双因素方差分析与多因素的交互作用

夏云天 张晓萌 马嫣然

CONTENT

01

无重复双因素方差分析

数学模型

设因素 A，B 作用于试验指标。因素 A 有 r 个水平 A_1，A_2，…，A_r，因素 B 有 s 个水平 B_1，B_2，…，B_s. 对因素 A，B 的每一个水平的一对组合（A_i，B_j），(i=1，2，…，r，j=1，2，…，s)只进行一次实验，得到 rs 个试验结果 X_{ij}，列于下表中

数学模型

表 8-2-1

试验结果 因素B / 因素A	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}

假设前提

与单因素方差分析的假设前提相同，仍假设：

1) $X_{ij} \sim N(\mu_{ij}, \sigma^2)$，$\mu_{ij}, \sigma^2$ 未知，$i=1,\cdots,r; j=1,\cdots,s.$

2) 每个总体的方差相同；

3) 各 X_{ij} 相互独立，$i=1,\cdots,r; j=1,\cdots,s.$

要判断A、B的影响是否显著，就是要检验同一因素的各个总体的均值是否相等，故检验假设为

$$H_{0A}: \ \mu_{1j} = \mu_{2j} = \cdots = \mu_{rj} = \mu_{\cdot j} \quad j=1,\cdots,s,$$

$$H_{0B}: \ \mu_{i1} = \mu_{i2} = \cdots = \mu_{is} = \mu_{i\cdot} \quad i=1,\cdots,r.$$

备择假设为：

$H_{1A}: \mu_{1j}, \mu_{2j}, \cdots, \mu_{rj}$不全相等。

$H_{1B}: \mu_{i1}, \mu_{i2}, \cdots, \mu_{is}$不全相等。

构建模型

假设：对无重复试验不考虑交互作用

$$\begin{cases} X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}, (i = 1,2,\ldots,r, j = 1,2,\ldots,s) \\ \varepsilon_{ij} \sim N(0,\sigma^2), \mu_{ij}, \sigma^2 \text{未知，各} \varepsilon_{ij} \text{相互独立，} \\ \sum_{i=1}^{r} \alpha_i = 0, \sum_{j=1}^{s} \beta_j = 0. \end{cases}$$

α_i，β_j分别表示因素A和因素B的各水平对指标的影响，ε_{ij}为抽样随机误差。

检验F统计量

当H_{0A}为真时，可以证明

$$F_A=\frac{S_A/(r-1)}{S_E/((r-1)(s-1))}\sim F(r-1,(r-1)(s-1));$$

取显著性水平为α，得假设H_{0A}的拒绝域为

$$F_A=\frac{S_A/(r-1)}{S_E/((r-1)(s-1))}\geq F_\alpha(r-1,(r-1)(s-1));$$

类似地，当H_{0B}为真时，可以证明

$$F_B=\frac{S_B/(s-1)}{S_E/((r-1)(s-1))}\sim F(s-1,(r-1)(s-1));$$

取显著性水平为α，得假设H_{0B}的拒绝域为

$$F_B=\frac{S_B/(s-1)}{S_E/((r-1)(s-1))}\geq F_\alpha(s-1,(r-1)(s-1));$$

检验F统计量

表 8-2-2 无重复试验双因素方差分析表

方差来源	平方和	自由度	均方和	F比
因素A	S_A	$r-1$	$\bar{S}_A=\dfrac{S_A}{r-1}$	$F_A=\bar{S}_A/\bar{S}_E$
因素B	S_B	$s-1$	$\bar{S}_B=\dfrac{S_B}{s-1}$	$F_B=\bar{S}_B/\bar{S}_E$
误差	S_E	$(r-1)(s-1)$	$\bar{S}_E=\dfrac{S_E}{(r-1)(s-1)}$	
总和	S_T	$rs-1$		

Lab

```
aov(formula, data=NULL, projections=FALSE, qr=TRUE,
contrasts=NULL, ...)
```

formula为方差分析公式；
data为数据框；
projection设置是否返回预测结果；
qr设置是否返回QR分解结果；
contrasts为公式中一些因子的列表。

