第四章 方差分析

Variance analysis

- 在科学试验和工农业生产中，常常需要分析哪几种因素对产品的产量或质量有显著性影响，并希望知道这些因素处于什么条件时生产状态最佳。
- 为了解决上述问题，一般需要做两方面的工作:
- 一是设计一个试验使其充分反映我们所感兴趣的因素的作用，并使试验次数尽可能的少，以节省人力和物力；
- 二是如何分析、利用试验结果对我们所关心的问题（如因素作用的大小）作出合理的推断。
- 前者称为试验设计，后者称为方差分析。

第四章 方差分析

4.1 单因素方差分析

4.2 多因素方差分析

§4.1 单因素方差分析

Variance analysis of Single factor

问 题 的 提 出

建 立 模 型

参 数 估 计

统 计 检 验

§4.1 单因素方差分析

引例. 某农科所为了比较四种不同的肥料对农作物产量的影响，进行了下面的试验，他们选用一块肥沃程度比较均匀的土地，并将这块土地按表 4-1 的方式等分成十六小块。同时为了尽可能减少土地原有肥沃程度的影响，采用表 4-1 的方式安排试验。

$A_i(i=1,2,3,4)$表示在这块小土地上施第i种肥料,显然肥料均用在四块小土地上。

表 4-1

A_1	A_2	A_3	A_4
A_2	A_3	A_4	A_1
A_3	A_4	A_1	A_2
A_4	A_1	A_2	A_3

§4.1 单因素方差分析

试验结果——作物产量由表 4-2 给出，我们希望通过表 4-2 给出的数据来推断肥料对该作物产量是否有显著影响？如果影响显著，那么施用哪一种肥料最好？

表 4-2

肥料种类 A_i	收获量 X_{ij}				平均产量 $\overline{X}_i$
A_1	98	96	91	66	87.75
A_2	60	69	50	35	53.50
A_3	79	64	81	70	73.50
A_4	90	70	79	88	81.75

§4.1 单因素方差分析

在统计上，我们把所考虑的因素（如肥料）称为因子，因子所处的状态称为水平。在引例中肥料为一个因子，它有四个水平A_1、A_2、A_3、A_4。施用不同种类的肥料所得的产量不能视为同一个总体中产生的子样，而应看作是分别从四个总体X_1，X_2，X_3，X_4中抽得长度为4的样本。

通常我们假定总体服从正态分布，即

$$X_i \sim N(\mu_i, \sigma^2)\ (i = 1,2,3,4)$$

我们想知道四种不同的肥料对农作物产量的影响有无差异，此问题可归结为下述检验：

H_0：$\mu_1 = \mu_2 = \mu_3 = \mu_4$。

§4.1 单因素方差分析

设因子A有k个水平$A_1, A_2, \cdots, A_k$，在水平A_i的条件下，做了r次试验（只讨论每个水平所作试验次数相同的情形），试验结果为$y_{i1}, y_{i2}, \cdots, y_{ir}$
$(i=1,2,3,\cdots, k)$，所有试验结果列入表示 4-3 中

表 4-3

结果 次数 因子水平	1	2	……	r
A_1	y_{11}	y_{12}	……	y_{1r}
A_2	y_{21}	y_{22}	……	y_{2r}
⋮	⋮	⋮	……	⋮
A_k	y_{k1}	y_{k2}	……	y_{kr}

§4.1 单因素方差分析

一. 建立模型

modeling

我们可将 y_{ij} 分解为

$$y_{ij} = \mu_i + \varepsilon_{ij} \left(i = 1,2,3,\cdots,k，\ j = 1,2,\cdots,r\right) \quad （4\text{–}1）$$

其中 μ_i 为 y_{ij} 的均值, 即 $Ey_{ij} = \mu_i$， ε_{ij} 为试验误差, 它是一个 $r.v$，通常假定 $\varepsilon_{ij}\left(i = 1,2,3,\cdots,k，\ j = 1,2,\cdots，r\right)$

独立同分布 $N（0，\sigma^2）$。

令 $\mu = \dfrac{1}{k}\sum_{i=1}^{k}\mu_i$ ，$\alpha_i = \mu_i - \mu \left(i = 1,2,\cdots,k\right)$ （4–2）

显然，$\sum_{i=1}^{k}\alpha_i = 0$，于是

§ 4.1 单因素方差分析

$$y_{ij} = \mu \ + \alpha_i + \varepsilon_{ij} \quad (i = 1,2,3,\cdots,k\text{，}\quad j = 1,2,\cdots,r)$$

