

연습문제 요약

©심송용

1. 기술통계 및 용어 등에 대한 연습문제는 교재-한림강의노트 시리즈, 한림대학교 출판부 033-248-2875- 참고
2. 지속성 안면 마비와 같은 신경계통의 치료를 하면 말을 더듬는 사람에게 효과가 있다고 한다. 이 신경계통의 치료를 위하여 아드레날린을 저하시켜 주는 프로프라노롤(propranolol)과 같은 약제를 사용하여 그 효과를 검증하고자 한다. 이를 위해 말을 더듬는 환자 10 명을 대상으로 2주간 약을 투여하고 투여하기 전과 후의 말 더듬는 정도를 측정하였더니 다음과 같은 결과를 얻었다.

환자	1	2	3	4	5	6	7	8	9	10
치료전	5	1	0	4	4	5	3	3	2	4
치료후	4	0	0	2	1	3	3	1	2	1

- (a) 이 약의 효과가 있는지 검정하여 보아라. ($\alpha = 0.01$)
- (b) 모평균 차에 대한 95% 신뢰구간을 구해 보아라.

풀이

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 PRE	3,10	10	1,663	,526
POST	1,70	10	1,337	,423

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 PRE & POST	10	,714	,020

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 PRE - POST	1,40	1,174	,371	,56	2,24	3,772	9	,004

그림 1: SPSS의 두 표본 쌍체 비교 t 검정 결과

- (a) SPSS에서 얻어진 양측 유의확률이 0.004이고 t 값이 양수이며 대립가설이 $H_1 : \mu_{pre} - \mu_{post} > 0$ 이므로 실제 유의확률은 0.004의 반인 0.002가 되어 유의수준 5% 에서 귀무가설을 기각한다.
- (b) 평균차에 대한 95% 신뢰구간은 출력에서 보는 것과 같이 (0.56, 2.24)이며 이 신뢰구간은 보기 2와 같은 결과를 얻을 수 있다.
3. 다음은 한 마리의 쥐의 왼쪽과 오른쪽 귀를 임의로 선택하여 한 쪽 귀에는 에탄올(ethanol)을 다른 쪽 귀에는 캡사이신(capsaicin)을 처리한 후 10분 뒤에 양쪽의 귀의 혈류량을 비교한 자료이다.

처리 \ 귀	1	2	3	4	5	6	7	8	9
에탄올	208	285	181	251	277	281	232	135	240
캡사이신	791	572	604	766	942	664	643	372	559

캡사이신이 에탄올보다 혈류량을 많이 증가시키는지 검정하여라. 또 평균차에 대한 95% 신뢰구간도 구해보아라.

풀이 SPSS의 메뉴에서 분석 → 평균비교 → 짝비교 순서로 메뉴를 선택한 결과 SPSS 출력이

Paired Samples Statistics

Pair	Mean	N	Std. Deviation	Std. Error Mean
1 ETHANOL	232,22	9	50,306	16,769
1 CAPSAI	657,00	9	162,853	54,284

Paired Samples Correlations

Pair 1	N	Correlation	Sig.
ETHANOL & CAPSAI	9	,566	,112

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	ETHANOL - CAPSAI	-424,78	140,647	46,882	-532,89	-316,67	-9,061	8	,000

와 같고 귀무가설 $H_0 : \mu_E - \mu_C = 0$ 대 대립가설 $H_1 : \mu_E - \mu_C < 0$ 의 검정이므로 관찰된 검정통계량 $t = -9.061$ 에 대한 유의확률은 $0.000/2=0.000$ 이므로 귀무가설을 기각한다. 즉, capsaicine을 적용한 경우 혈류량을 증가한다는 근거가 충분하다고 할 수 있다. 출력에 의하면 $\mu_E - \mu_C$ 의 95% 신뢰구간은 (-532.89, -316.67)이다.

4. 다음은 신장결석이 있는 환자와 없는 환자의 요에서 calcium concentration을 조사한 결과이다.(단위: millimoles/litre) 신장결석이 있는 환자와 그렇지 않은 환자 사이에 calcium concentration의 차이가 있다고 할 수 있겠는가?

