

第二节 单因素方差分析与协方差分析实验

方差分析是检验两个或两个以上的多个样本均值间差异是否具有统计意义的一种统计分析方法. 统计分析软件 SAS 和 SPSS 均提供了多种进行方差分析的过程. 本节介绍分别用 SAS 软件和 SPSS 软件进行各种单因素方差分析与协

方差分析的方法和步骤.

一、单因素完全随机试验的方差分析

1. 单因素完全随机等重复试验的方差分析

只考察一个因素 A , 参试水平或处理为 a 个, 设置重复, 这样的试验称为单因素完全随机试验. 如果试验的每个水平安排相同数目的试验单位, 即等重复试验, 则这种试验设计称为平衡设计.

例 1 设有五种治疗荨麻疹的药, 要比较它们的疗效. 为此, 将 30 个病人随机分成五组, 每组 6 人, 同组的病人使用一种药物, 并记录下病人从用药开始到痊愈所需天数 x (见下表). 试检验五种药物的疗效有无显著差异.

药物 A	治愈所需天数					
A1	6	8	7	7	10	8
A2	4	6	6	3	5	6
A3	6	4	4	5	3	2
A4	7	4	6	6	3	5
A5	9	4	5	7	7	6

(1) 在 SAS 软件中可用 ANOVA 过程进行单因素完全随机试验的方差分析. ANOVA 过程的调用格式为

```
PROC ANOVA 选项;
CLASS 变量表;
MODEL 观测变量表 = 效应表;
MEANS 效应表/选项;
TEST H = 效应表 E = 效应;
BY 变量表;
```

前三个语句是必需的. MEANS 和 TEST 语句只能在 MODEL 语句之后, 并可使用多次. 其中

PROC ANOVA 选项; 用于调用分析过程, 该语句常用的选项有:

DATA = 数据集: 指定一个要分析的输入数据集.

OUTSTAT = 数据集: 指定一个输出数据集, 其中包括平方和、F 统计量及模型中每个效应的概率水平等信息.

CLASS 变量表; 语句声明方差分析中因素的分类水平处理变量, 即分类变量. 分类变量必须首先在本语句中指明. 这些变量可以是数值型的, 也可以是字符型的, 但字符变量长度不得超过 16 个字符.

MODEL 观测变量表 = 效应表; 语句指定方差分析中观测变量及因素效

应名称.

MEANS 效应表/选项;语句计算由效应表所列因素效应的各水平均值,并对两两均值相等的假设作显著性检验.斜线“/”后常用的选项有

BON:指定效应均值差异的 BONFERRONI 检验.

DUNCAN:指定效应均值差异的 DUNCAN 多重比较检验.

LSD:最小显著差数法,进行配对 t 检验.

T:指定效应均值差异的配对 t 检验.

ALPHA = P:指定均值差异的显著性水平,默认为 $P = 0.05$.

CLDIFF:要求对所有配对均值差异以置信区间给出.

TEST H=效应表 E=效应;在方差分析中,如果以其他效应作为误差项进行 F 检验,则可用 TEST 语句.其中

H=效应表:指定 MODEL 语句中某些效应作为假设效应,这些效应的平方和作为 F 统计量的分子.

E=效应:指定一个效应,此效应作为误差项,相应的平方和作为 F 统计量的分母.

BY 变量表;对 BY 语句指定的变量组分别进行分析.

对于本例,用 ANOVA 过程进行分析的程序如下

```
DATA EX01;
INPUT A X@@@;
CARDS;
1 6 2 4 3 6 4 7 5 9
1 8 2 6 3 4 4 4 5 4
1 7 2 6 3 4 4 6 5 5
1 7 2 3 3 5 4 6 5 7
1 10 2 5 3 3 4 3 5 7
1 8 2 6 3 2 4 5 5 6
;
PROC ANOVA;
CLASS A;
MODEL X = A;
MEANS A/LSD;
RUN;
```

程序中 A 为分组变量, X 为治愈所需天数. MEANS 语句要求对 X 的均值用最小显著差数法进行多重比较. 程序输出的主要分析结果如下

Analysis of Variance Procedure

Dependent Variable: X

Source	DF	Sum of Squares	F Value	Pr > F
Model	4	47.46666667	5.55	0.0025
Error	25	53.50000000		
Corrected Total	29	100.96666667		

Analysis of Variance Procedure

Dependent Variable: X

R-Square	C. V.	X Mean
0.470122	25.96818	5.63333333

Analysis of Variance Procedure

Dependent Variable: X

Source	DF	Anova SS	F Value	Pr > F
A	4	47.46666667	5.55	0.0025

Analysis of Variance Procedure

T tests (LSD) for variable: X

NOTE: This test controls the type I comparisonwise error rate
not the experimentwise error rate.

Alpha = 0.05 df = 25 MSE = 2.14

Critical Value of T = 2.06

Least Significant Difference = 1.7395

Means with the same letter are not significantly different.