α_i 反映了水平 A_i 对试验结果的影响，称作水平 A_i 的主效应。由前所述得方差分析的一般数学模型：

$$\begin{cases} y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad i = 1,2,\cdots,k, \quad j = 1,2,\cdots,r \\ \sum\limits_{i=1}^{k} \alpha_i = 0 \\ \varepsilon_{ij} \text{独立同分布} N(0,\sigma^2) \end{cases} \qquad (4\text{–}3)$$

μ, α_i $(i = 1,2,3,\cdots,k)$ 是未知参数.要解决的问题是：

（1）估计未知参数 μ，α_i （$i = 1,2,3,\cdots$，k）；

（2）考察 k 个因子水平对试验结果的影响有无显著差异。即检验 $H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$。

二. 参数估计 Parameter estimate

记 $\bar{y} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} y_{ij}$， $\bar{y}_i = \frac{1}{r}\sum_{j=1}^{r} y_{ij}$，由于

$$E\bar{y} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r} Ey_{ij} = \frac{1}{kr}\sum_{i=1}^{k}\sum_{j=1}^{r}(\mu+\alpha_i) = \mu$$

$$E(\bar{y}_i - \bar{y}) = E\bar{y}_i - E\bar{y} = \frac{1}{r}\sum_{j=1}^{r} Ey_{ij} - \mu$$

$$= \frac{1}{r}\sum_{j=1}^{r}(\mu+\alpha_i) - \mu = \alpha_i$$

因此μ， α_i的矩估计为

$$\hat{\mu} = \bar{y}，\quad \hat{\alpha}_i = \bar{y}_i - \bar{y}，\quad i = 1,2$$

显然它们分别是μ， α_i的无

一. 建立模型 modeling

我们可将y_{ij}分解为

$$y_{ij} = \mu_i + \varepsilon_{ij}\ (i = 1,2,3,\cdots,k，\ j = 1,2,\cdots,r) \qquad (4\text{-}1)$$

其中μ_i为y_{ij}的均值，即$Ey_{ij}=\mu_i$， ε_{ij}为试验误差，它是一个$r.v$，通常假定$\varepsilon_{ij}\ (i = 1,2,3,\cdots,k，\ j = 1,2,\cdots，r)$独立同分布$N$（0， σ^2）。

令 $\mu = \frac{1}{k}\sum_{i=1}^{k}\mu_i$ ，$\alpha_i = \mu_i - \mu\ (i = 1,2,\cdots,k)$ （4-2）

显然，$\sum_{i=1}^{k}\alpha_i = 0$，于是

三. 统计检验 Statistical tests

记全部数据的总变差为

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y})^2=\sum_{i=1}^{k}\sum_{j=1}^{r}y_{ij}^2-\frac{T_{..}^2}{kr} \qquad （4–5）$$

其中 $T_{..}=\sum_{i=1}^{k}\sum_{j=1}^{r}y_{ij}$ ，对 $S_{总}^2$ 进行平方和分解得

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y})^2=\sum_{i=1}^{k}\sum_{j=1}^{r}[(y_{ij}-\bar{y}_i)+(\bar{y}_i-\bar{y})]^2$$

$$=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)^2+\sum_{i=1}^{k}\sum_{j=1}^{r}(\bar{y}_i-\bar{y})^2$$

$$+\boxed{2\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)(\bar{y}_i-\bar{y})}$$

=0

三. 统计检验

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y})^2 =\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)^2+\sum_{i=1}^{k}\sum_{j=1}^{r}(\bar{y}_i-\bar{y})^2$$

$$=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)^2+r\sum_{i=1}^{k}(\bar{y}_i-\bar{y})^2 \triangleq S_{误}^2+S_{组间}^2$$

其中

$$S_{组间}^2=r\sum_{i=1}^{k}(\bar{y}_i-\bar{y})^2=\sum_{i=1}^{k}\frac{T_{i\cdot}^2}{r}-\frac{T_{\cdot\cdot}^2}{rk} \quad \left(T_{i\cdot}=\sum_{j=1}^{r}y_{ij}\right) \quad (4\text{–}6)$$

$$S_{误}^2=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)^2=S_{总}^2-S_{组间}^2 \quad (4\text{–}7)$$

$$\boxed{\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)(\bar{y}_i-\bar{y})=0}$$

三. 统计检验

$$\boxed{\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)(\bar{y}_i-\bar{y})}=\sum_{i=1}^{k}\boxed{(\bar{y}_i-\bar{y})\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)}$$

$$=\sum_{i=1}^{k}(\bar{y}_i-\bar{y})\boxed{(\sum_{j=1}^{r}y_{ij}-r\bar{y}_i)}=\sum_{i=1}^{k}(\bar{y}_i-\bar{y})(r\bar{y}_i-r\bar{y}_i)=0$$