있음	2.17	0.17	0.83	3.04	1.06	3.93	5.38	3.53	4.54
	3.98	1.02	3.46	1.19	5.64	2.66	1.22	2.64	2.31
	4.49								
없음	6.96	13.00	5.54	6.19	7.31	14.34	4.74	2.50	1.27
	4.18	3.10	3.01	6.81	8.28	2.33	7.18	5.67	12.68
	8.94	3.16	3.30	6.99					

풀이 SPSS의 메뉴에서 분석 → 평균비교 → 독립인 두 표본 순서로 메뉴를 선택한 결과 SPSS 출력이

Group Statistics

결석여부	N	Mean	Std. Deviation	Std. Error Mean
CALCIUM 1	19	2.8268	1.65959	.38074
2	22	6.2491	3.56822	.76075

Independent Samples Test

	Levene's Test for Equality of Variances	t-test for Equality of Means								
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
CALCIUM	Equal variances assumed	6.236	.017	-3.833	39	.000	-3.4222	.89283	-5.22817	-1.61633
	Equal variances not assumed			-4.023	30.598	.000	-3.4222	.85070	-5.15819	-1.68631

와 같고 귀무가설 $H_0 : \mu_1 - \mu_2 = 0$ 대 대립가설 $H_1 : \mu_1 - \mu_2 \neq 0$ 의 검정이다. 등분산성 검정의 결과 유의확률이 0.017로 유의수준 5% 에서 등분산성을 가정할 수 없으므로 $t = -4.023$ 이고 유의확률은 0.000이므로 귀무가설을 기각한다. 즉, 신장결석의 유무에 따른 요의 calcium concentration이 다르다고 할 수 있다.

5. (위 문제의 연속) Calcium concentration의 차이에 대한 신뢰구간을 구해 보고 이 신뢰구간의 결과와 가설검정의 결과가 일치하는지 비교 하여라. 가설검정의 결과와 비교할 신뢰구간의 신뢰도와 앞의 유의수준 사이의 관계를 말하여라.

풀이 위 연습문제의 컴퓨터 출력에서 95% 신뢰구간은 (-5.158, -1.686)이므로 신뢰구간에 0이 포함되지 않으며 이는 유의수준 5% 인 가설검정에서 두 그룹의 평균이 같다고 할 수 없다는 결론과 일치하는 결론이다.

6. 암로디핀(amlodipine) 성분은 혈관의 긴장도를 낮추어 혈압을 낮추는 약효를 나타내며 심장의 부담을 덜어 준다. 암로디핀은 심장에 혈액과 산소의 공급을 증가시켜 줌으로써 협심증 발작을 치료, 예방하는 역할을 하는 것으로 알려져 있으며 고혈압, 협심증의 치료제로 사용하고 있다. 어느 제약회사에서 제조비용이 기존의 방법보다 적게 들면서 치료효과에는 차이가 없을 것으로 믿어지는 새로운 암로디핀 성분의 약을 개발하여 이 약의 효과를 기존의 약과 비교하고자 한다. 이를 위해 확장기 혈압이 90mmHg 이상인 고혈압 환자를 두 그룹으로 나누어 한쪽 그룹은 기존의 약을 처방하고 다른 그룹은 새로 개발한 약을 처방하여 6주 후에 두 그룹의 확장기 혈압을 조사하였더니 다음과 같았다.

기준 약	89 78 82 94 81 82 70 93 79 86 87 81 80 89 79 78 84
신약	85 89 75 85 99 74 82 92 92 81 77 85 77 84 83 88 85 78 88 82 94 77 79 84 73 83

두 약의 치료효과가 다르다고 할 수 있겠는가? 확장기 혈압의 차이에 대한 95% 신뢰구간을 구해 보고 이 구간에 0이 포함되는지 확인하여 가설검정의 결과와 비교하여 보아라.

풀이 그림 2의 SPSS 출력을 보면 등분산이 가정을 한 후 귀무가설

집단통계량

약의종류	N	평균	표준편차	평균의 표준오차
확장기혈압 기준약	17	83.0588	6.06703	1.47147
신약	26	83.5000	6.43894	1.26278

독립표본검정

	Levene의 등분산검정		평균의 동일성에 대한 t-검정						
	F	유의확률	t	자유도	유의확률 (양측)	평균차	차이의 표준오차	차이의 95% 신뢰구간	
								하한	상한
확장기혈압 등분산이 가정됨	.025	.876	-.225	41	.823	-.44118	1.96389	-4.40733	3.52498
등분산이 가정되지 않음			-.228	35.813	.821	-.44118	1.93903	-4.37442	3.49207

그림 2: 결과물

$H_0 : \mu_1 = \mu_2$ 와 대립가설 $H_1 : \mu_1 \neq \mu_2$ 을 검정하기 위한 출력은 두번째 표의 첫번째 줄을 사용한다. 따라서 양측검정의 유의확률 0.823이 유의수준보다 작으므로 귀무가설을 기각하지 못한다. 즉, 약의 종류에 따라 확장기혈압의 평균이 같다고 주장할 충분한 근거가 된다. 또한 차이의 95%신뢰구간이 (-4.40733, 3.52498)로 이 구간에 0이 포함되지 않아서 귀무가설을 기각하지 못한다. 따라서 이 결과는 위의 가설검정의 결과와 일치한다.