T	Grouping	Mean	N	A
	A	7.6667	6	1
	A			
B	A	6.3333	6	5
B				
B	C	5.1667	6	4
B	C			
B	C	5.0000	6	2
	C			
	C	4.0000	6	3

上述结果中,方差分析表中的 F 值为 5.55,显著性水平达 0.0025;中间给出

了 R^2 值、变异系数(C.V.)和 $\bar{x}$ 的均值,最后是最小显著差数法(LSD)的多重比较 t 检验结果。

(2) 在 SPSS 软件中可用 One-Way ANOVA 过程进行单因素完全随机试验的方差分析。One-Way ANOVA 过程要求因变量属于正态分布总体,如果因变量的分布明显是非正态的,则不能使用该过程,它仅适用于进行单因素完全随机试验的方差分析。本例用 One-Way ANOVA 过程进行方差分析的步骤为:

- 在数据窗口中建立数据文件,定义两个数值型变量 fact(标签为“因素”)、x(标签为“疗效天数”),并输入相应数据。变量 fact 格式为 F4.0,取值 1,2,3,4,5 分别代表药物的五个水平;变量 x 格式为 F4.0,其值为治愈所需天数。

- 在系统的“Statistics”菜单中,选择“Compare Means”子菜单项下的“One-Way ANOVA”过程,打开主对话框。然后在该对话框中的“Dependent List”中指定进行方差分析的因变量 x,在“Factor”中指定因素变量(分类变量)fact。

- 主对话框中还有一些可选项的选择按钮,用户可以根据不同需要对这些选项作出选择。这些选项分为三类,分别对应于主对话框中的三个按钮“Contrasts”、“Post Hoc”和“Options”,各选项及功能分别介绍如下:

① 进行均值的多项式比较选择项“Contrasts”

在主对话框中单击“Contrasts”按钮,打开相应的对话框。均值的多项式比较是包括两个或更多个均值的比较。例如多项式

$$1.0 * \text{mean1} - 1.0 * \text{mean2} - 1.0 * \text{mean3} + 1.0 * \text{mean4}$$

表示检验第一、四组的效应和与第二、三组效应和之间是否有显著性差异。多项式的系数由用户根据需要输入。具体操作如下

1) 选择“polynomial”选项。

2) 单击“Degree”项右边向下的箭头向下拉出一菜单,选择其中的一种多项式,如线性(Linear)。

3) 在“Coefficients”框中,依次输入多项式指定各组均值的系数,每输入一个系数后单击“Add”按钮,输入的系数依次显示在下面的显示框中。也可以同时建立多个多项式。当一个多项式的一组系数输入结束后单击“Next”按钮,准备接收下一组系数数据。

4) 单击“Previous”或“Next”按钮,显示输入的各组系数,检查无误后单击“Continue”按钮确认,并返回主对话框。

② 各组均值的多重比较选择项“Post Hoc”

在主对话框中单击“Post Hoc”按钮,打开相应对话框。在该对话框中选择进行均值多重比较的方法,同时还可以选择样本含量的估计方法。

对话框的“Equal Variances Assumed”框中列出了多种多重比较的方法,选择时只需单击所选项前面的复选按钮,这些选项可以同时选择多个。

对话框的“Sample Size Estimate”框中,列有样本含量估计方法的两个选项,用户可根据需要选择其一。

③ 输出统计量的选项“Options”

在主对话框中单击“Options”按钮打开相应对话框,对话框中列出了输出统计量的选项和对缺值处理的选项等。其中“Statistics”框中列出的输出统计量选项有

1) “Descriptive”:要求输出描述统计量,选择此项会计算并输出观测量数目、均值、标准差、标准误、最小值、最大值、各组中每个因变量的 95% 置信区间。

2) “Homogeneity-of-Variance”:要求进行方差齐次性检验,并输出检验结果。该项采用“Levene Lest”检验,即计算每个观测量与其均值之差,然后对这些差值进行一维方差分析。

在“Missing Values”框中列出的选择项有

1) “Exclude cases analysis by analysis”:对含有缺失值的观测量,根据缺失值是因变量还是自变量从有关的分析中剔除。

2) “Exclude cases listwise”:对含有缺失值的观测量从所有分析中剔除。

对话框中还有一个“Means Plot”选项,选择此项可绘出因变量均值对因素水平的散点折线图。

以上三项选择完成后,单击“Continue”按钮,返回主对话框。

本例中,可单击“Post Hoc”按钮,在弹出的对话框中选择“LSD”、“Tukey”和“Duncan”多重比较方法,按“Continue”按钮回到主对话框中,其他使用默认值。最后按“OK”按钮,执行该过程,则在“OUTPUT”窗口中有如下主要运行结果

ANOVA

疗效天数

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	47.467	4	11.867	5.545	.002
Within Groups	53.500	25	2.140		
Total	100.967	29			

若在主对话框中单击“Paste”按钮,该过程的命令程序出现“Syntax”窗口,单击“Syntax”窗口中的“Run”按钮也可执行该过程。

2. 单因素完全随机不等重复试验的方差分析

在单因素完全随机试验中,设试验因素 A 共有 a 个处理 $A_1, A_2, \dots, A_a$, 每个处理分别重复 $r_1, r_2, \dots, r_a$, 它们完全不等或不完全相等,即试验的每个水平安排的试验单元数目不同,这种不等重复设计的试验称为不平衡设计试验。

疗效天数					
	因素	N	Subset for alpha = .05		
			1	2	3
Tukey HSD ^a	3	6	4.00		
	2	6	5.00		
	4	6	5.17		
	5	6	6.33	6.33	
	1	6		7.67	
	Sig.		.072	.524	
Duncan ^a	3	6	4.00		
	2	6	5.00	5.00	
	4	6	5.17	5.17	
	5	6		6.33	6.33
	1	6			7.67
	Sig.		.203	.147	.127

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 6.000.