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{r}(y_{ij}-\bar{y}_i)^2+r\sum_{i=1}^{k}(\bar{y}_i-\bar{y})^2\triangleq S_{误}^2+S_{组间}^2$$

记$n=kr$，可证当H_0成立时

$$F=\frac{S_{组间}^2/(k-1)}{S_{误}^2/(n-k)}\sim F(k-1,\ \ n-k) \qquad （4\text{–}8）$$

因此可利用F检验来检验$H_0:\alpha_1=\alpha_2=\cdots=\alpha_k=0$，

三. 统计检验

对给定水平α，由$P\{F>\lambda\}=\alpha$查$F(k-1, n-k)$表得λ。若$F>\lambda$，则拒绝H_0，即检验效果显著；否则接受H_0，即检验效果不显著.通常将有关结果列为一张表，称为方差分析表，如表 4-4

表 4-4 方差分析表

方差来源	平方和	自由度	平均平方和	F 值
因素 A(组间)	$S^2_{组间}$	$k-1$	$\dfrac{S^2_{组间}}{k-1}$	$F=\dfrac{S^2_{组间}/(k-1)}{S^2_{误}/(n-k)}$
误差（组内）	$S^2_{误}$	$n-k$	$\dfrac{S^2_{误}}{n-k}$	
总　　和	$S^2_{总}$	$n-1$		

例 4-1 利用表 4-2 对引例进行方差分析。

表 4-2

肥料种类 A_i	收 获 量 X_{ij}				平均产量 $\overline{X}_i$
A_1	98	96	91	66	87.75
A_2	60	69	50	35	53.50
A_3	79	64	81	70	73.50
A_4	90	70	79	88	81.75

解： 由于 $k=4$，$r=4$，$n=kr=16$，经计算得

$$T_{1\cdot}=\sum_{j=1}^{r} y_{1j}=351 \qquad T_{2\cdot}=\sum_{j=1}^{r} y_{2j}=214$$

$$T_{3\cdot}=\sum_{j=1}^{r} y_{3j}=294 \qquad T_{4\cdot}=\sum_{j=1}^{r} y_{4j}=327$$

$$T_{\cdot\cdot} = \sum_{i=1}^{k}\sum_{j=1}^{r} y_{ij} = 1186$$

$$S_{总}^2 = \sum_{i=1}^{k}\sum_{j=1}^{r} y_{ij}^2 - \frac{T_{\cdot\cdot}^2}{kr} = 4413.75$$

$$S_{组间}^2 = \sum_{i=1}^{k} \frac{T_{i\cdot}^2}{r} - \frac{T_{\cdot\cdot}^2}{rk} = 2678.25$$

$$S_{误}^2 = \sum_{i=1}^{k}\sum_{j=1}^{r} (y_{ij} - \bar{y}_i)^2 = S_{总}^2 - S_{组间}^2 = 1735.5$$

$$F = \frac{S_{组间}^2/(k-1)}{S_{误}^2/(n-k)} = \frac{2678.25/3}{1735.5/12} \approx 6.173$$

方差来源	平方
因素 A(组间)	267
误差(组内)	173
总　　和	441

对给定水平

$F(3,12)$分布表得

H_0，即认为肥

若检验效果

方打“*”，如果

右上方打“**”

α=0.10

Fα k1 / k2	1	2	3	4	5	6	8	12	24	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	59.44	60.71	62.00	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.37	9.41	9.45	9.49
3	5.54	5.46	5.36	5.32	5.31	5.28	5.25	5.22	5.18	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.95	3.90	3.83	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.34	3.27	3.19	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	2.98	2.90	2.82	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.75	2.67	2.58	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.59	2.50	2.40	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.47	2.38	2.28	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.38	2.28	2.18	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.30	2.21	2.10	1.97
12	3.18	2.81	$\boxed{2.61}$	2.48	2.39	2.33	2.24	2.15	2.04	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.20	2.10	1.98	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.15	2.05	1.94	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.12	2.02	1.90	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.09	1.99	1.87	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.06	1.96	1.84	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.04	1.93	1.81	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.02	1.91	1.79	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.00	1.89	1.77	1.61

例 4-1 利用表 4-2 对引例进行方差分析。

表 4-2

肥料种类 A_i	收 获 量 X_{ij}				平均产量 $\overline{X}_i$
A_1	98	96	91	66	87.75
A_2	60	69	50	35	53.50
A_3	79	64	81	70	73.50
A_4	90	70	79	88	81.75