Lab

formula公式的表示：（y为因变量，ABC为分组因子）

符号	用法
~	分隔符号，左边为响应变量，右边为解释变量eg：y~A+B+C
+	分隔解释变量
:	表示变量的交互项eg：y~A+B+A:B
*	表示所有可能交互项eg：y~A*B*C可展开为：y~A+B+C+A:B+A:C+B:C+A:B:C
^	表示交互项达到次数eg：y~(A+B+C)^2展开为：y~A+B+C+A:B+A:C+B:C
.	表示包含除因变量外的所有变量eg：若一个数据框包括变量y,A、B和C，代码y~.可展开为y~A+B+C

Lab

常见研究设计的表达式：（小写字母表示定量变量，大写字母表示组别因子，Subject是对被试者独有的标识变量）

设计	表达式
单因素ANOVA	y~A
含单个协变量的单因素ANCOVA	y~x+A
双因素ANOVA	y~A*B
含两个协变量的双因素ANCOVA	y~x1+x2+A*B
随机化区组	y~B+A，B为区组因子
单因素组内ANOVA	y~A+Error（Subject/A）
含单个组内因子（W）和单个组间因子（B）的重复测量ANOVA	y~B*W+Error（Subject/W）

例题

1、将一种生长激素配成M1，M2，M3，M4和M5五个浓度，并用H1、H2和H3三种时间浸泡某大豆品种的种子，出苗45天后得各处理每一植株的平均干物质（g）如下。试问生长激素浓度与浸泡时间是否对植株的平均干物质重量有影响？

	H1	H2	H3
M1	13	14	14
M2	12	12	13
M3	3	3	3
M4	10	9	10
M5	2	5	4

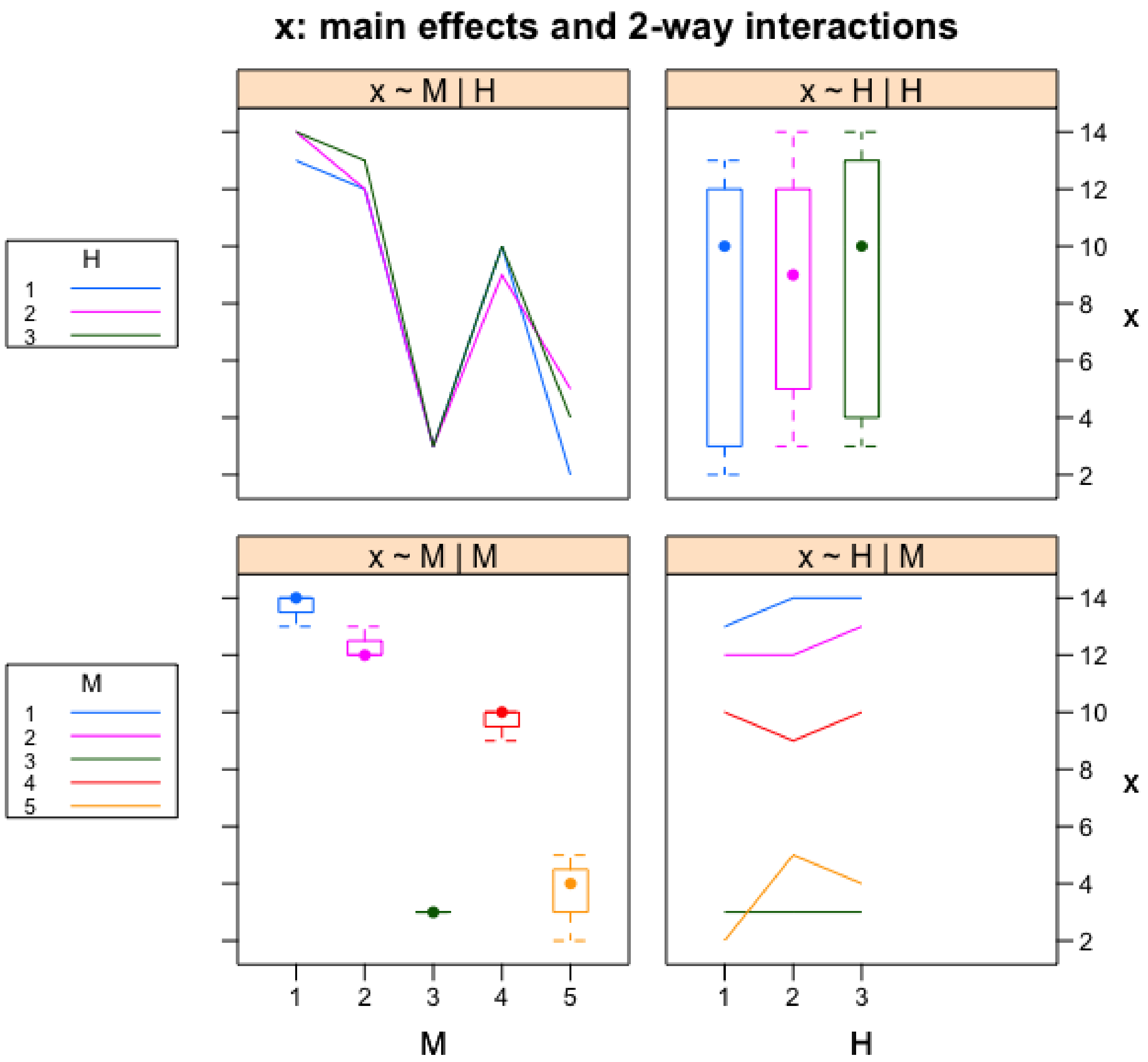
例题

```
drug<-data.frame(
x=c(13,14,14,
   12,12,13,
   3,3,3,
   10,9,10,
   2,5,4),
M=gl(5,3),
H=gl(3,1,15)
)
> bartlett.test(x~M,data=drug)

        Bartlett test of homogeneity of variances

data:  x by M
Bartlett's K-squared = Inf, df = 4, p-value < 2.2e-16

> bartlett.test(x~H,data=drug)

        Bartlett test of homogeneity of variances

data:  x by H
Bartlett's K-squared = 0.0492, df = 2, p-value = 0.9757
```

