

● 분산분석

○ 분산분석 > 일원배치분산분석

- 하나의 요인을 가지는 분산분석 모형에 대한 추론 결과를 제공함
- 분석품
  - '변수목록'에서 '분석변수' 목록과 '요인'에 해당 변수를 전달함. 분석변수 목록에 여러 개가 있으면 각각의 변수에 대해 분산분석을 실시함
  - 요인의 성질에 따라 '고정효과(fixed effect)'와 변량효과(random effect)'를 선택할 수 있으며 변량효과를 선택하면 다중비교와 평균추정을 하지 않음
  - 잔차검진에서 등분산성검정 방법으로 'Levene'와 'Bartlett' 방법을 지원하고 정규성검정은 Shapiro-Wilk 검정을 실시함
  - 다중비교 방법으로 Fisher의 LSD, Bonferroni, Scheffe 방법이 지원됨
  - '언어변환'은 같은 결과를 얻기 위한 'R' 또는 'SAS' 명령어를 출력해 줌
- 결과정리표에는 별도의 작업을 하지 않고 보고서 첨부할 수 있는 형태의 결과표를 제공함
- 변량효과를 선택했을 때에는 아래와 같은 모수추정값들을 제공함



[그림 1] 일원배치 분산분석 분석품

【분석결과 예제】

변량효과모형 모수추정				
변수	평균	오차분산	처리분산	ICC
Y	48.192	33.264	40.676	0.290

일원배치분산분석: 예제8-7(고급)						
분산분석표(ANOVA Table)						
변수	요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
Y	모형	2	81.352	40.676	1.223	0.339
	오차	9	299.378	33.264		
등분산성검정(Levene 검정)						
변수	MSTR	MSE	검정통계량	자유도	p-값	
Y	0.041	11.228	0.004	2,9	0.996	
정규성검정(Shapiro-Wilk 검정)						
변수	표본크기	왜도	첨도	검정통계량	p-값	
Y	12	-0.342	1.891	0.929	0.375	
다중비교: Scheffe						
변수	비교대상	평균차	신뢰하한	신뢰상한	차이여부(O:차이있음, X:차이없음)	
Y	A vs B	-6.375	-18.274	5.524	X	
	A vs C	-3.350	-15.249	8.549	X	
	B vs C	3.025	-8.874	14.924	X	
결과정리표						
변수	수준	표본크기	평균	표준편차	F-통계량	p-값
Y	A	4	44.950	5.262	1.223	0.339
	B	4	51.325	6.241		
	C	4	48.300	5.758		

○ 분산분석 > 이원배치분산분석

- 요인이 두 개인 분산분석 모형에 대한 추론 결과를 제공함
- 분석품
  - '변수목록'에서 분석할 변수와 요인들을 선택하여 해당 목록이나 텍스트박스로 이동시킴. '분석변수' 목록에 여러 개의 변수가 있으면 각각의 변수에 대해 이원배치 분산분석을 실시함
  - '모형설정'에서 요인들이 고정효과인지, 변량효과인지를 선택하고 반복이 있는 경우 상호작용항을 포함할 것인지를 선택함
  - 유효효과선택을 선택하면 상호작용이 있는 경우 상호작용의 유효성을 검정하고 유효하면 종료, 아니면 상호작용을 제거하고 주효과에 대한 검정을 실시하여 최종모형을 선택함
  - '잔차검진'에서는 잔차에 대해 Levene, Bartlett에 의한 등분산성, Shapiro-Wilk에 의한 정규성 검정을 선택함
- 혼합효과나 변량효과모형의 경우 주효과에 대한 유의성 검정은 MSE와 비교하는 것이 아니라 상호작용항의 MS와 비교함
- 상호작용항이 있는 모형에서 반복수가 다른 불균형 자료이면 SS계산에서 Type I, II, III를 모두 제공함

이원배치분산분석 V1.0

변수목록: [빈 공간]

분석변수: 활동지수

요인: 아동 (요인1), 투여약 (요인2)

모형설정: 요인1 (고정효과), 요인2 (고정효과), 상호작용항 포함, 유효효과선택

잔차검진: 등분산성검정 (Levene), 정규성검정(Shapiro-Wilk)

추가분석: 다중비교 (Fisher LSD), 신뢰수준 (95 %), 모수추정

언어변환: R, SAS, Python

[그림 2] 이원배치 분산분석 분석품

【분석결과 예제】

이원배치분산분석: 예제8-8(고급)(활동지수)					
분석모형: 균형자료 & 고정효과모형					
분산분석표					
요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
아동	1	121.000	121.000	8.000	0.015
투여약	1	42.250	42.250	2.793	0.121
상호작용	1	930.250	930.250	61.504	0.000
오차	12	181.500	15.125		
전체	15	1275.000			
위의 분산분석결과가 최종모형입니다.					