7. 다음은 세 종류의 약을 투여한 후 약효를 측정된 자료이다.

A	13 9 4 9 13 13 8
B	8 10 16 12 9 14 13 12
C(위약)	24 9 4 6 13 8 12 10 2

이 자료를 이용하여

- (a) 세 그룹의 분산이 같은지 검정하여라.
- (b) 위약과 나머지 두 약의 약효가 같은지 검정하여라.(세그룹의 분산이 같다고 가정할 때와 그렇지 않을 때의 결과를 비교하여라.)

풀이:

- (a) 그림 3의 결과에서 보는 것처럼 귀무가설 $H_0 : \sigma_1 = \sigma_2 = \sigma_3$ 와 대립가설 $H_1 : \text{not } H_0$ 을 검정하기 위한 Levene의 등분산 결과 유의확률 0.263으로 유의수준보다 크므로 귀무가설을 기각하지 못한다. 즉, 세 집단의 분산이 같다고 주장할 충분한 근거가 된다. i

분산의 동질성 검정

약효			
Levene 통계량	df1	df2	유의확률
1.425	2	21	.263

그림 3: 등분산검정

(b) 그림 4에서 보는 것처럼 세 분산이 같다고 가정한 경우 유의확률 0.604로 유의수준보다 크므로 귀무가설을 기각하지 못한다. 따라서 위약과 나머지 두 약의 약효는 같다고 주장할 충분한 근거가 된다.

세 분산이 같지 않다고 가정한 경우 유의확률 0.604로 유의수준보다 크므로 귀무가설을 기각하지 못한다. 따라서 위약과 나머지 두 약의 약효는 같다고 주장할 충분한 근거가 된다.

대비검정

약효	대비	대비 값	표준 오차 오류	t	df	유의 확률 (양측)
등분산 가정	1	-2.0516	3.89780	-.526	21	.604
등분산을 가정하지 않습니다.	1	-2.0516	4.56421	-.449	10.221	.662

그림 4: 대비

- MRSA(Methicillin Resistance Staphylococcus Aureus (메치실린이라는 항생제에도 내성이 있는 Staphylococcus Aureus 라는 균)에 감염이 되면 보통의 항생제로는 치료가 되지 않고 반코마이신(Vancomycin)이나 Teicoplanin, Linezolid라는 항균제를 사용하고 있으나 이들에게도 내성이 생기는 경우에는 치료 방법이 없는 실정이다. 또한 반코마이신 등은 신장쪽에 부작용이 생길 수 있다.

나폴레옹 군대의 야전 의사는 구더기의 상처치유력을 기록하고 있으며 1차 대전과 미국의 남북전쟁에서도 구더기가 부상한 병사들의 상처치유에 이용되었다. 구더기는 괴사하거나 손상된 조직만 먹어 치우지 건강한 조직은 파먹지 않는다고 한다. MRSA에 감염된 상처를 가진 환자를 대상으로 네 가지 치료방법을 이용하여 치료효과의 차이를 검정하기 위해서 다음과 같은 자료를 얻었다(가상 자료).

	구더기	Vancomycin	Teicolanin	Linezolid
치료점수	99	105	91	91
	113	105	95	80
	99	128	120	79
	89	123	55	130
	84	62	101	103
	82	103	100	87

네 방법에 따른 차이가 있는가?
 풀이 그림 5의 분산분석표에 따라 귀무가설 $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ 와

분산분석					
치료점수	제곱합	df	평균 제곱	거짓	유의확률
집단-간	455.333	3	151.778	.403	.752
집단-내	7524.000	20	376.200		
합계	7979.333	23			

그림 5: 등분산검정

대립가설 $H_1 : \text{not } H_0$ 을 검정하기 위한 분산분석 결과 유의확률 0.752으로 유의수준보다 크므로 귀무가설을 기각하지 못한다. 즉, 세 집단의 치료점수가 같다고 주장할 충분한 근거가 된다.

9. 다음은 스키장 의무실에서 치료를 필요로 하는 머리부상자의 머리부상 정도와 헬멧 사용여부를 조사한 것이다.