例 2 用四种不同饲料喂猪,共 19 头分为四组,每组用一种饲料,喂养一段时间后称重,猪体重增加数据见下表,比较四种饲料对猪体重增加的作用有无不同。

饲 料	A	B	C	D
增重	133.8	151.2	193.4	225.8
	125.3	149.0	185.3	224.6
	143.1	162.7	182.8	220.4
	128.9	143.8	188.5	212.3
	135.7	153.5	198.6	

(1) 在 SAS 软件中仍然可用 ANOVA 过程进行不等重复的单因素完全随机试验的方差分析,ANOVA 过程的分析程序如下

```
DATA ex02;
INPUT cat x@@;
CARDS;
1 133.8 2 151.2 3 193.4 4 225.8
1 125.3 2 149.0 3 185.3 4 220.4
```

```

1 143.1 2 162.7 3 182.8 4 220.4
1 128.9 2 143.8 3 188.5 4 212.3
1 135.7 2 153.5 3 198.6
;
PROC ANOVA;
CLASS cat;
MODEL x = cat;
MEANS cat/duncan;
RUN;

```

执行该程序后可得如下主要结果

```

Analysis of Variance Procedure
Class Level Information
Class Levels Values
CAT 4 1 2 3 4
Number of observations in data set = 19
Analysis of Variance Procedure
Dependent Variable: X

```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	20128.35839474	6709.45279825	158.93	0.0001
Error	15	633.25950000	42.21730000		
Corrected Total	18	20761.61789474			

R-Square	C. V.	Root MSE	X Mean
0.969499	3.793277	6.49748413	171.28947368

Source	DF	Anova SS	Mean Square	F Value	Pr > F
CAT	3	20128.35839474	6709.45279825	158.93	0.0001

(2) 在 SPSS 软件中进行不等重复的单因素完全随机试验的方差分析同样可用 One-Way ANOVA 过程. 本例用 One-Way ANOVA 过程进行方差分析的步骤为:

- 在数据窗口中建立数据文件, 定义两个数值型变量 fact(标签为“饲料”)、x(标签为“增重”), 并输入相应数据. 变量 fact 格式为 F4.0, 取值 1, 2, 3, 4 分别代表四种不同饲料; 变量 x 格式为 F8.2, 其值为猪增重.

- 在系统的“Statistics”菜单中, 选择“Compare Means”子菜单项下的“One-Way ANOVA”过程, 打开主对话框. 然后在该对话框中的“Dependent List”中指定进行方差分析的因变量 x, 在“Factor”中指定因素变量(分类变量)fact.

· 单击“Post Hoc”按钮,在弹出的对话框中选择“LSD”、“Duncan”和“Dunnett”的双尾检验多重比较方法,显著水平输入 0.01. 按“Continue”按钮回到主对话框中,再单击“Options”按钮,选择“Homogeneity-of-Variance”方差同质检验,并选择“Means Plot”复选框,其他使用默认值,最后按“OK”按钮执行该过程. 也可在主对话框中单击“Paste”按钮,激活“Syntax”窗口,单击“Syntax”窗口中的“Run”按钮执行该过程. 这时在“OUTPUT”窗口中有如下运行结果

ANOVA

增重

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	20538.698	3	6846.233	157.467	.000
Within Groups	652.159	15	43.477		
Total	21190.858	18			

增重

	饲料	N	Subset for alpha = .01			
			1	2	3	4
Duncan ^{a,b}	1	5	133.3600			
	2	5		152.0400		
	3	5			189.7200	
	4	4				220.7750
	Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 4.706.

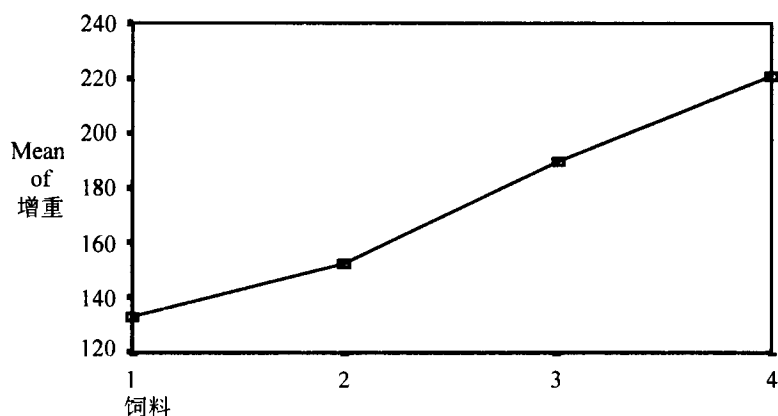
b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

二、单因素随机区组试验的方差分析

由于单因素完全随机试验要求试验单元之间没有系统误差,因而仅在一些实验室或试验空间不大的田间试验中适用. 如果试验单元之间存在系统误差,就必须采用区组化技术,进行单因素随机区组设计,用区组因素把系统误差效应分离出来,以保证试验的精确度和准确度.

1. 单因素随机完全区组试验的方差分析

单因素随机完全区组试验要求每个区组内包含因素的所有处理,故每个区



组的容量正好是处理水平数. 设因素 A 有 a 个处理, 每个处理均重复 r 次, 共需 ar 个试验单元. 把 ar 个试验单元随机划分成相对均匀的 r 个区组, 使每个处理在该区组中均能出现, 于是完成单因素随机完全区组试验设计.

