었다면 어떤 대비에서 차이가 나타나는지를 찾아내는 통계적 방법인 사후비교분석(post-hoc comparison analysis)을 진행할 수 있다.
④ 귀무가설이 기각되어 사후검사를 실시할 때 연구자가 가능한 모든 대비에 대한 차이 검정을 실시할 필요는 없다. 관심의 대상이 되는 대비만 검정할 수 있다.

[예] Scheffé 방법 ; 모든 짝비교 검정 방법

[다중 비교]

종속변수: 수학점수
Scheffe

(I)교수법	(J)교수법	평균차(I-J)	표준오차	유의확률	95% 신뢰구간 하한값	95% 신뢰구간 상한값
설명식교수법	시청각교수법	−2.50	1.46	.282	−6.77	1.77
	개인교수법	−4.50*	1.28	.021	−8.25	−.75
시청각교수법	설명식교수법	2.50	1.46	.282	−1.77	6.77
	개인교수법	−2.00	1.40	.398	−6.08	2.08
개인교수법	설명식교수법	4.50*	1.28	.021	.75	8.25
	시청각교수법	2.00	1.40	.398	−2.08	6.08

* .05수준에서 평균차가 큼

[수학점수]

Scheffe[a, b]

교수법	N	유의수준=.05에 대한 부집단	
		1	2
설명식	4	2.50	
시청각	3	5.00	5.00
개인	5		7.00
유의확률		.248	.391

동일 집단군에 있는 집단에 대한 평균이 표시됩니다.
a. 조화평균표본크기 = 3.830을(를) 사용
b. 집단크기가 같지 않습니다. 집단크기의 조화평균이 사용됩니다. 제1종 오류 수준은 보장할 수 없습니다.

㉠ 짝비교 결과, 설명식 방법과 개인교수법만 통계적으로 유의한 차이를 보이는 것으로 나타남
㉡ 즉, 설명식 교수법과 시청각 교수법은 유의한 차이가 없으므로 하나의 동일집단으로, 시청각 교수법과 개인교수법도 유의한 차이가 없으므로 하나의 동일집단으로 볼 수 있다는 의미.

⑤ 그러나! 귀무가설이 기각되지 않았을 경우 사후비교분석을 실시할 수 없다.

5-3. 등분산 가정 검정

① 분산분석을 하기 전에 두 가지 가정 '정규성 가정'과 '등분산 가정'을 확인해야 한다.
② 만약 정규성 가정을 만족하면 모수적 방법인 분산분석을 선택하고 정규성 가정을 만족하지 않으면 비모수적 방법인 Krusakal-Wallis검정을 실시해야 한다.
③ 등분산이란 그룹간의 분산이 서로 같다는 의미이다. 분산분석은 집단간 변동과 집단내 변동을 이용해서 분석하는데, 집단간 변동이 집단내 변동보다 크다면 집단간에 차이가 있다고 결정하는 방법이다. 따라서 분산분석은 집단간 분산의 동질성 가정에 민감한 편이다. 그리고 등분산 가정이 만족되는 경우에만 분산분석을 수행하는 것이 좋고, 만약 그렇지 않다면 Welch's ANOVA를 수행하는 것이 좋다.
④ 등분산 가정을 확인하는 검정으로는 레빈의 검정(Levene's test)과 바틀렛 검정(Bartlett's test) 그리고 F검정이 대표적이다.
⑤ Levene의 검정에 의한 방법은 Brown and Forsythe (1974)이 Levene의 절차(Levene, 1960)를 수정한 것이다. 이 방법에서는 표본 평균 대신 표본 중위수로부터 관측치까지의 거리를 사용한다. 표본 평균 대신 표본 중위수를 사용하면 표본 크기가 작을수록 보다 견고한 검정이 수행된다.
⑥ 데이터를 정규 분포에서 얻은 경우에는 Bartlett의 검정을 사용한다. Bartlett의 검정 통계량은 자유도를 기반으로 각 표본 분산의 가중 산술 평균과 가중 기하 평균을 계산한다. 평균의 차이가 클수록 표본의 분산이 같지 않을 확률이 높다.
⑦ 데이터가 정규분포를 따르며 수준이 두 개뿐이면 F검정에 의한 방법을 사용한다. 이때, F값이 1에 가까우면 두 모집단의 분산은 다르지 않음을 예견할 수 있으며, 매우 크거나 혹은 0에 가까운 작은 값을 가지면 두 모집단의 분산이 다를 것이라고 예견할 수 있다.

[예] A교수법(5명)과 B교수법(6명)의 학업성취도 점수 비교에서, 두 모집단의 분산이 같은지 유의수준 0.05에서 검정하여라.

A 교수법	B 교수법
3	7
4	6
6	8
5	7
5	5
	8

[풀이]

㉠ 가설세우기

H_0 : 두 모집단의 분산은 같다.

$$\sigma_A^2 = \sigma_B^2$$

H_1 : 두 모집단의 분산은 같지 않다.

$$\sigma_A^2 \neq \sigma_B^2$$

㉡ 두 교수법에 의한 분산

A교수법에 의한 학업성취 점수의 분산$=1.3$, 5명

B교수법에 의한 학업성취 점수의 분산$=1.366$, 6명

㉢ F통계값 $F=\dfrac{s_1^2}{s_2^2}=\dfrac{1.3}{1.366}=0.952$

㉣ 결론 : 유의수준 0.05와 $\nu_1=4$, $\nu_2=5$인 F분포의 양방적 검정의 기각값은 각각 $F_{0.025}(4,\ 5)=7.39$, $F_{0.975}(5,\ 4)=0.017$이므로 귀무가설을 기각할 수 없다. 즉, '두 모집단의 분산은 같다'고 결론 내릴 수 있다.