例题

```
> summary(aov(x~M+H,data=drug))
            Df Sum Sq Mean Sq F value
Pr(>F)
M            4 289.07   72.27 117.189
3.91e-07 ***
H            2   1.73    0.87   1.405
0.3
Residuals    8   4.93    0.62
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01
‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> qf(0.95,4,8)
[1] 3.837853
> qf(0.95,2,8)
[1] 4.45897
```

例题



x: main effects and 2-way interactions
x ~ M | H
x ~ H | H
x ~ M | M
x ~ H | M
H
1
2
3
M
1
2
3
4
5
14
12
10
8
6
4
2
x
1
2
3
4
5
M
1
2
3
H

02

可重复双因素方差分析

理论部分

无重复实验的双因素方差分析和有重复实验的双因素方差分析:

1、设计方面,无重复实验,每个组合下只做一次实验,得到一个数据,而重复实验,是每个组合下做多次实验,得到多个数据.

2、分析方面：无重复,由于每个组合下只有一个数据，只能估计交互作用和偶然误差的联合偏差；而重复的能估计偶然误差造成的偏差，可以考虑这个组合下的抽样误差的大小,从而将总变异分解的更细,也即将试验因素的作用体现的更明显,可以看出两因素的独立作用与交互作用。

3、变异分无重复的双因素方差分析,总变异分解为：处理组、配伍组、误差重复的双因素方差分析,总变异分解为：处理组、配伍组、两因素的交互作用、误差

数学模型

$$
\begin{cases}
X_{ijk} = \mu + a_i + b_j + (ab)_{ij} + \varepsilon_{ijk}, \\
\varepsilon_{ijk} \sim N(0, \sigma^2), \text{各} \varepsilon_{ijk} \text{独立}, \\
i = 1, 2, \cdots, r, j = 1, 2, \cdots, s, k = 1, 2, \cdots, l, \\
\sum_{i=1}^{r} a_i = 0, \sum_{j=1}^{s} b_j = 0, \sum_{i=1}^{r} (ab)_{ij} = 0, \sum_{j=1}^{s} (ab)_{ij} = 0.
\end{cases}
$$

假设检验

$$\begin{cases} H_{01}: a_1 = a_2 = \cdots = a_r = 0, \\ H_{11}: a_1, a_2, \cdots, a_r \text{不全为零}. \end{cases}$$

$$\begin{cases} H_{02}: b_1 = b_2 = \cdots = b_s = 0, \\ H_{12}: b_1, b_2, \cdots, b_s \text{不全为零}. \end{cases}$$

$$\begin{cases} H_{03}: (ab)_{11} = (ab)_{12} = \cdots = (ab)_{rs} = 0, \\ H_{13}: (ab)_{11}, (ab)_{12}, \cdots, (ab)_{rs} \text{不全为零}. \end{cases}$$