例 3 用 A_1, A_2, A_3, A_4 四种饲料分别饲养四群雏鸡, 每群 100 只, 测定其增重. 试验分别在四个鸡场进行. 试分析四种饲料效应的差异显著性, 试验结果如下表所示

	B_1	B_2	B_3	B_4
A_1	1133	1182	1319	1104
A_2	1102	1186	1338	1092
A_3	1182	1199	1336	1125
A_4	1184	1259	1328	1139

(1) 利用 SAS 的 ANOVA 过程分析程序如下

```

DATA EX03;
DO A = 1 TO 4;
  DO B = 1 TO 4;
    INPUT Y@; OUTPUT;
  END;
END;
CARDS;
1133 1182 1319 1104
1102 1186 1338 1092
1182 1199 1336 1125
1184 1259 1328 1139
;
PROC ANOVA;

```

```

CLASS A B;
MODEL Y = A B;
MEANS A/DUNCAN;

RUN;

```

程序中 A 表示饲料种类, B 表示区组(鸡场), CLASS 语句指定 A 和 B 为方差分析中的因素变量. MEANS 语句要求对各种饲料增重的均值作 DUNCAN 法差异显著性检验. 程序的执行结果如下

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	F Value	Pr > F
Model	6	112929.500000	41.80	0.0001
Error	9	4052.500000		
Corrected Total	15	116982.000000		
	R-Square	C. V.	Y Mean	
	0.965358	1.767576	1200.50000	

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Anova SS	F Value	Pr > F
A	3	6104.000000	4.52	0.0340
B	3	106825.500000	79.08	0.0001

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: Y

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha = 0.05 df = 9 MSE = 450.2778

Number of Means 2 3 4

Critical Range 33.94 35.43 36.28

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	A
A	1227.50	4	4
A			
B	1210.50	4	3
B			

B	1184.50	4	1
B			
B	1179.50	4	2

上述分析结果前面是方差分析表,其中变量 A(饲料种类)的 F 值为 4.52,显著性水平达 0.0340,区组变量 B 的 F 值为 79.08,显著性水平达 0.0001;最后是 DUNCAN'S 多重比较的结果。

(2) 在 SPSS 中可用 GLM-General Factorial 过程进行单因素随机完全区组试验的方差分析。GLM-General 过程使用的一般步骤如下:

- 建立数据文件、定义变量并输入数据。
- 输入因变量、固定因素变量(因素)、随机因素变量(区组),如果有协变量和权重变量,还应在相应框中输入。
- 若不再选择其他可选项,单击主对话框中的“OK”按钮,就可按系统默认值对数据窗口中的数据进行方差分析,在输出窗口中即可得到分析结果。

GLM-General Factorial 过程的选项有六类,包括“Model”模型定义、“Contrasts”设置对照、“Plots”作图、“Post Hoc”多重比较、“Save”保存结果为新变量和“Options”一般选项。各类选项中包含的选项如下:

① “Model”模型定义选项

1) Specify Model 用于指定模型方式,包括 Full Factorial 全因子模式和 Custom 自定义模式,默认为全因子模式,即对所有可能因子及可能的交互作用进行分析。

2) Sum of 指定四种求和类型中的一种,默认为第三类。

3) Include Intercept in Model 选择是否包含截距。

② “Contrasts”设置对照选项

1) Factor 选择有对照的因子。

2) Change Contrast 改变对照方式(共六种方式)及参考对照(开头或最后)。

③ “Plots”作图选项用于选择作图竖轴变量及分隔线与分隔图变量。

④ “Post Hoc”多重比较选项对所选择的固定效应因子在方差相等时可选择 14 种多重比较方法,方差不等时可选择 4 种多重比较方法。

⑤ “Save”选项可将一些结果如不同类型残差,不同类型预测值及诊断值等保存在数据区作为新变量,以便作为其他运算数据。

⑥ “Options”用于选择要显示均值的因子及各种统计量。

本例在指定模型中只选两个主效应,不含截距;多重比较选择 LSD 和 DUNCAN 两种,其他使用默认值。主要结果如下

Tests of Between-Subjects Effects

Dependent Variable: 鸡增重

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
B	Hypothesis	106850.250	3	35616.750	79.724	.000
	Error	4020.750	9	446.750 ^a		
A	Hypothesis	6040.750	3	2013.583	4.507	.034
	Error	4020.750	9	446.750 ^a		

a MS(Error)

Expected Mean Squares^{a,b}

Source	Variance Component		
	Var(B)	Var(Error)	Quadratic Term
B	4.000	1.000	
A	4.000	1.000	A
Error	.000	1.000	

a For each source, the expected mean square equals the sum of the coefficients in the cells times the variance components, plus a quadratic term involving effects in the Quadratic Term cell.

b Expected Mean Squares are based on the Type III Sums of Squares.

2. 单因素 BIB 试验的方差分析

若区组容量小于处理水平数,即一个区组内不能完全安排所有处理水平时,采用的不完全区组设计称为单因素 BIB 试验设计。

例 4 用同窝、同性别的 4 头仔猪作为一个区组,共五窝,进行五种饲料比较研究的 BIB 设计及其方差分析,仔猪增重见下表

区 组	处 理				
	A ₁	A ₂	A ₃	A ₄	A ₅
B ₁	35	34	40	37	
B ₂	38	30	52		18
B ₃	39	13		33	29
B ₄	13		49	30	1
B ₅		25	39	41	4

(1) 利用 SAS 的 ANOVA 过程分析程序如下

```
DATA EX04;
```

```
INPUT A B Y@@;
```

```

CARDS;
1 1 35 1 2 38 1 3 39 1 4 13
2 1 34 2 2 30 2 3 13 2 5 25
3 1 40 3 2 52 3 4 49 3 5 39
4 1 37 4 3 33 4 4 30 4 5 41
5 2 18 5 3 29 5 4 1 5 5 4
;
PROC ANOVA;
    CLASS A B;
    MODEL Y = A B;
    MEANS A/DUNCAN T;
RUN;

