

微型计算机BASIC算法语言通用程序选介 (续)

刘 乐 融

(中国科学院西北水土保持研究所)

II —— 方差分析

方差分析是分析试验(或观测)数据的一种方法。它所解决的基本问题是通过数据的分析,弄清与研究对象有关的各个因素以及各个因素之间交互作用对该对象的影响。

在方差分析中,分“单因素方差分析”和“多因素方差分析”介绍。在方法概要中对数学推导过程不作详细介绍,仅给出计算公式。

一、单因素方差分析

1、方法概要。若有统计假设 H_0 ,并视 X_{ij} 是一个随机变量的一批观测值,则有单因素试验的数据汇总表1。

表1 单因素试验数据汇总表

	1	2	j	b	合 计	平 均
1	x_{11}	x_{12}	x_{1j}	x_{1b}	T_1	\bar{X}_1
2	x_{21}	x_{22}	x_{2j}	x_{2b}	T_2	\bar{X}_2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
i	x_{i1}	x_{i2}	x_{ij}	x_{ib}	T_i	\bar{X}_i
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
a	x_{a1}	x_{a2}	x_{aj}	x_{ab}	T_a	\bar{X}_a

其中:

$$T_i = \sum_{j=1}^b X_{ij} \qquad i = 1, 2, \dots, a$$

$$\bar{X}_i = \frac{1}{b} \sum_{j=1}^b X_{ij} = \frac{T_i}{b}$$

$$\bar{X} = \frac{1}{a \cdot b} \sum_{j=1}^b \sum_{i=1}^a X_{ij} = \frac{T}{a \cdot b}$$

这样,我们可建立单因素方差分析计算表2。

表2

单因素方差分析计算表

	离 差 平 方 和	自 由 度	平均离差平方和	F
组与组间	$Q_1 = \sum_{i=1}^a \frac{T_i^2}{b_i} - \frac{T^2}{A}$	$a - 1$	$S_1 = \frac{Q_1}{a-1}$	$F = \frac{Q_1/(a-1)}{Q_2/(A-a)}$
组内部	$Q_2 = \sum_{i=1}^a \sum_{j=1}^{b_i} X_{ij}^2 - \sum_{i=1}^a \frac{T_i^2}{b_i}$	$A - a$	$S_2 = \frac{Q_2}{A-a}$	
总 和	$Q = \sum_{i=1}^a \sum_{j=1}^{b_i} (X_{ij} - \bar{X})^2$	$A - 1$		

其中:

$$A = \sum_{i=1}^a b_i = \text{试验单元的总计数}$$

如果 $F > F_\alpha$, 则统计假设 H_0 不成立; 如果 $F \leq F_\alpha$, 则统计假设 H_0 成立。这就是单因素方差分析的基本原理与做法。

2. 程序清单

```

1000 CLS
1010 INPUT "N=";N:
      INPUT "K=";K
1020 DIM X(N),NI(K),MC(K)
1030 FOR I=1 TO N
1040 READ X(I):LPRINT X(I),
1050 NEXT I:PRINT " "
1060 FOR I=1 TO K
1070 INPUT "N"; I,"=";NI(I)
      :LPRINT "N",
      I,"=";NI(I)
1080 NEXT I:LPRINT " "
1090 N=0
1100 FOR I=1 TO K
1110 MI=0:NC=NI(I)
1120 FOR J=1 TO NC
1130 ST=X(N+J):MI=MI+ST
      :SW=SW+ST^2
1140 NEXT J
1150 MX=MX+MI:N=N+NC
1160 MC(I)=MI/NC:
      SB=SB+MI^2/NC

```

```

1170 NEXT I
1180 MX=MX/N:SW=SW-SB
1190 SB=SB-MX^2*N:ST=SW+SB
1200 F=SB*(N-K)/(SW*(K-1))
1210 FOR I=1 TO K
1220 LPRINT "X",I="MC(I),
1230 NEXT I:LPRINT:" "
1240 LPRINT "Q1=",SB,"S1=",
      SB/(K-1),"V1=",K-1
1250 LPRINT "Q2=",SW,"S2=",
      SW/(N-K),"V2="N-K
1260 LPRINT "Q=",ST,"V=",N-
      1,:LPRINT "X=",MX
1270 LPRINT "F=",F
1280 RETURN
1290 END

```

3. 使用说明

N—试验数据总数
 K—因子分组数
 N_i—各组试验数据的个数
 (i=1,2,...b)
 X_i—各组均值即 \bar{X}_i

($i = 1, 2, \dots, b$)