方差分析表

双因素重复试验方差分析表

误差来源	平方和	自由度	均方	F 值	显著性
因素 A	$S_A = Q_A - P$	$r-1$	$MS_A = \dfrac{S_A}{r-1}$	$F_A = \dfrac{MS_A}{MS_E}$	
因素 B	$S_B = Q_B - P$	$s-1$	$MS_B = \dfrac{S_B}{s-1}$	$F_B = \dfrac{MS_B}{MS_e}$	
交互作用 $A \times B = I$	$S_I = R - Q_A - Q_B + P$	$(r-1)(s-1)$	$MS_I = \dfrac{S_I}{(r-1)(s-1)}$	$F_I = \dfrac{MS_I}{MS_e}$	
误差 e	$S_e = W - R$	$rs(l-1)$	$MS_e = \dfrac{S_e}{rs(l-1)}$		
总和	$S_T = W - P$	$rsl-1$			

例题

例、下面给出某化工产品在3种浓度、4种温度水平下得到的质量指标：

浓度 温度	10		24		38		52	
2	14	10	11	11	13	9	10	12
4	9	7	10	8	7	11	6	10
6	5	11	13	14	12	13	14	10

试在水平α =0.05下检验：浓度的差异对质量指标有无影响；温度的差异对质量指标有无影响；交互作用对质量指标有无影响。

例题

```
qual<-data.frame(
  x=c(14,10,11,11,13,9,10,12,9,7,10,8,7,11,6,10,5,11,13,14,12,13,14,10),
  den=gl(3,8),
  tem=gl(4,2,24)
)
> head(qual)
   x den tem
1 14   1   1
2 10   1   1
3 11   1   2
4 11   1   2
5 13   1   3
6  9   1   3
> table(den, tem)
   tem
den 1 2 3 4
  1 2 2 2 2
  2 2 2 2 2
  3 2 2 2 2
```

例题

```
> bartlett.test(x~den,data=qual)

	Bartlett test of homogeneity of variances

data:  x by den
Bartlett's K-squared = 2.8585, df = 2, p-value =
0.2395

> bartlett.test(x~tem,data=qual)

	Bartlett test of homogeneity of variances

data:  x by tem
Bartlett's K-squared = 0.7479, df = 3, p-value =
0.8619
```

方差齐次性检验，p>0.05。

例题

```
> fit<-aov(x~den*tem,data=qual)
> summary(fit)
            Df Sum Sq Mean Sq F value Pr(>F)  
den          2  44.33  22.167   4.092 0.0442 *
tem          3  11.50   3.833   0.708 0.5657  
den:tem      6  27.00   4.500   0.831 0.5684  
Residuals   12  65.00   5.417                 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 
‘ ’ 1
> qf(0.95,2,12)
[1] 3.885294
> qf(0.95,3,12)
[1] 3.490295
> qf(0.95,6,12)
[1] 2.99612
```

