



실험계획과 분석

심송용(한림대학교 데이터과학스쿨)

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

a개 그룹 비교-일원배치 ANOVA

SAS를 사용한 분석

```
data a; /* oneway1.sas */  
  input program sales @@;  
  cards;  
1 74 1 67 1 83 1 77 1 71  
2 94 2 82 2 69 2 78 2 68  
3 62 3 75 3 59 3 79 3 68  
4 80 4 82 4 75 4 90 4 72  
;  
proc anova data = a;  
  class program;  
  model sales = program;  
run;
```

- PROC ANOVA: 분산분석에서 사용하는 SAS 프로시저 중의 하나. 주로 균형설계(일반적

a개 그룹 비교-일원배치 ANOVA

으로 모든 수준에서의 반복수가 동일)인 경우에 사용

- CLASS: 처리(그룹)을 표현하는데 사용된 변수를 설정
- MODEL: 반응변수 = 처리변수(들) 형태로 반응변수와 설명변수를 설정

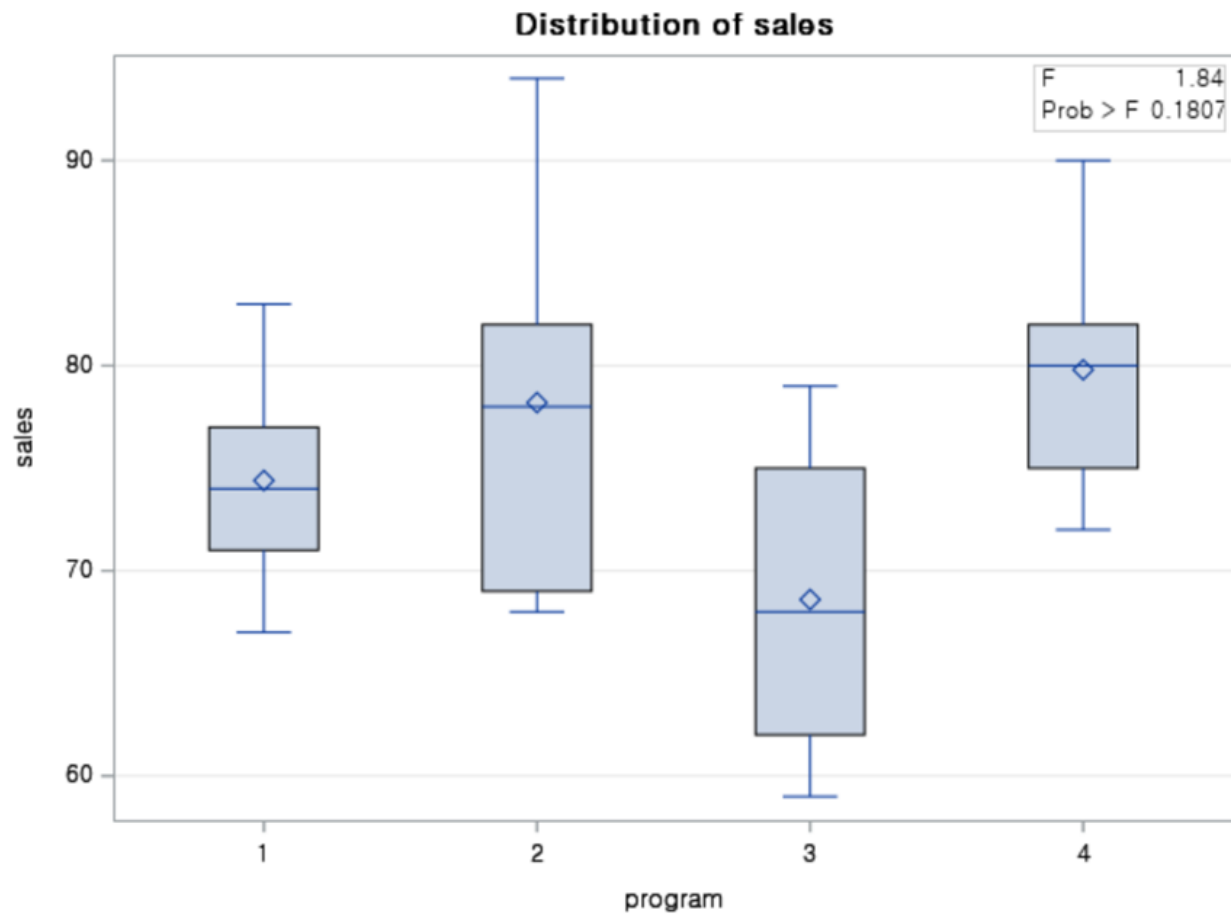
Dependent Variable: sales

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	371.750000	123.916667	1.84	0.1807
Error	16	1078.000000	67.375000		
Corrected Total	19	1449.750000			

R-Square	Coeff Var	Root MSE	sales Mean
0.256424	10.90794	8.208228	75.25000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
program	3	371.7500000	123.9166667	1.84	0.1807

a개 그룹 비교-일원배치 ANOVA



a개 그룹 비교-일원배치 ANOVA

```
proc anova data = a;  
  class program;  
  model sales = program;  
  means program;  
run;
```