\bar{X} —总均值

Q_1 —各组间差方和

Q_2 —各组内差方和

Q —总差方和

S_i —平均差方和

V_i —自由度

例题 研究 6 种氮肥施用法对小麦的效应, 每种施肥法种 5 盆小麦, 完全随机设计, 分别测定它们的含氮量(毫克) 其结果如下, 试作方差分析。

施 氮 法					
1	2	3	4	5	6
12.9	14.0	12.6	10.5	14.6	14.0
12.3	13.8	13.2	10.8	14.6	13.3
12.2	13.8	13.4	10.7	14.4	13.7
12.5	13.6	13.4	10.8	14.4	13.5
12.7	13.6	13.0	10.5	14.4	13.7

主程序:

```
100 DATA 12.9,12.3,12.2,12.5,12.7,14.0,
      13.8,13.8
110 DATA 13.6,13.6,12.6,13.2,13.4,13.4,
      13.0,10.5
120 DATA 10.8,10.7,10.8,10.5,14.6,14.6,
      14.4,14.4
130 DATA 14.4,14.0,13.3,13.7,13.5,13.7
140 DATA GOSUB 1000
150 END
```

```
键入      显示
RUN RETURN N=? RETURN
30 RETURN K=? RETURN
6 RETURN N1=? RETURN
5 RETURN N2=? RETURN
5 RETURN N3=? RETURN
5 RETURN N4=? RETURN
```

```
5 RETURN N5=? RETURN
5 RETURN N6=? RETURN
5 RETURN
N=30      K=6
DATA
2.9 12.3 12.2 12.5 12.7 14.0 13.8 13.8
13.6 13.6 12.6 13.2 13.4 13.4 13.0 10.5
10.8 10.7 10.8 10.5 14.6 14.6 14.4 14.4
14.4 14.0 13.3 13.7 13.5 13.7
N1=5      N2=5      N3=5
N4=5      N5=5      N6=5
X1=12.52  X2=13.76  X3=13.12
X4=10.66  X5=14.48  X6=13.64
Q1=44.463 S1=8.8926 V1=5
Q2=1.29996 S2=0.0541649 V2=24
Q=45.7629 V=29      X=3.03
F=164.176
```

若查F表得 $F_{0.01} = 3.90$, $F \gg F_{0.01}$, 故有推断: 否定 H_0 , 即 6 种施氮法的植株含氮有显著差异。

本程序仍适应分组数据不等情况。

二、多因素方差分析

1、方法概要。设有A、B两因素, X_{ij} 为一个随机变量的一批观测值。那么就有两因素无重

复试验的数据汇总表 3。

表3 两因素无重复试验数据汇总表

A \ B							$T_{i.}$	$\bar{X}_{i.}$
	1	2	j	b		
1	X_{11}	X_{12}	X_{1j}	X_{1b}	$T_{1.}$	$\bar{X}_{1.}$
2	X_{21}	X_{22}	X_{2j}	X_{2b}	$T_{2.}$	$\bar{X}_{2.}$
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots	\vdots
i	X_{i1}	X_{i2}	X_{ij}	X_{ib}	$T_{i.}$	$\bar{X}_{i.}$
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots	\vdots
a	X_{a1}	X_{a2}	X_{aj}	X_{ab}	$T_{a.}$	$\bar{X}_{a.}$
$T_{.j}$	$T_{.1}$	$T_{.2}$	$T_{.j}$	$T_{.b}$	T	
$\bar{X}_{.j}$	$\bar{X}_{.1}$	$\bar{X}_{.2}$	$\bar{X}_{.j}$	$\bar{X}_{.b}$		

其中： $T_{i.}$ 代表第 i 行的数据 ($X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{ib}$) 之和， $T_{.j}$ 代表按第 j 列的数据 ($X_{1j}, X_{2j}, \dots, X_{ij}, \dots, X_{aj}$) 之和， $\bar{X}_{i.}$ 代表数据按第 i 行的平均， $\bar{X}_{.j}$ 代表数据按第 j 列的平均。
即：

$$\bar{X}_{i.} = T_{i.}/b \quad \bar{X}_{.j} = T_{.j}/a$$

$$T = \sum_{j=1}^b T_{.j} = \sum_{i=1}^a T_{i.} = \sum_{i=1}^a \sum_{j=1}^b X_{ij}$$

据推导有两因素无交互作用的方差分析计算表 4。

如果设有统计假设 H_{01}, H_{02} ，若 $F_1 \leq F_{.1}$ ，则假设 H_{01} 成立；若 $F_2 \leq F_{.2}$ ，则假设 H_{02} 成立。反之，若 $F_1 > F_{.1}$ ，则假设 H_{01} 不成立；若 $F_2 > F_{.2}$ ，则假设 H_{02} 不成立。这就是两因素无交互作用时的方差分析方法。