解： 由于 $k=4$，$r=4$，$n=kr=16$，经计算得

$$T_{1\cdot}=\sum_{j=1}^{r}y_{1j}=351 \qquad T_{2\cdot}=\sum_{j=1}^{r}y_{2j}=214$$

$$T_{3\cdot}=\sum_{j=1}^{r}y_{3j}=294 \qquad T_{4\cdot}=\sum_{j=1}^{r}y_{4j}=327$$

$$T_{..} = \sum_{i=1}^{k}\sum_{j=1}^{r} y_{ij} = 1186$$

$$S_{总}^2 = \sum_{i=1}^{k}\sum_{j=1}^{r} y_{ij}^2 - \frac{T_{..}^2}{kr} = 4413.75$$

$$S_{组间}^2 = \sum_{i=1}^{k} \frac{T_{i.}^2}{r} - \frac{T_{..}^2}{rk} = 2678.25$$

$$S_{误}^2 = \sum_{i=1}^{k}\sum_{j=1}^{r} (y_{ij} - \bar{y}_i)^2 = S_{总}^2 - S_{组间}^2 = 1735.5$$

$$F = \frac{S_{组间}^2/(k-1)}{S_{误}^2/(n-k)} = \frac{2678.25 / 3}{1735.5 / 12} \approx 6.173$$

方差来源	平方
因素A(组间)	267
误差(组内)	173
总　　和	441

对给定水平

$F(3,12)$分布表

H_0，即认为肥

若检验效果

方打“*”，如果

右上方打“**”

α=0.10

Fα k1 / k2	1	2	3	4	5	6	8	12	24	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	59.44	60.71	62.00	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.37	9.41	9.45	9.49
3	5.54	5.46	5.36	5.32	5.31	5.28	5.25	5.22	5.18	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.95	3.90	3.83	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.34	3.27	3.19	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	2.98	2.90	2.82	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.75	2.67	2.58	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.59	2.50	2.40	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.47	2.38	2.28	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.38	2.28	2.18	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.30	2.21	2.10	1.97
12	3.18	2.81	$\boxed{2.61}$	2.48	2.39	2.33	2.24	2.15	2.04	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.20	2.10	1.98	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.15	2.05	1.94	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.12	2.02	1.90	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.09	1.99	1.87	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.06	1.96	1.84	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.04	1.93	1.81	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.02	1.91	1.79	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.00	1.89	1.77	1.61

练习： 某厂有三条生产线，从三条生产线生产的纤维中分别抽取了一些样品，测得纤维强度数据见表 4–6，试考察它们生产的纤维在强度上是否有显著差异？

表 4-6

自动生产线	纤维强度					
甲	7.0	7.4	6.1	6.5	7.5	
乙	5.5	6.7	7.2	5.8		
丙	6.7	7.2	8.2	7.3	7.5	6.9

分析：由于从三条生产线中抽取的样品个数不同，即相当不同水平所做重复试验次数 r 不相等，因此需对例 7–1 中使用的公式进行适当的修正。

解：令 $r_1=5, r_2=4, r_3=6$，由表 4–6 的数据可以算出：

$$T_{..}=\sum_{i=1}^{k}\sum_{j=1}^{r_i} y_{ij}=103.5$$

$$S_{总}^2=\sum_{i=1}^{k}\sum_{j=1}^{r_i} y_{ij}^2-\frac{T_{..}^2}{\sum_{i=1}^{k} r_i}$$

$$=\sum_{i=1}^{3}\sum_{j=1}^{r_i} y_{ij}^2-\frac{T_{..}^2}{5+4+6}=721.21-\frac{1}{15}\times 103.5^2=7.06$$

$$T_{1.}=\sum_{j=1}^{5} y_{1j}=34.5 \quad T_{2.}=\sum_{j=1}^{4} y_{2j}=25.2 \quad T_{3.}=\sum_{j=1}^{6} y_{3j}=43.8$$

$$S^2_{组间}=\sum_{i=1}^{k}\frac{T_{i\cdot}^2}{r_i}-\frac{T_{\cdot\cdot}^2}{\sum r_i}$$

$$=\frac{34.5^2}{5}+\frac{25.2^2}{4}+\frac{43.8^2}{6}-\frac{103.5^2}{15}=2.4$$

$$S^2_{误}=\sum_{i=1}^{k}\sum_{j=1}^{r_i}(y_{ij}-\bar{y}_i)^2=S^2_{总}-S^2_{组间}=7.06—2.4=4.66$$

$$F=\frac{S^2_{组间}/(k-1)}{S^2_{误}/(n-k)}=\frac{2.4/2}{4.66/12}\approx 3.09$$

将上述计算结果列入表 4-7，得

表 4-7 方差分析表

方差来源	平方和	自由度	平均平方和	F 值
因素（组间）	2.4	3—1	1.2	3.09
误差（组内）	4.66	15-3	0.3883	
总 和	7.06	15—1		