```

程序的主要执行结果如下

Analysis of Variance Procedure

Class Level Information

Class	Levels	Values
A	5	1 2 3 4 5
B	5	1 2 3 4 5

Number of observations in data set = 20

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Sum of	Mean		
		Squares	Square	F Value	Pr > F
Model	8	2725.000000	340.6250000	4.02	0.0181
Error	11	931.0000000	84.6363636		
Corrected Total	19	3656.0000000			

R-Square	C. V.	Root MSE	Y Mean
0.745350	30.66601	9.1998024	30.000000

Source	DF	Anova SS	Mean Square	F Value	Pr > F
A	4	2253.5000000	563.3750000	6.66	0.0056
B	4	471.5000000	117.8750000	1.39	0.2992

Analysis of Variance Procedure

T tests (LSD) for variable; Y

NOTE: This test controls the type I comparisonwise error rate not the

experimentwise error rate.

Alpha = 0.05 df = 11 MSE = 84.63636

Critical Value of T = 2.20

Least Significant Difference = 14.318

Means with the same letter are not significantly different.

T	Grouping	Mean	N	A
	A	45.000	4	3
	A			
B	A	35.250	4	4
B	A			
B	A	31.250	4	1
B				
B	C	25.500	4	2
	C			
	C	13.000	4	5

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: Y

NOTE: This test controls the type I comparisonwise error rate, not

the experimentwise error rate

Alpha = 0.05 df = 11 MSE = 84.63636

Number of Means 2 3 4 5

Critical Range 14.32 14.98 15.37 15.63

Means with the same letter are not significantly different.

Duncan Grouping		Mean	N	A
	A	45.000	4	3
	A			
B	A	35.250	4	4
B	A			
B	A	31.250	4	1
B				
B	C	25.500	4	2
	C			

C 13.000 4 5

(2) 在 SPSS 中, 同样可用 GLM-General Factorial 过程进行单因素 BIB 试验的方差分析. 首先定义处理变量 A、区组变量 B 和仔猪增重变量 X, 然后录入数据, 执行“GLM-General Factorial”过程, 打开相应主对话框. 在主对话框中选择 X 为因变量, A 为固定因素变量, B 为随机因素变量, 多重比较选择 LSD 和 DUNCAN 方法, 按“OK”按钮得如下主要结果

Tests of Between-Subjects Effects

Dependent Variable: 增重

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
B	Hypothesis	477.300	4	119.325	1.419	.291
	Error	925.200	11	84.109 ^a		
A	Hypothesis	2259.300	4	564.825	6.715	.005
	Error	925.200	11	84.109 ^a		

a MS(Error)

Expected Mean Squares^{a,b}

	Variance Component		
Source	Var(B)	Var(Error)	Quadratic Term
B	3.750	1.000	
A	.000	1.000	A
Error	.000	1.000	

a For each source, the expected mean square equals the sum of the coefficients in the cells times the variance components, plus a quadratic term involving effects in the Quadratic Term cell.

b Expected Mean Squares are based on the Type III Sums of Squares.

三、单因素系统分组设计试验的方差分析

系统分组设计是把试验空间逐级向低次方向划分的试验设计方法, 也称为巢式设计. 若为单因素, 共有 a 个处理, 把试验空间分为 a 个组, 每组随机安排一个处理, 再把每个组分为 m 个亚组, 每个亚组设置 r 个重复试验单元, 这样的设计称为单因素二级系统分组设计试验.

1. 次级样本相等的单因素二级系统分组设计试验的方差分析

例 5 研究 A_1, A_2, A_3, A_4 四种水生蔬菜对砷污染的“抗性”, 每种蔬菜种 3 盆, 每盆 5 株. 生长期间施用一次有机砷农药, 在收获时对每一株植物的砷含量作一次分析, 得下表. 试作方差分析.

蔬菜品种	A ₁			A ₂			A ₃			A ₄		
盆 号	A ₁₁	A ₁₂	A ₁₃	A ₂₁	A ₂₂	A ₂₃	A ₃₁	A ₃₂	A ₃₃	A ₄₁	A ₄₂	A ₄₃
含 砷 量	0.7	0.9	0.8	1.2	1.1	1.5	0.6	0.5	0.6	4.2	2.9	3.6
	0.6	0.9	0.6	1.4	0.9	1.4	0.6	0.8	1.2	3.7	3.5	3.5
	0.9	0.7	0.9	1.6	1.3	0.9	0.8	0.9	0.8	2.9	3.8	4.0
	0.5	1.1	1.0	1.2	1.2	1.3	0.9	1.0	0.9	3.5	3.1	3.3
	0.6	0.7	0.8	1.5	1.0	1.6	0.7	0.6	1.0	3.6	3.5	3.7

(1) 在 SAS 中的分析程序如下

```

DATA EX05;
DO A = 1 TO 4;
  DO B = 1 TO 3;
    DO C = 1 TO 5;
      INPUT Y@;OUTPUT;
    END;
  END;
END;
CARDS;
0.7 0.6 0.9 0.5 0.6 0.9 0.9 0.7 1.1 0.7 0.8 0.6 0.9 1.0 0.8
1.2 1.4 1.6 1.2 1.5 1.1 0.9 1.3 1.2 1.0 1.5 1.4 0.9 1.3 1.6
0.6 0.6 0.8 0.9 0.7 0.5 0.8 0.9 1.0 0.6 0.6 1.2 0.8 0.9 1.0
4.2 3.7 2.9 3.5 3.6 2.9 3.5 3.8 3.1 3.5 3.6 3.5 4.0 3.3 3.7
;
PROC ANOVA;
  CLASS A B;
  MODEL Y = A B(A);
  MEANS A/DUNCAN;
RUN;