2. 程序清单

```

1000 CLS
1010 INPUT "R=";R:INPUT
      "N=";N
1020 DIM X(N),LE(R),T(2^R),N(2^R),
      MS(2^R),F(2^R),S(2^R)
1030 DIM I(R),TI(R),TL(R),
      Q(N,N),A(N),B(N)
1040 FOR J=1 TO N
1050 READ X(J):LPRINT X(J);
1060 NEXT J:LPRINT " " :
      LPRINT " "
1061 FOR J=1 TO R
1062 INPUT "L";J;"=";LE(J):
      LPRINT "L";J;"=";LE(J),
1063 NEXT J:LPRINT " " :
      LPRINT " "
1070 GOTO 1140
1080 T=0
1090 FOR J=R TO 1 STEP -1
1100 T=T*LE(J)+S(J)-1
1110 NEXT J
1120 T=T+1
1130 RETURN
1140 FOR F=1 TO R

```

表4

两因素无交互作用的方差分析表

	离 差 平 方 和	自 由 度	平均离差平方和	F
各行作用	$Q_1 = b \sum_{i=1}^a (\bar{x}_{i.} - \bar{x})^2$ $= \frac{1}{b} \sum_{i=1}^a T_{i.}^2 - \frac{T^2}{ab}$	$a - b$	$S_1 = Q_1 / (a - 1)$	$F_1 = \frac{Q_1 / (a - 1)}{Q_3 / (a - 1)(b - 1)}$
各列作用	$Q_2 = a \sum_{j=1}^b (\bar{x}_{.j} - \bar{x})^2$ $= \frac{1}{a} \sum_{j=1}^b T_{.j}^2 - \frac{T^2}{ab}$	$b - 1$	$S_2 = Q_2 / (b - 1)$	$F_2 = \frac{Q_2 / (b - 1)}{Q_3 / (a - 1)(b - 1)}$
剩 余	$Q_3 = \sum_{i=1}^a \sum_{j=1}^b (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$ $= \sum_{i=1}^a \sum_{j=1}^b x_{ij}^2 - \frac{1}{a} \sum_{j=1}^b T_{.j}^2$ $- \frac{1}{b} \sum_{i=1}^a T_{i.}^2 + \frac{T^2}{ab}$	$(a - 1)(b - 1)$	$S_3 = \frac{Q_3}{(a - 1)(b - 1)}$	
总 和	$Q = \sum_{i=1}^a \sum_{j=1}^b (x_{ij} - \bar{x})^2$ $= \sum_{i=1}^a \sum_{j=1}^b x_{ij}^2 - \frac{T^2}{ab}$	$ab - 1$		

```

1150 NI=LE(F)
1160 FOR K=1 TO NI
1170 Q(K,1)=1/SQR(NI)
1180 NEXT K
1190 FOR J=2 TO NI
1200 FOR K=1 TO J-1
1210 Q(K,J)=-1/SQR(J*(J-1))
1220 NEXT K
1230 Q(J,J)=SQR((J-1)/J)
1240 FOR K=J+1 TO NI
1250 Q(K,J)=0
1260 NEXT K
1270 NEXT J

```

```

1280 FOR J=1 TO R
1290 I(J)=1
1300 NEXT J
1310 FOR K1=1 TO NI
1320 I(F)=K1
1330 FOR J1=1 TO R
1340 S(J1)=I(J1):
      L(J1)=LE(J1)
1350 NEXT J1
1360 GOSUB 1080
1370 B1=T
1380 A(K1)=X(B1)
1390 NEXT K1

