




试验数据的方差分析

- 
- 方差分析（analysis of variance，简称ANOVA）
 - 检验试验中有关因素对试验结果影响的显著性
 - 试验指标（experimental index）
 - 衡量或考核试验效果的参数
 - 因素（experimental factor）
 - 影响试验指标的条件
 - 可控因素（controllable factor）
 - 水平（level of factor）
 - 因素的不同状态或内容



3.1 单因素试验的方差分析 (one-way analysis of variance)

3.1.1 单因素试验方差分析基本问题

(1) 目的：检验一个因素对试验结果的影响是否显著性

(2) 基本命题：

- 设某单因素A有r种水平： A_1, A_2, \dots, A_r ，在每种水平下的试验结果服从正态分布
- 在各水平下分别做了 n_i ($i=1, 2, \dots, r$) 次试验
- 判断因素A对试验结果是否有显著影响

(3) 单因素试验数据表

试验次数	A_1	A_2	...	A_i	...	A_r
1	x_{11}	x_{21}	...	x_{i1}	...	x_{r1}
2	x_{12}	x_{22}	...	x_{i2}	...	x_{r2}
...
j	x_{1j}	x_{2j}	...	x_{ij}	...	x_{rj}
...
n_i	x_{1n1}	x_{2n2}	...	x_{ini}	...	x_{rnr}



3.1.2 单因素试验方差分析基本步骤

(1) 计算平均值

■ 组内平均值：

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

■ 总平均：

$$\bar{x} = \frac{1}{n} \sum_{i=1}^r \sum_{j=1}^{n_i} x_{ij}$$

(2) 计算离差平方和

①总离差平方和 SS_T (sum of squares for total)


$$SS_T = \sum_{i=1}^r \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

- 表示了各试验值与总平均值的偏差的平方和
- 反映了试验结果之间存在的总差异

②组间离差平方和 SS_A (sum of square for factor A)

$$SS_A = \sum_{i=1}^r \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 = \sum_{i=1}^r n_i (\bar{x}_i - \bar{x})^2$$

- 反映了各组内平均值之间的差异程度
- 由于因素A不同水平的不同作用造成的



③ 组内离差平方和 SS_e (sum of square for error)

$$SS_e = \sum_{i=1}^r \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

- 反映了在各水平内，各试验值之间的差异程度
- 由于随机误差的作用产生

三种离差平方和之间关系：

$$SS_T = SS_A + SS_e$$

(3) 计算自由度 (degree of freedom)

- 总自由度： $df_T = n - 1$
- 组间自由度： $df_A = r - 1$
- 组内自由度： $df_e = n - r$

三者关系： $df_T = df_A + df_e$

(4) 计算平均平方

- 均方 = 离差平方和除以对应的自由度

$$MS_A = SS_A / df_A \quad MS_e = SS_e / df_e$$

MS_A ——组间均方

MS_e ——组内均方/误差的均方

(5) F 检验

$$F_A = \frac{\text{组间均方}}{\text{组内均方}} = \frac{MS_A}{MS_e}$$

- 服从自由度为 (df_A, df_e) 的F分布 (F distribution)
- 对于给定的显著性水平 α , 从F分布表查得临界值 $F_\alpha(df_A, df_e)$
- 如果 $F_A > F_\alpha(df_A, df_e)$, 则认为因素A对试验结果有显著影响否则认为因素A对试验结果没有显著影响

(6) 方差分析表

单因素试验的方差分析表

差异源	SS	df	MS	F	显著性
组间 (因素A)	SS_A	$r-1$	$MS_A = SS_A / (r-1)$	MS_A / MS_e	
组内 (误差)	SS_e	$n-r$	$MS_e = SS_e / (n-r)$		
总和	SS_T	$n-1$			

- 若 $F_A > F_{0.01}(df_A, df_e)$ ，称因素A对试验结果有非常显著的影响，用 “* *” 号表示；
- 若 $F_{0.05}(df_A, df_e) < F_A < F_{0.01}(df_A, df_e)$ ，则因素A对试验结果有显著的影响，用 “*” 号表示；
- 若 $F_A < F_{0.05}(df_A, df_e)$ ，则因素A对试验结果的影响不显著



3.1.3 Excel在单因素试验方差分析中的应用

- 利用Excel“分析工具库”中的“单因素方差分析”工具

3.2 双因素试验的方差分析

- 讨论两个因素对试验结果影响的显著性，又称“二元方差分析”