```

程序的 MODEL 语句中用 B(A)表示 B 嵌套于 A 中,执行结果如下

Analysis of Variance Procedure

Class Level Information

Class	Levels	Values
A	4	1 2 3 4
B	3	1 2 3

Number of observations in data set = 60

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	F Value	Pr > F
Model	11	77.36583333	116.41	0.0001
Error	48	2.90000000		
Corrected Total	59	80.26583333		

R-Square	C. V.	Y Mean
0.963870	15.44281	1.59166667

Source	DF	Anova SS	F Value	Pr > F
A	3	76.73916667	423.39	0.0001
B(A)	8	0.62666667	1.30	0.2679

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: Y

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha = 0.05 df = 48 MSE = 0.060417

Number of Means 2 3 4

Critical Range .1805 .1898 .1959

Analysis of Variance Procedure

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	A
A	3.52000	15	4
B	1.27333	15	2
C	0.79333	15	3
C			
C	0.78000	15	1

上面分析结果中,A对应的是不同品种间的方差;B(A)对应的是同一品种中盆间的方差.单因素二级系统分组设计试验的方差分析,还可以用嵌套设计的方差分析过程 NESTED 进行分析处理,对此就不介绍了.

(2) 在 SPSS 中,可用 GLM-General Factorial 过程进行单因素二级系统分组设计试验的方差分析.首先在数据窗口定义蔬菜品种变量 A、盆号 B 和重复 R,砷含量为 X,然后录入数据,执行“GLM- General Factorial”过程.在主对话框中选择 X 为因变量,A 为固定因素,B 与 R 为随机因素;按“Models”按钮选择模型,选效应 A、效应 A * B,使用第一类求平方和方法,包含回归截距;按“Post Hoc”按钮选择 LSD 和 DUNCAN 多重比较方法,其他都用默认值,按“OK”按钮

得如下主要结果

Tests of Between-Subjects Effects

Dependent Variable: 含砷量

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	152.004	1	152.004	1940.479	0.00
	Error	.627 ^a	8	7.833E-02		
A	Hypothesis	76.739	3	25.580	326.550	.000
	Error	.627 ^a	8	7.833E-02		
A * B	Hypothesis	.627	8	7.833E-02	1.297	.268
	Error	2.900 ^b	48	6.042E-02		

a MS(A * B) b MS(Error)

2. 次级样本不等的单因素二级系统分组设计试验的方差分析

例 6 六头公猪所配二十六头母猪,其女儿平均窝产活仔数资料如下表. 试进行方差分析.

(1) 在 SAS 中的分析程序如下

	1	2	3	4	5	6	7	8	9	10	11	1	2
1	10.2	9.0	6.8	9.5	10.5	8.0	12.0	10.0	12.7	9.0	9.5	13.5	6.0
2	9.4	9.0	9.3			6.0				8.3		8.5	9.3
3	9.5	9.0	6.7										
4	6.7												

	1	2	1	2	3	4	5	1	2	3	4	1	2
1	8.0	12.0	12.4	9.8	10.7	12.0	9.7	9.0	12.0	6.0	8.8	7.5	7.5
2			8.8		10.5				10.3	8.5			
3			10.4		9.8				10.0	12.5			
4			11.7										

DATA EX06;

INPUT A B Y@@@;

CARDS;

1 1 10.2 1 1 9.4 1 1 9.5 1 1 6.7 1 2 9.1 2 9.1 2 9

1 3 6.8 1 3 9.3 1 3 6.7 1 4 9.5 1 5 10.5 1 6 8.1 6 6

1 7 12.1 8 10.1 9 12.7 1 10 9.1 10 8.3 1 11 9.5

2 12 13.5 2 12 8.5 2 13 6.2 13 9.3 3 14 8.3 15 12

4 16 12.4 4 16 8.8 4 16 10.4 4 16 11.7 4 17 9.8

```

4 18 10.7 4 18 10.5 4 18 9.8 4 19 12 4 20 9.7
5 21 9 5 22 12 5 22 10.3 5 22 10 5 23 6 5 23 8.5
5 23 12.5 5 24 8.8 6 25 7.5 6 26 7.5
;
PROC ANOVA;
  CLASS A B;
  MODEL Y = A B(A);
  MEANS A/DUNCAN;
RUN;

```

程序执行主要结果如下

Analysis of Variance Procedure

Class Level Information

Class	Levels	Values
A	6	1 2 3 4 5 6
B	26	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

Number of observations in data set = 46

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	25	94.09851449	3.76394058	1.19	0.3517
Error	20	63.46083333	3.17304167		
Corrected Total	45	157.55934783			

	R-Square	C. V.	Root MSE	Y Mean
	0.597226	18.78065	1.7813034	9.4847826

Source	DF	Anova SS	Mean Square	F Value	Pr > F
A	5	24.38759783	4.87751957	1.54	0.2232
B(A)	20	69.71091667	3.48554583	1.10	0.4178

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: Y

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha = 0.05 df = 20 MSE = 3.173042

WARNING: Cell sizes are not equal.

Harmonic Mean of cell sizes = 3.934426
 Number of Means 2 3 4 5 6
 Critical Range 2.649 2.781 2.864 2.923 2.966
 Means with the same letter are not significantly different.