浓度：p=0.0442<0.05，显著

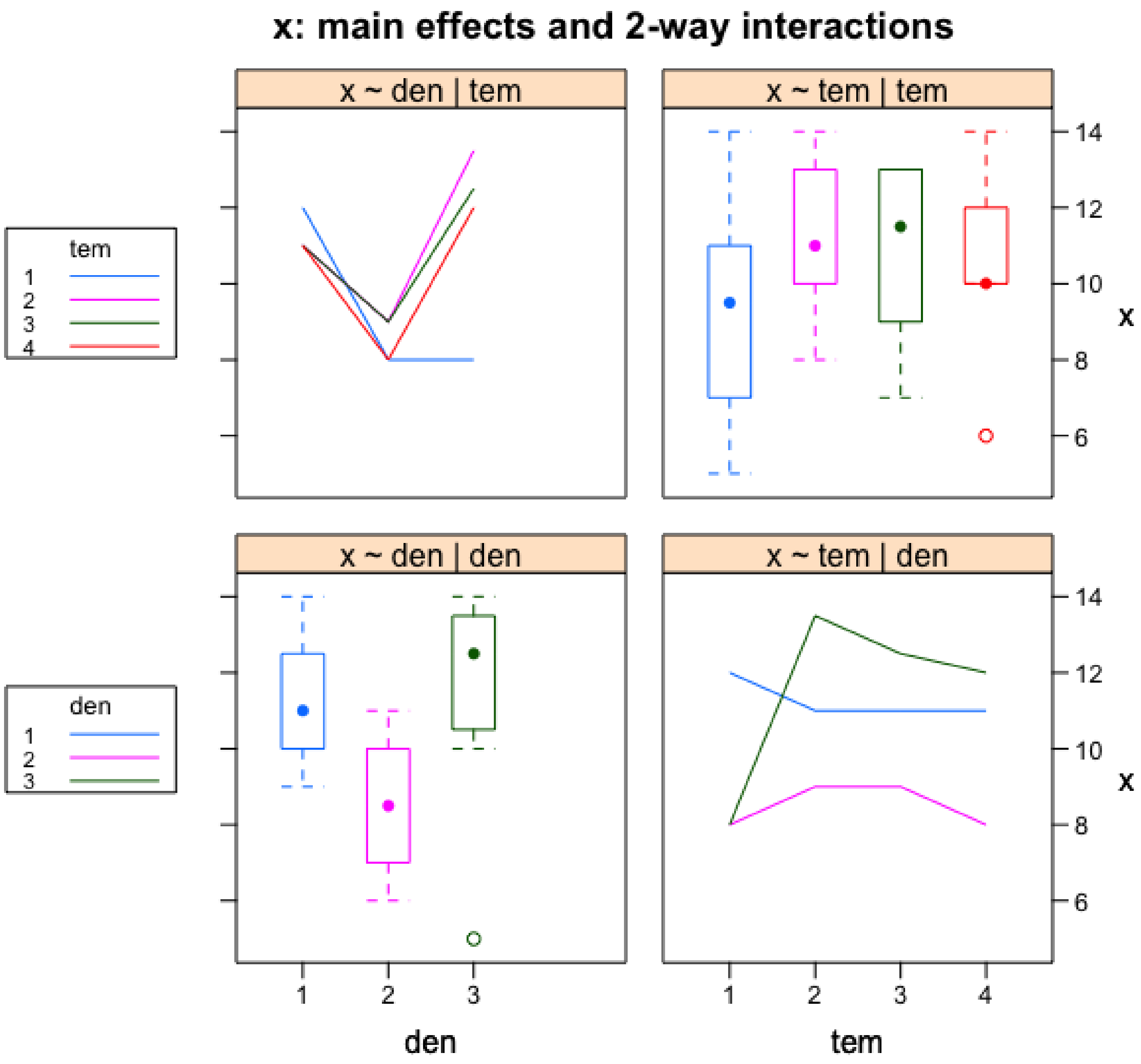
温度：p=0.5657>0.05，不显著

交互作用：p=0.5684>0.05，不显著

```
>library(HH)
>p <- interaction2wt(x~den*tem,data=qual)
>plot(p)
```

图形化展示两因素方差分析的交互效应

例题



x: main effects and 2-way interactions
x ~ den | tem
x ~ tem | tem
x ~ den | den
x ~ tem | den
tem
1
2
3
4
den
1
2
3
14
12
10
8
6
x
1
2
3
den
1
2
3
4
tem

03

习题

习题

1、某汽车轮胎厂在研制新型车胎橡胶配方中，考虑3种不同的促进剂：（A）4种不同分量的氧化锌；（B）三种不同的促进剂。每种配方各做一次实验，测得300%定强如下：

	B1	B2	B3	B4
A1	31	34	35	39
A2	33	36	37	38
A3	35	37	39	42

试检验促进剂、氧化锌对定强有无显著影响。

习题

2、下表记录了3位操作工分别在不同机器上操作3天的日产量，取显著性水平α =0.05，试分析操作工之间、机器之间以及交互作用之间有无显著性差异？

	甲			乙			丙		
A1	15	15	17	19	19	16	16	18	21
A2	17	17	17	15	15	15	19	22	22
A3	15	17	16	18	17	16	18	18	18
A4	18	20	22	15	16	17	17	17	17

THANK YOU