이원배치분산분석: 예제8-8(고급)(활동지수)					
분석모형: 불균형자료 & 고정효과모형					
분산분석표					
요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
처리효과	3	1002.567	334.189	20.291	0.000
오차	11	181.167	16.470		
전체	14	1183.733			
Type I					
요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
아동	1	78.019	78.019	4.737	0.052
투여약	1	75.214	75.214	4.567	0.056
상호작용	1	849.333	849.333	51.569	0.000
Type II					
요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
아동	1	67.089	67.089	4.073	0.069
투여약	1	75.214	75.214	4.567	0.056
상호작용	1	849.333	849.333	51.569	0.000
Type III					
요인	자유도	제곱합(SS)	균제곱합(M)	F-통계량	p-값
아동	1	108.333	108.333	6.578	0.026
투여약	1	41.026	41.026	2.491	0.143
상호작용	1	849.333	849.333	51.569	0.000

## ○ 분산분석 > 공분산분석 모형

- 공변량(covariate)과 요인(factor)이 있는 공분산분석 결과를 제공하며 두 변수의 상호작용(interaction)항을 모형에 추가할 수 있도록 함. 분산분석메뉴에도 동일한 공분산분석모형 분석도구를 추가함

공분산분석 V1.0 ×

**변수목록**

id

**반응변수**

salary

**공변량(Covariate)**

educ  
salbegin  
jobtime  
prevexp

**요인(Factor)**

gender  
jobcat  
minority

**분석** **출력옵션** **재설정**

**도움말** **종료**

**상호작용(Interaction)**

educ \* salbegin  
gender \* prevexp  
jobcat \* minority

**추가** **삭제**

**모형설정 및 분석**

☒ 절편포함  
  
☒ 잔차검진

☒ 반응변수변환 자연로그

☒ 유의성검정

☐ 예측값-잔차 출력

**언어변환**

☐ R
 ☐ SAS
 ☐ Python

[그림 3] 공분산분석 분석폼

- 분석폼 '변수목록'에서 분석할 반응변수와 공변량 및 요인들을 선택하여 해당 텍스트 박스나 목록으로 이동시킴. 공변량과 요인 목록에서 두 변수를 선택하여 상호작용 좌측의 추가 버튼을 누르면 두 변수가 결합된 변수가 상호작용 목록에 추가되고 상호작용에서 선택한 것을 제거 버튼을 통해 제거할 수 있음
- 공변량이나 요인 변수 중 변수목록으로 이동시키면 해당변수가 포함된 상호작용항을 자동 삭제됨
- 절편을 포함한 것으로 가정하고 모든 분석을 실시함
- 유의성 검정의 경우 완전모형(full model)과 각 변수를 제거한 축소모형(reduced model) 간의 설명력을 비교하는 F-검정을 실시한 결과를 출력함
- 잔차검진에서는 잔차에 대해 Levene에 의한 등분산성, Shapiro-Wilk, Jacque-Bera에 의한 정규성 검정을 실시함
- 예측값-잔차 출력을 통해 분석결과의 예측값과 잔차를 원자료의 시트 마지막 열 다음에 출력할 수 있음
- 분석결과는 전체 설명변수가 포함된 분산분석표와 개별 변수에 대한 회귀계수에 대한 t-검정 결과를 기본적으로 출력해 줌. 요인의 경우 수준을 가나다(알파벳)순으로 정렬하여 가정 앞에 있는 수준을 기준으로 분석을 진행함. 출력된 결과 값의 자리수는 출력옵션에서 수정할 수 있으며 디폴트는 소수점 3자리수까지 표시하도록 함

【분석결과 예제】

공분산분석결과: data					
분석모형: salary = 절편+educ+salbegin+jobtime+prevexp+gender+jobcat+minority+educ * salbegin+gender * prevexp+jobcat * minority					
분산분석표					
소스	자유도	제곱합(SS)	평균제곱(MF-통계량	p-값(F >)	
모형	12	63.170	5.264	210.936	0.000
오차	461	11.505	0.025		
전체	473	74.675			
결정계수	0.846				
수정결정계수	0.842				

회귀계수추론				
설명변수	회귀추정량	표준오차	T-통계량	p-값
절편	9.135	0.147	62.302	0.000
educ	0.055	0.007	7.587	0.000
salbegin	0.000	0.000	7.454	0.000
jobtime	0.004	0.001	5.142	0.000
prevexp	-0.001	0.000	-6.626	0.000
gender_여성	-0.119	0.023	-5.272	0.000
jobcat_관리직	0.031	0.064	0.478	0.633
jobcat_사무직	-0.267	0.032	-8.327	0.000
minority_Yes	0.117	0.084	1.399	0.162
educ:salbegin	0.000	0.000	-4.715	0.000
prevexp:gender_여성	0.000	0.000	1.600	0.110
jobcat_관리직:minority_Yes	-0.212	0.105	-2.017	0.044
jobcat_사무직:minority_Yes	-0.174	0.086	-2.009	0.045

설명변수 유의성검정					
설명변수	자유도	SS	MS	F-통계량	p-값
educ	1	1.436	1.436	57.560	0.000
salbegin	1	1.387	1.387	55.568	0.000
jobtime	1	0.660	0.660	26.443	0.000
prevexp	1	1.096	1.096	43.904	0.000
gender	1	0.694	0.694	27.792	0.000
jobcat	2	2.596	1.298	52.015	0.000
minority	1	0.049	0.049	1.958	0.162
educ * salbegin	1	0.555	0.555	22.229	0.000
gender * prevexp	1	0.064	0.064	2.559	0.110
jobcat * minority	2	0.114	0.057	2.277	0.104
※ 완전모형과 해당변수 제거한 축소모형 간의 설명력 비교					

잔차검진			
가정	검정방법	검정통계량	p값
정규성	Shapiro-W	0.980	0.000
정규성	Jacque-Bei	91.744	0.000
등분산성	Breusch-Pa	17.686	0.126