

單因子變異數分析例題

我們想了解學生對不同教學法喜愛程度是否有差異，A組以PPT教學，B組直接黑板演練，C組使用線上教學，隨機抽取三種教學法學生各5個人，以1－10的分數請他們評分如下：

A組	分數	B組	分數	C組	分數
A1	8	B1	8	C1	4
A2	4	B2	5	C2	4
A3	5	B3	9	C3	6
A4	5	B4	7	C4	6
A5	4	B5	9	C5	5
平均=5.2		平均=7.6		平均=5	
總平均=5.9					

組間: $df = 2$

$$5 \cdot (5.2 - 5.9)^2 + 5 \cdot (7.6 - 5.9)^2 + 5 \cdot (5 - 5.9)^2 \\ = 5 \cdot (0.7)^2 + 5 \cdot (1.7)^2 + 5 \cdot (0.9)^2 = 20.95$$

組內: $df = 12$

$$(8 - 5.2)^2 + (4 - 5.2)^2 + (5 - 5.2)^2 + (5 - 5.2)^2 + (4 - 5.2)^2 \\ + (8 - 7.6)^2 + (5 - 7.6)^2 + (9 - 7.6)^2 + (7 - 7.6)^2 + (9 - 7.6)^2 \\ + (4 - 5)^2 + (4 - 5)^2 + (6 - 5)^2 + (6 - 5)^2 + (5 - 5)^2 = 26$$

總: $20.95 + 26 = 46.95$ $df = 14$

ANOVA分析表

來源	平方和	自由度	均方和	F值
組間	20.95	2	10.475	4.83
組內	26	12	2.167	
總和	46.95	14		

組間: $df = 2$ 組內: $df = 12$ 總: $df = 14$
 $= 20.95$ $= 26$ $= 46.95$

	1	2	3	4	5	6	7	8	9
1	161.4	199	216	225	230	234	236.8	239	241
2	18.51	19	19.2	19.2	19.3	19.3	19.35	19.4	19.4
3	10.13	9.55	9.28	9.12	9.01	8.94	8.887	8.85	8.81
4	7.709	6.94	6.59	6.39	6.26	6.16	6.094	6.04	6
5	6.608	5.79	5.41	5.19	5.05	4.95	4.876	4.82	4.77
6	5.987	5.14	4.76	4.53	4.39	4.28	4.207	4.15	4.1
7	5.591	4.74	4.35	4.12	3.97	3.87	3.787	3.73	3.68
8	5.318	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39
9	5.117	4.26	3.86	3.63	3.48	3.37	3.293	3.23	3.18
10	4.965	4.1	3.71	3.48	3.33	3.22	3.135	3.07	3.02
11	4.844	3.98	3.59	3.36	3.2	3.09	3.012	2.95	2.9
12	4.747	3.89	3.49	3.26	3.11	3	2.913	2.85	2.8
13	4.667	3.81	3.41	3.18	3.03	2.92	2.832	2.77	2.71

- $SS_{TOT} = \sum_{i=1}^n (Y_i - \bar{Y})^2 = (\mathbf{Y} - \bar{\mathbf{Y}})'(\mathbf{Y} - \bar{\mathbf{Y}})$, where

$$\bar{\mathbf{Y}} = \bar{Y} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

- $SS_{REG} = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 = (\hat{\mathbf{Y}} - \bar{\mathbf{Y}})'(\hat{\mathbf{Y}} - \bar{\mathbf{Y}})$

- $SS_{ERR} = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = (\mathbf{Y} - \hat{\mathbf{Y}})'(\mathbf{Y} - \hat{\mathbf{Y}})$

Source	SS	d.f.	MS	F
Model	SS_{REG}	k	$MS_{REG} = \frac{SS_{REG}}{k}$	$\frac{MS_{REG}}{MS_{ERR}}$
Error	SS_{ERR}	$n - k - 1$	$MS_{ERR} = \frac{SS_{ERR}}{n - k - 1}$	
Total	SS_{TOT}	$n - 1$		

```

```{r}
x1 = c(8,4,5,5,4)
x2 = c(8,5,9,7,9)
x3 = c(4,4,6,6,5)
z = c(x1,x2,x3)
a = factor(rep(1:3,each = 5))
print(z)
print(a)
```

```

```

[1] 8 4 5 5 4 8 5 9 7 9 4 4 6 6 5
[1] 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3
Levels: 1 2 3

```

```

```{r}
summary(aov(z~a))
```

```

```

          Df Sum Sq Mean Sq F value Pr(>F)
a           2   20.93   10.467    4.831 0.0289 *
Residuals   12   26.00    2.167
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

$$\bar{Y} = \text{mean}(c(x1,x2,x3)) = 5.9$$

$$X1: \hat{Y}_1 = \text{mean}(X1) = 5.2$$

$$SS_{reg} = \text{sum}((\hat{Y} - \bar{Y})^2) = 5 \cdot (5.2 - 5.9)^2 + 5 \cdot (7.6 - 5.9)^2 + 5 \cdot (5 - 5.9)^2 = 5 \cdot (0.7)^2 + 5 \cdot (1.7)^2 + 5 \cdot (0.9)^2 = 20.95$$

$$SS_{err} = \text{sum}((Y - \hat{Y})^2) = (8 - 5.2)^2 + (4 - 5.2)^2 + (5 - 5.2)^2 + (5 - 5.2)^2 + (4 - 5.2)^2 + (8 - 7.6)^2 + (5 - 7.6)^2 + (9 - 7.6)^2 + (7 - 7.6)^2 + (9 - 7.6)^2 + (4 - 5)^2 + (4 - 5)^2 + (6 - 5)^2 + (6 - 5)^2 + (5 - 5)^2 = 26$$