Duncan Grouping		Mean	N	A
	A	10.580	10	4
	A			
B	A	10.000	2	3
B	A			
B	A	9.638	8	5
B	A			
B	A	9.325	4	2
B	A			
B	A	9.055	20	1
B	A			
B		7.500	2	6
B				

(2) 在 SPSS 中,可用 GLM-General Factorial 过程进行次级样本不等的单因素二级系统分组设计试验的方差分析.首先在数据窗口定义公猪 A、母猪 B 和女儿 C,活仔数为 X,然后录入数据,执行“GLM-General Factorial”过程,在主对话框中选择 X 为因变量,A 为固定因素,B 与 R 为随机因素;按“Models”按钮选择模型,选效应 A、效应 A * B,使用第一类求平方和方法,包含回归截距;按“Post Hoc”按钮选择 LSD、Tukey 和 DUNCAN 多重比较方法,其他都用默认值,按“OK”按钮得如下主要结果

Tests of Between-Subjects Effects

Dependent Variable: 活仔数

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	4138.211	1	4138.211	1149.229	.000
	Error	38.682 ^a	10.743	3.601		
A	Hypothesis	24.388	5	4.878	1.388	.278
	Error	59.705 ^b	16.991	3.514		
A * B	Hypothesis	69.711	20	3.486	1.098	.418
	Error	63.461 ^c	20	3.173		

a 1.369 MS(A * B) - .369 MS(Error)

b 1.091 MS(A * B) - 9.063E-02 MS(Error)

c MS(Error)

Expected Mean Squares ^{a,b}			
Source	Variance Component		
	Var(A * B)	Var(Error)	Quadratic Term
Intercept	2.348	1.000	Intercept, A
A	1.870	1.000	A
A * B	1.715	1.000	
Error	.000	1.000	

- a For each source, the expected mean square equals the sum of the coefficients in the cells times the variance components, plus a quadratic term involving effects in the Quadratic Term cell.
- b Expected Mean Squares are based on the Type I Sums of Squares.

四、拉丁方试验的方差分析

将不同符号排成行列,使每个符号在每一行、列仅出现一次的方阵,叫做拉丁方.拉丁方设计较随机区组设计更进一步,它可以从行和列两个方向进行局部控制,使行列皆成区组,以剔除两个方向的系统误差,因而有较好的准确度与精确度.

1. 单个拉丁方试验的方差分析

例7 用5头奶牛检验5种饲料A,B,C,D,E对泌乳量的影响,由于5头牛的泌乳力不太一致,且泌乳期有差异,因而用拉丁方设计试验.其中把泌乳期分成5个适当的阶段,试验结果如下表所示.

产乳量		泌 乳 期				
		1	2	3	4	5
牛号	1	E 300	A 320	B 390	C 390	D 380
	2	D 420	C 390	E 280	B 370	A 270
	3	B 350	E 360	D 400	A 260	C 400
	4	A 280	D 400	C 390	E 280	B 370
	5	C 400	B 380	A 350	D 430	E 320

(1) 在SAS中的分析程序如下

```
DATA EX07;
DO L = 1 TO 5;
DO K = 1 TO 5;
INPUT V$ P@@@; OUTPUT;
END;
END;
CARDS;
```

```

E 300 A 320 B 390 C 390 D 380
D 420 C 390 E 280 B 370 A 270
B 350 E 360 D 400 A 260 C 400
A 280 B 400 C 390 E 280 B 370
C 400 B 380 A 350 D 430 E 320

```

```
;
```

```
PROC ANOVA;
```

```
CLASS L K V;
```

```
MODEL P = V;
```

```
MEANS V/LSD DUNCAN;
```

```
RUN;
```

程序中,变量 V 表示饲料、K 代表泌乳期、L 代表奶牛、P 表示产乳量。程序执行结果如下

Analysis of Variance Procedure

Dependent Variable: P

Source	DF	Sum of Squares	F Value	Pr > F
Model	12	55872.0000000	7.60	0.0007
Error	12	7352.0000000		
Corrected Total	24	63224.0000000		

R-Square	C.V.	P Mean
0.883715	6.968498	355.200000

Source	DF	Anova SS	F Value	Pr > F
L	4	3224.0000000	1.32	0.3193
K	4	2144.0000000	0.87	0.5071
V	4	50504.0000000	20.61	0.0001

Analysis of Variance Procedure

T tests (LSD) for variable: P

NOTE: This test controls the type I comparisonwise error rate
not the experimentwise error rate.

Alpha = 0.05 df = 12 MSE = 612.6667

Critical Value of T = 2.18

Least Significant Difference = 34.108

Means with the same letter are not significantly different.

T Grouping	Mean	N	V
------------	------	---	---

A	406.00	5	D
A	394.00	5	C
A	372.00	5	B
B	308.00	5	E
B	296.00	5	A

Duncan's Multiple Range Test for variable: P

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha = 0.05 df = 12 MSE = 612.6667

Number of Means 2 3 4 5

Critical Range 34.11 35.70 36.67 37.31

Duncan Grouping	Mean	N	V
A	406.00	5	D
A	394.00	5	C
A	372.00	5	B
B	308.00	5	E
B	296.00	5	A

(2) 在 SPSS 软件中可用 GLM-General Factorial 过程进行拉丁方试验设计的方差分析. 在数据窗口定义变量 A(行), B(列), C(因素)和 X(产乳量), 并输入数据. 执行“GLM-General Factorial”过程, 打开相应主对话框. 在主对话框中选择 X 为因变量, C 为固定因素变量, A, B 为随机因素变量, 多重比较选择 DUNCAN 方法, 按“OK”按钮得如下主要结果