```

```

1400 FOR K2=1 TO NI
1410 B(K2)=0
1420 FOR J2=1 TO NI
1430 B(K2)=B(K2)+Q(J2,K2)
      *A(J2)
1440 NEXT J2
1450 NEXT K2
1460 FOR K3=1 TO NI
1470 I(F)=K3
1480 FOR J3=1 TO R
1490 S(J3)=I(J3) :
      L(J3)=LE(J3)
1500 NEXT J3
1510 GOSUB 1080
1520 B1=T
1530 X(B1)=B(K3)
1540 NEXT K3
1550 FOR J4=1 TO R
1560 IF J4=F THEN 1590
1570 I(J4)=I(J4)+1
1580 IF I(J4)<=LE(J4) THEN
      1310 ELSE I(J4)=1
1590 NEXT J4
1600 NEXT F
1610 FOR J=1 TO R
1620 TI(J)=1 : TL(J)=2
1630 NEXT J
1640 FOR J5=1 TO R
1650 I(J5)=TI(J5) : S(J5)=TI
      (J5) : L(J5)=TL(J5)
1660 NEXT J5
1670 GOSUB 1080
1680 B1=T
1690 T(B1)=0
1700 FOR J6=1 TO R
1710 S(J6)=I(J6) :
      L(J6)=LE(J6)
1720 NEXT J6
1730 GOSUB 1080
1740 B2=T
1750 T(B1)=T(B1)+X(B2)*X(B2)
1760 FOR J7=1 TO R
1770 IF TI(J7)<>2 THEN 1800

1780 I(J7)=I(J7)+1
1790 IF I(J7)<=LE(J7) THEN
      1700 ELSE I(J7)=2
1800 NEXT J7
1810 FOR J8=1 TO R
1820 TI(J8)=TI(J8)+1
1830 IF TI(J8)<=2 THEN
      1640 ELSE TI(J8)=1
1840 NEXT J8
1850 FOR J=1 TO R
1860 TI(J)=1
1870 NEXT J
1880 GOTO 1990
1890 DF=1
1900 FOR K4=1 TO R
1910 IF TI(K4)=2 THEN
      DF=DF*(LE(K4)-1)
1920 NEXT K4
1930 FOR K5=1 TO R
1940 S(K5)=TI(K5) : L(K5)=TL(K5)
1950 NEXT K5
1960 GOSUB 1080
1970 B1=T
1980 N(B1)=DF : MS(B1)=T(B1)/N(B1)
1990 FOR K6=1 TO R
2000 TI(K6)=TI(K6)+1
2010 IF TI(K6)<=2 THEN
      1890 ELSE TI(K6)=1
2020 NEXT K6
2030 B2=2^R-1
2040 FOR J=2 TO B2
2050 F(J)=MS(J)/MS(B2+1)
2060 NEXT J
2070 LPRINT TAB(1)"N";TAB(15)
      "Q";TAB(31)"V";
      TAB(46)"S";TAB(62)"F"
2080 FOR J=2 TO B2+1
2090 LPRINT TAB(1)J-1;TAB(12)
      T(J);TAB(31)N(J);
      TAB(43)MS(J);TAB(59)F(J)
2100 NEXT J
2120 RETURN

```

3、使用说明

Q—离差平方和

V—自由度

S—平均离差平方和

Li—各因子的水平数

NO—各因子的号数

例题 用生长素处理豌豆，共6个处理。种子发芽后，分别在每一木箱中移植4株，每组6个木箱，每箱一个处理。共4组24箱，试验组排于温室中，使各箱环境条件一致。记录各箱见第一朵花时4株豌豆的总节间数其结果如表5，试作方差析。

表 5

处 理 (A)	组 (B)			
	I	II	III	IV
对照 (未处理)	60	62	61	60
赤 霉 素	65	65	68	65
动 力 精	63	61	61	60
咧 噪 乙 酸	64	67	63	61
硫 酸 腺 嘌呤	62	65	62	64
马 来 酸	61	62	62	65

主程序:

```
100 DATA 60,65,63,64,62,61,62,65,61,
      67,65,62
```

```
110 DATA 61,68,61,63,62,62,60,65,60,
      61,64,65
```

```
120 GOSUB 1000
```

```
130 END
```

键入

显示

```
RUN RETURN R=? RETURN
```

```
2 RETURN N=? RETURN
```

```
24 RETURN L1=? RETURN
```

```
6 RETURN L2=? RETURN
```

```
4 RETURN
```

```
R = 2 N = 24
```

-----DATA-----

```
60 65 63 64 62 61 62 65 61 67 65 62 61
```

```
68 61 63 62 72 60 65 60 61 64 65
```

L1 = 6

L2 = 4

NO.	Q	V	S	F
1	65.875	5	13.175	4.56497
2	5.45829	3	1.81943	0.630409
3	43.2917	15	2.88611	0

这样有方差分析结果:

表6

	离差平方和	自 由 度	平均离差平方和	F	$F_{0.05}$
处理 (A) 作用	$Q_1 = 65.875$	5	$S_1 = 13.175$	4.565	2.90
组 (B) 作用	$Q_2 = 5.458$	3	$S_2 = 1.8194$	0.631	3.29
误 差	$Q_3 = 43.292$	15	$S_3 = 2.8861$		
总 合	$Q = 114.625$	23			

从表6可看出, $F_1 > F_{0.05}$, $F_2 < F_{0.05}$ 。这说明, 不同生长素的处理有显著差异, 而组间环境条件无显著差异。

(续完)