对给定水平 α =0.05，由 $P\{F>\lambda\}=0.05$ 查 $F(2，12)$ 分布表得 λ =3.89。显然 $F<\lambda$，所以不能认为三条生产线生产的纤维在强度上有显著差异。

4. Matlab方差分析

- 在Matlab中，anova1函数用于单因素方差分析。
- 原假设：为样本观测矩阵的各列所对应的总体具有相同的均值。
- 样本观测矩阵的列数表示因素A的水平数。
- 假设单因素方差分析条件满足：
- 样本服从正态总体；正态总体的方差相同；样本相互独立。

4. Matlab方差分析

anova1函数的调用格式为

p=anova1(X，group)

[p,table]=anova1(X，group)

[p,table,stats]=anova1(X，group)

其中，当X为样本观测矩阵时，列数表示因素A的水平数；这种调用仅对于每个水平下，实验次数相同的单因素方差分析，此时输入参数group可省略。

4. Matlab方差分析

- 输入参数group 是一个列向量，长度与X的长度相同。
- X中具有相同group值的元素为同组元素。
- 输出参数为检验的p值，若p小于显著性水平alpha，则拒绝原假设；否则接受原假设。
- 输出参数table为元胞数组形式的方差分析表。
- 输出参数stats为结构体变量，用于后续的多重比较。

4. Matlab方差分析

利用 Matlab 对引例的表 4-2 进行方差分析。

表 4-2

肥料种类 A_i	收 获 量 X_{ij}				平均产量 $\overline{X}_i$
A_1	98	96	91	66	87.75
A_2	60	69	50	35	53.50
A_3	79	64	81	70	73.50
A_4	90	70	79	88	81.75

4. Matlab方差分析

```
clear

X=[98,96,91,66;60,69,50,35;79,64,81,70;90,70,79,88]';

[p,table]=anova1(X)
```

4. Matlab方差分析

ANOVA 表

来源	SS	df	MS	F	p 值(F)
列	2678.25	3	892.75	6.17	0.0088
误差	1735.5	12	144.625		
合计	4413.75	15			

方差来源	平方和	自由度	平均平方和	F 值
因素 A(组间)	2678.25	3	892.75	6.173*
误差(组内)	1735.50	12	144.625	
总　　和	4413.75	15		

P=0.0088<0.1, 拒绝“原假设: 样本观测矩阵的各列所对应的总体具有相同的均值”；可得：因子影响显著

4. Matlab方差分析

练习： 某厂有三条生产线，从三条生产线生产的纤维中分别抽取了一些样品，测得纤维强度数据见表 4-6，试考察它们生产的纤维在强度上是否有显著差异？

表 4-6

自动生产线	纤 维 强 度					
甲	7.0	7.4	6.1	6.5	7.5	
乙	5.5	6.7	7.2	5.8		
丙	6.7	7.2	8.2	7.3	7.5	6.9

解：令 $r_1 = 5, r_2 = 4, r_3 = 6$，由表 4-6 的数据可以算出：

4. Matlab方差分析

```
X=[7.0,7.4,6.1,6.5,7.5,5.5,6.7,7.2,5.8,6.7,7.2,8.2,7.3,7.5,
6.9]';

group=[1,1,1,1,1,2,2,2,2,3,3,3,3,3,3]';

[p,table]=anova1(X,group)
```

4. Matlab方差分析

ANOVA 表

来源	SS	df	MS	F	p 值(F)
组	2.4	2	1.2	3.09	0.0827
误差	4.66	12	0.38833		
合计	7.06	14			

表 4-7 方差分析表

方差来源	平方和	自由度	平均平方和	F 值
因素（组间）	2.4	3—1	1.2	3.09
误差（组内）	4.66	15-3	0.3883	
总 和	7.06	15—1		

对给定水平 α =0.05，由 $P\{F > \lambda\} = 0.05$ 查 $F(2，12)$ 分

P=0.0827>0.05，接受“原假设：样本观测矩阵的各列所对应的总体具有相同的均值”；可得：因子影响不显著