3.2.1 双因素无重复试验的方差分析

(1) 双因素无重复试验

	B_1	B_2	...	B_s
A_1	x_{11}	x_{12}	...	x_{1s}
A_2	x_{21}	x_{22}	...	x_{2s}
...
A_r	x_{r1}	x_{r2}	...	x_{rs}

(2) 双因素无重复试验方差分析的基本步骤

① 计算平均值

■ 总平均：

$$\bar{x} = \frac{1}{rS} \sum_{i=1}^r \sum_{j=1}^s x_{ij}$$

■ A_i 水平时：

$$\bar{x}_{i\cdot} = \frac{1}{S} \sum_{j=1}^s x_{ij}$$

■ B_j 水平时：

$$\bar{x}_{\cdot j} = \frac{1}{r} \sum_{i=1}^r x_{ij}$$

②计算离差平方和

- 总离差平方和:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^s (x_{ij} - \bar{x})^2 = SS_A + SS_B + SS_e$$

- 因素A引起离差的平方和:

$$SS_A = \sum_{j=1}^s \sum_{i=1}^r (\bar{x}_{i\cdot} - \bar{x})^2 = s \sum_{i=1}^r (\bar{x}_{i\cdot} - \bar{x})^2$$

- 因素B引起离差的平方和:

$$SS_B = \sum_{i=1}^r \sum_{j=1}^s (\bar{x}_{\cdot j} - \bar{x})^2 = r \sum_{j=1}^s (\bar{x}_{\cdot j} - \bar{x})^2$$

- 误差平方和:

$$SS_e = \sum_{i=1}^r \sum_{j=1}^s (x_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{x})^2$$

③计算自由度

- SS_A 的自由度: $df_A = r - 1$
- SS_B 的自由度: $df_B = s - 1$
- SS_e 的自由度: $df_e = (r - 1)(s - 1)$
- SS_T 的自由度: $df_T = n - 1 = rs - 1$
- $df_T = df_A + df_B + df_e$

④计算均方

$$MS_A = \frac{SS_A}{df_A} = \frac{SS_A}{r - 1} \qquad MS_B = \frac{SS_B}{df_B} = \frac{SS_B}{s - 1}$$

$$MS_e = \frac{SS_e}{df_e} = \frac{SS_e}{(r - 1)(s - 1)}$$

⑤ F检验

$$F_A = \frac{MS_A}{MS_e} \qquad F_B = \frac{MS_B}{MS_e}$$

- F_A 服从自由度为 (df_A, df_e) 的F分布;
- F_B 服从自由度为 (df_B, df_e) 的F分布;
- 对于给定的显著性水平 α ，查F分布表:

$$F_\alpha (df_A, df_e), \quad F_\alpha (df_B, df_e)$$

- 若 $F_A > F_\alpha (df_A, df_e)$ ，则因素A对试验结果有显著影响，否则无显著影响;
- 若 $F_B > F_\alpha (df_B, df_e)$ ，则因素B对试验结果有显著影响，否则无显著影响;

⑥无重复试验双因素方差分析表

无重复试验双因素方差分析表

差异源	SS	df	MS	F	显著性
因素A	SS_A	$r - 1$	$MS_A = \frac{SS_A}{r - 1}$	$F_A = \frac{MS_A}{MS_e}$	
因素B	SS_B	$s - 1$	$MS_B = \frac{SS_B}{s - 1}$	$F_B = \frac{MS_B}{MS_e}$	
误差	SS_e	$(r - 1)(s - 1)$	$MS_e = \frac{SS_e}{(r - 1)(s - 1)}$		
总和	SS_T	$rs - 1$			

3.2.2 双因素重复试验的方差分析

(1) 双因素重复试验方差分析试验表

双因素重复试验方差分析试验表

因素	B_1	B_2	...	B_s
A_1	$x_{111}, x_{112}, \dots, x_{11c}$	$x_{121}, x_{122}, \dots, x_{12c}$...	$x_{1s1}, x_{1s2}, \dots, x_{1sc}$
A_2	$x_{211}, x_{212}, \dots, x_{21c}$	$x_{221}, x_{222}, \dots, x_{22c}$...	$x_{2s1}, x_{2s2}, \dots, x_{2sc}$
...
A_r	$x_{r11}, x_{r12}, \dots, x_{r1c}$	$x_{r21}, x_{r22}, \dots, x_{r2c}$...	$x_{rs1}, x_{rs2}, \dots, x_{rsc}$