헬멧	부상정도			합
	경상(2주 이하)	중상(4주 이하)	중상(5주 이상)	
착용	10	17	7	34
미착용	23	12	5	40
합	33	29	12	74

이 표에서 헬멧사용 여부와 부상정도가 독립인지 알아보아라.

풀이 SPSS에서 분할표는 분석 → 기술통계량 → 교차분석 순서로 메뉴에서 선택한다. 이 순서로 메뉴를 선택하면 그림 6과 같은 대화창이 만들어지는데 이 표에서 분할표를 만들 행과 열을 지정한다. 이 그림은 자료가 이미 일차로 정리된 자료로서 표의 각 칸에 해당하는 숫자만큼의 빈도수가 있다는 의미이다. 따라서 문제의 표를 SPSS의 데이터 창에 입력을 하게되면 그림 7와 같이 입력된다. 이 그림의 두 개의 변수에 대해서 교차표를 만들면 각각의 경우에 대해서 한 번의 빈도밖에 나타나지 않게 된다. 그림 7의 자료가 의미하는 것은 헬멧과 부상정도의 각각에 값에 해당하는 케이스의 수가 빈도라는 의미이므로 이를 SPSS에 알려주어야 한다. 따라서 표의 빈도수 만큼 가중치를

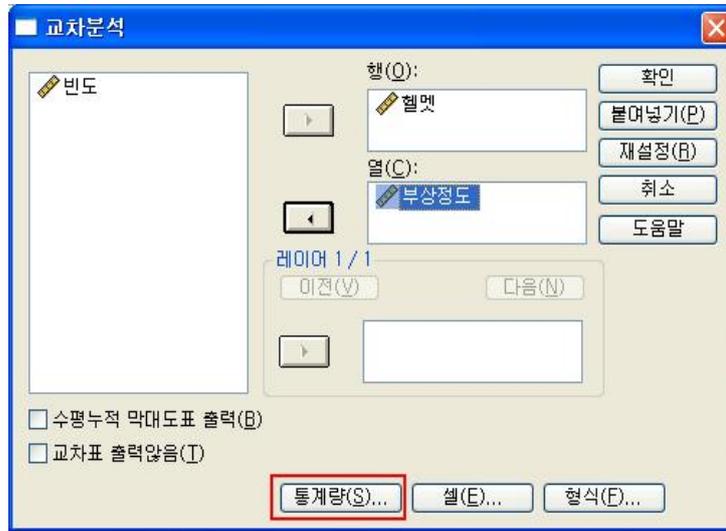


그림 6: SPSS에서의 분할표 대화창 설정

	헬멧	부상정도	빈도	변수
1	1	1	10	
2	1	2	17	
3	1	3	7	
4	2	1	23	
5	2	2	12	
6	2	3	5	
7				
8				
9				

그림 7: SPSS에서의 분할표 분석을 위한 자료의 형태

주어야 한다. 가중치 지정은 데이터 → 가중 케이스 순서로 메뉴를 선택하면 그림 8와 같은 대화창이 만들어지고 이 대화창에 ‘가중 케이스

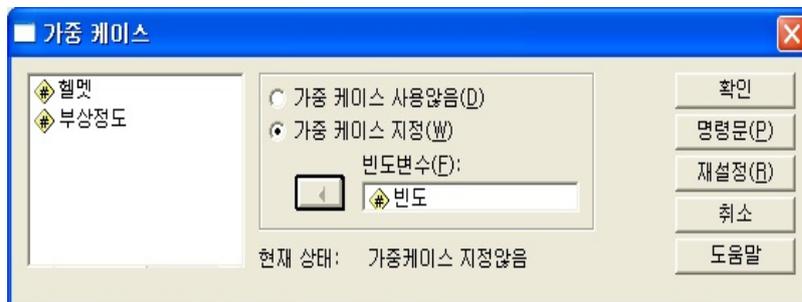


그림 8: SPSS에서의 가중치 주기

지정’을 을 선택하면 그림에서와 같이 가중치를 지정할 수 있게 된다.



그림 9: 교차표에 관련된 통계량 출력 지정

그림 6의 **통계량** 버튼을 클릭하면 독립성 검정에 필요한 각종 통계량을 지정할 수 있는 대화창이 그림 9가 생성되는데 이 그림의 **A**에서 카이제곱을 선택하면 앞에서 설명된 카이제곱 통계량은 $\chi^2 = 5.8687$ 이다. 계산한 유의확률 p 값은 0.0532이므로 유의수준 5%에서 귀무가설 헬멧착용여부와 부상정도는 독립이라는 귀무가설을 기각할 수 없다. (False sense of security 참조)