



실험계획과 분석

심송용(한림대학교 데이터과학스쿨)

<http://jupiter.hallym.ac.kr>

a개 그룹 비교-일원배치 ANOVA

보기: 교육프로그램에 따른 판매실적

어느 회사에서 교육프로그램에 따라 영업실적이 차이가 있는지 알아보기 위해 20명의 신입사원을 대상으로 서로 다른 네 가지 판매전략 교육프로그램에 임의로 배치한 후 이들의 영업실적을 조사하여 다음 결과를 얻었다.

프로그램 1	프로그램 2	프로그램 3	프로그램 4
74	94	62	80
67	82	75	82
83	69	59	75
77	78	79	90
71	68	68	72

$$\sum \sum y_{ij} = 1505, N = na = 20 \text{이므로 } CT = \frac{y_{..}^2}{N} = \frac{1505^2}{20} = 113251.2 \text{이며(반올림)}$$

$$\sum \sum y_{ij}^2 = 114701, y_{i.} \text{은 순서대로 } 372, 391, 343, 399 \text{이므로}$$

$$\sum \frac{y_{i.}^2}{n_i} = \frac{1}{5}(372^2 + 391^2 + 343^2 + 399^2) = \frac{568115}{5} = 113623$$

a개 그룹 비교-일원배치 ANOVA

이다. 따라서

$$SST = \sum \sum y_{ij}^2 - CT = 114701 - 113251.2 = 1449.75$$

$$SSTrt = \sum \frac{y_{i.}^2}{n_i} - CT = 113623 - 113251.2 = 371.75$$

$SSE = SST - SSTrt = 1078$ 을 얻을 수 있다. 따라서 분산분석표는

요인	제곱합	자유도	평균제곱	F	유의확률
처리	371.75	3	123.916	1.839	0.181
오차	1078	16	67.375		
전체	1449.75	19			

로 얻을 수 있다.

검정통계량 $F_0 = 1.839 < F_{3,16,0.05} = 3.24$ 이므로 귀무가설 $H_0: \tau_1 = \tau_2 = \dots = \tau_4 = 0$ 를 기각하지 못한다. 즉, 교육프로그램에 따른 차이가 없다고 결론을 얻는다.

위의 유의확률을 보면 0.181로 유의수준 0.05보다 작지 않으므로 같은 결론을 얻는다.

a개 그룹 비교-일원배치 ANOVA

참고로 위의 분산분석표의 유의확률은 R에서

> 1-pf(1.839, 3,16)

[1] 0.1807352

로 계산하였다.

- 첫 번째 교육프로그램에 따른 평균 μ_1 의 95% 신뢰구간은 $\bar{y}_1 = 74.4$ 이고 $t_{16;0.025} = 2.12$ 이므로

$$\mu_1 = \bar{y}_1 \pm t_{N-\alpha;\alpha/2} \frac{\sqrt{MSE}}{\sqrt{n_1}} = 74.4 \pm 2.12 \frac{\sqrt{67.375}}{\sqrt{5}} = (66.62, 82.18) \text{ 이다.}$$

- $H_0 : \mu_1 = 75.25$ 대 $H_1 : \mu_1 < 75.25$ 을 검정해보면(75.25는 전체 평균)

$$t_0 = \frac{\bar{y}_1 - \mu_{i0}}{\sqrt{MSE} / \sqrt{n_1}} = \frac{74.4 - 75.25}{\sqrt{67.375} / \sqrt{5}} = -0.2315 \text{ 가 } -t_{16;0.05} = -1.745884 \text{보다 작지 않으므로}$$

귀무가설을 기각하지 못한다. 즉, 첫 번째 교육프로그램의 평균이 75.25보다 작다고 할 수 없다.

a개 그룹 비교-일원배치 ANOVA

- 첫 번째 교육프로그램과 두 번째 교육프로그램의 평균차에 대한 95% 신뢰구간을 구해보면(그룹별 평균은 순서대로 74.4, 78.2, 68.6, 79.8)

$$\begin{aligned}\mu_1 - \mu_2 &= (\bar{y}_1 - \bar{y}_2) \pm t_{N-\alpha/2} \sqrt{MSE} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = (74.4 - 78.2) \pm 2.12 \sqrt{67.375} \sqrt{\frac{1}{5} + \frac{1}{5}} \\ &= (-14.80564, 7.205639)\end{aligned}$$

이다.

- 첫 번째 교육프로그램과 두 번째 교육프로그램의 평균차이가 있는지 유의수준 5%에서 검정해보자. $H_0 : \mu_1 - \mu_2 = 0$, $H_1 : \mu_1 - \mu_2 \neq 0$

검정통계량이

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{MSE} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{74.4 - 78.2}{\sqrt{67.375} \sqrt{\frac{1}{5} + \frac{1}{5}}} = -0.7319884 \text{ 이고}$$

$|t_0| > t_{16;0.025} = 2.12$ 이지 않으므로 귀무가설을 기각하지 못한다.