(2) 双因素重复试验方差分析的基本步骤

① 计算平均值

■ 总平均：
$$\bar{x} = \frac{1}{rsc} \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^c x_{ijk}$$

■ 任一组合水平 (A_i, B_j) 上：
$$\bar{x}_{ij\bullet} = \frac{1}{c} \sum_{k=1}^c x_{ijk}$$

■ A_i 水平时：
$$\bar{x}_{i\bullet\bullet} = \frac{1}{sc} \sum_{j=1}^s \sum_{k=1}^c x_{ijk}$$

■ B_j 水平时：
$$\bar{x}_{\bullet j\bullet} = \frac{1}{rc} \sum_{i=1}^r \sum_{k=1}^c x_{ijk}$$

②计算离差平方和

■ 总离差平方和:

$$SS_T = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^c (x_{ijk} - \bar{x})^2 = SS_A + SS_B + SS_{A \times B} + SS_e$$

■ 因素A引起离差的平方和:

$$SS_A = sc \sum_{i=1}^r (\bar{x}_{i..} - \bar{x})^2$$

■ 因素B引起离差的平方和:

$$SS_B = rc \sum_{j=1}^s (\bar{x}_{.j.} - \bar{x})^2$$

■ 交互作用A×B引起离差的平方和:

$$SS_{A \times B} = c \sum_{i=1}^r \sum_{j=1}^s (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x})^2$$

■ 误差平方和:

$$SS_e = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^c (x_{ijk} - \bar{x}_{ij.})^2$$

③计算自由度

- SS_A 的自由度: $df_A = r - 1$
- SS_B 的自由度: $df_B = s - 1$
- $SS_{A \times B}$ 的自由度: $df_{A \times B} = (r - 1)(s - 1)$
- SS_e 的自由度: $df_e = rs(c - 1)$
- SS_T 的自由度: $df_T = n - 1 = rsc - 1$
- $df_T = df_A + df_B + df_{A \times B} + df_e$

④计算均方

$$MS_A = \frac{SS_A}{r-1}$$

$$MS_B = \frac{SS_B}{s-1}$$

$$MS_{A \times B} = \frac{SS_{A \times B}}{(r-1)(s-1)}$$

$$MS_e = \frac{SS_e}{rs(c-1)}$$

⑤ F检验

$$F_A = \frac{MS_A}{MS_e}$$

$$F_B = \frac{MS_B}{MS_e}$$

$$F_{A \times B} = \frac{MS_{A \times B}}{MS_e}$$

- 若 $F_A > F_{\alpha} (df_A, df_e)$ ，则认为因素A对试验结果有显著影响，否则无显著影响；
- 若 $F_B > F_{\alpha} (df_B, df_e)$ ，则认为因素B对试验结果有显著影响，否则无显著影响；
- 若 $F_{A \times B} > F_{\alpha} (df_{A \times B}, df_e)$ ，则认为交互作用A×B对试验结果有显著影响，否则无显著影响。

⑥重复试验双因素方差分析表

有交互作用双因素试验的方差分析表

差异源	SS	df	MS	F	显著性
因素 A	SS_A	$r-1$	$MS_A = \frac{SS_A}{r-1}$	$F_A = \frac{MS_A}{MS_e}$	
因素 B	SS_B	$s-1$	$MS_B = \frac{SS_B}{s-1}$	$F_B = \frac{MS_B}{MS_e}$	
交互作用 A×B	$SS_{A \times B}$	$(r-1)(s-1)$	$MS_{A \times B} = \frac{SS_{A \times B}}{(r-1)(s-1)}$	$F_{A \times B} = \frac{MS_{A \times B}}{MS_e}$	
误差	SS_e	$rs(c-1)$	$MS_e = \frac{SS_e}{rs(c-1)}$		
总和	SS_T	$rsc-1$			



3.2.3 Excel在双因素方差分析中的应用

(1) 双因素无重复试验方差分析

- 利用“分析工具库”中的“无重复双因素方差分析”工具

(2) 双因素重复试验方差分析

- 利用“分析工具库”中的“重复双因素方差分析”工具