- MEANS: 에 주어진 설명변수의 값에 따른 처리수준별 평균, 표준편차 계산 및 다중비교

Level of program	N	sales	
		Mean	Std Dev
1	5	74,4000000	6,0663004
2	5	78,2000000	10,6395489
3	5	68,6000000	8,4439327
4	5	79,8000000	6,9426220

a개 그룹 비교-일원배치 ANOVA

GLM 프로시저: ANOVA보다 범용으로 사용

DATA 정의는 oneway1.sas와 같으며 CLASS 문, MODEL, MEANS문의 기본사용법은 같음

```
/* oneway2.sas */
```

```
proc glm data = a;
```

```
  class program;
```

```
  model sales = program / solution clparm;
```

```
  means program;
```

```
run;
```

MODEL 문의 옵션

- solution: τ_i 들의 추정치 계산
- clparm: 신뢰구간 계산(기준범주와 차이에 대한 신뢰구간임에 유의)

출력에서 분산분석표, 평균/표준편차 계산, 상자그림 등은 PROC ANOVA와 같은 값이므로 생략함.

a개 그룹 비교-일원배치 ANOVA

Parameter	Estimate		Standard Error	t Value	Pr > t	95% Confidence Limits	
Intercept	79.80000000	B	3.67083097	21.74	<.0001	72.01818598	87.58181402
program 1	-5.40000000	B	5.19133894	-1.04	0.3137	-16.40514693	5.60514693
program 2	-1.60000000	B	5.19133894	-0.31	0.7619	-12.60514693	9.40514693
program 3	-11.20000000	B	5.19133894	-2.16	0.0465	-22.20514693	-0.19485307
program 4	0.00000000	B

빨간 박스: solution 옵션에 의한 결과

파란 박스: clparm 옵션에 의한 결과

- B(회귀계수)의 값은 intercept에 각 program에서의 추정치를 더해주면 됨.
즉, $\bar{y}_{1.} = 79.8 - 5.4 = 74.4$, $\bar{y}_{4.} = 79.8 + 0 = 79.8$ 등임.
- 표준오차는 intercept를 제외하면 모두 5.191인데 이 값은 기준범주(이 경우 program 4)와 차이에 대한 표준오차를 계산한 것으로 이 값은

a개 그룹 비교-일원배치 ANOVA

$$SE(\bar{y}_i - \bar{y}_4) = \sqrt{MSE} \sqrt{\frac{1}{n_i} + \frac{1}{n_4}} = \sqrt{67.375} \sqrt{\frac{1}{5} + \frac{1}{5}} = 5.191339 \text{ 임.}$$

따라서 차이에 대한 신뢰구간은

$$\begin{aligned} \mu_i - \mu_4 &= \bar{y}_i - \bar{y}_4 \pm t_{N-\alpha;\alpha} \sqrt{MSE} \sqrt{\frac{1}{n_i} + \frac{1}{n_4}} = (74.4 - 79.8) \pm 2.12 \times 5.191339 \\ &= (-16.403, 5.603) \quad (i = 1 \text{ 일 때, 반올림 오차 포함}) \end{aligned}$$

- intercept에 대한 표준오차는 기준범주 program 4의 평균 μ_i 에 대한 것으로

$$SE(\bar{y}_4) = \sqrt{MSE} / \sqrt{n_4} = \sqrt{67.375/5} = 3.670831 \text{ 이며}$$

$$\text{신뢰구간은 } \mu_4 = \bar{y}_4 \pm t_{N-\alpha;\alpha/2} SE(\bar{y}_4) = 79.8 \pm 2.12 \times 3.670831 = (72.01784, 87.58216) \text{ 임.}$$

a개 그룹 비교-일원배치 ANOVA

보기: 독립 이표본 t -검정 재검토

```
data score; /* oneway3.sas */
```

```
input gender$ score;
```

```
cards;
```

데이터 생략. ttest1.sas 와 같음

```
;
```

```
proc glm data=score;
```

```
class gender;
```

```
model score = gender;
```

```
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1960.000000	1960.000000	5.75	0.0433
Error	8	2726.000000	340.750000		
Corrected Total	9	4686.000000			

a개 그룹 비교-일원배치 ANOVA

R-Square	Coeff Var	Root MSE	score Mean
0.418267	5.768567	18.45941	320.0000

t -검정의 출력(ttest1.sas)

gender	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
F		334.0	311.7	356.3	17.9304	10.7427	51.5241
M		306.0	282.4	329.6	18.9737	11.3678	54.5219
Diff (1-2)	Pooled	28.0000	1.0780	54.9220	18.4594	12.4685	35.3640
Diff (1-2)	Satterthwaite	28.0000	1.0630	54.9370			

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	8	2.40	0.0433
Satterthwaite	Unequal	7.9745	2.40	0.0434

유의확률이 같음(0.0433)

$$t^2 = F \quad (2.40^2 = 5.75)$$

$$S_p^2 = MSE \quad (\text{출력에서는 제공된 값을 비교: } S_p = \sqrt{MSE} = 18.4594)$$