Tests of Between-Subjects Effects

Dependent Variable: 产乳量

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
A	Hypothesis	3224.000	4	806.000	1.316	.319
	Error	7352.000	12	612.667 ^a		
B	Hypothesis	2144.000	4	536.000	.875	.507
	Error	7352.000	12	612.667 ^a		
C	Hypothesis	50504.000	4	12626.000	20.608	.000
	Error	7352.000	12	612.667 ^a		

a MS(Error)

Expected Mean Squares ^{a,b}				
Source	Variance Component		Var(Error)	Quadratic Term
	Var(A)	Var(B)		
A	5.000	.000	1.000	
B	.000	5.000	1.000	
C	.000	.000	1.000	C
Error	.000	.000	1.000	

- a For each source, the expected mean square equals the sum of the coefficients in the cells times the variance components, plus a quadratic term involving effects in the Quadratic Term cell.
- b Expected Mean Squares are based on the Type III Sums of Squares.

产 乳 量

	因 素	N	Subset	
			1	2
Duncan ^{a,b}	1	5	296.0000	
	5	5	308.0000	
	2	5		372.0000
	3	5		394.0000
	4	5		406.0000
	Sig.		.458	.060

Means for groups in homogeneous subsets are displayed. Based on Type III Sum of Squares The error term is Mean Square(Error) = 612.667.

a Uses Harmonic Mean Sample Size = 5.000.

b Alpha = .05.

2. 重复拉丁方试验的方差分析

例 8 A, B, C, D 四个棉花品种, 在 U_1 , U_2 两块地各进行一次 4×4 拉丁方试验, U_1 为麦行间套种棉花, U_2 为麦后播种棉花, 播期相差 45 天, 计产面积 49 m^2 , 试验结果如下表所示, 试进行方差分析.

棉花 产量	U_1				U_2			
	1	2	3	4	1	2	3	4
1	C 5.9	B 4.9	A 6.1	D 7.6	A 4.2	C 3.6	B 3.6	D 4.8
2	D 7.9	C 5.6	B 5.8	A 7.8	B 3.0	D 4.0	C 3.4	A 3.9
3	B 5.0	A 7.6	D 8.2	C 6.5	D 4.6	B 3.0	A 3.6	C 3.5
4	A 7.3	D 7.2	C 6.0	B 4.9	C 3.4	A 4.1	D 4.4	B 3.1

(1) 在 SAS 中的分析程序如下

```

DATA EX08;
DO I = 1 TO 2;
  DO L = 1 TO 5;
    DO K = 1 TO 5;
      INPUT V$ P@@@; OUTPUT;
    END;
  END;
END;
CARDS;
C 5.9  B 4.9  A 6.1  D 7.6
D 7.9  C 5.6  B 5.8  A 7.8
B 5.0  A 7.6  D 8.2  C 6.5
A 7.3  D 7.2  C 6.0  B 4.9
A 4.2  C 3.6  B 3.6  D 4.8
B 3.0  D 4.0  C 3.4  A 3.9
D 4.6  B 3.0  A 3.6  C 3.5
C 3.4  A 4.1  D 4.4  B 3.1
;
PROC ANOVA;
  CLASS I L K V;
  MODEL P = V;
  MEANS V/LSD DUNCAN;
RUN;

```

程序执行结果为

Analysis of Variance Procedure Class Level Information

Class	Levels	Values
I	2	1 2
L	5	1 2 3 4 5
K	5	1 2 3 4 5
V	4	A B C D

Number of observations in data set = 32

Analysis of Variance Procedure

Dependent Variable: P		Sum of	Mean		
Source	DF	Squares	Square	F Value	Pr > F
Model	3	17.63593750	5.87864583	2.45	0.0845
Error	28	67.24125000	2.40147321		

Corrected Total	31	84.87718750			
	R-Square	C. V.	Root MSE	P Mean	
	0.207782	30.14553	1.5496687	5.1406250	
Source	DF	Anova SS	Mean Square	F Value	Pr > F
V	3	17.63593750	5.87864583	2.45	0.0845

Analysis of Variance Procedure

T tests (LSD) for variable: P

NOTE: This test controls the type I comparisonwise error rate not the experimentwise error rate.

Alpha = 0.05 df = 28 MSE = 2.401473 Critical Value of T = 2.05

Least Significant Difference = 1.5872

Means with the same letter are not significantly different.

T Grouping	Mean	N	V
A	6.0875	8	D
A			
B	5.5750	8	A
B			
B	4.7375	8	C
B			
B	4.1625	8	B

Analysis of Variance Procedure

Duncan's Multiple Range Test for variable: P

NOTE: This test controls the type I comparisonwise error rate, not the experimentwise error rate

Alpha = 0.05 df = 28 MSE = 2.401473

Number of Means 2 3 4

Critical Range 1.587 1.668 1.720

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	V
A	6.0875	8	D
B	5.5750	8	A
B			

B	A	4.7375	8	C
B		4.1625	8	B

(2) 在 SPSS 软件中仍用 GLM-General Factorial 过程进行重复拉丁方试验设计的方差分析. 在数据窗口定义变量 A(行), B(列), C(因素), D(重复) 和 X(棉花产量), 并输入数据. 执行“GLM-General Factorial”过程, 打开主对话框. 在主对话框中选择 X 为因变量, C 为固定因素变量, A, B, D 为随机因素变量, 多重比较选择 DUNCAN 方法, 按“OK”按钮得如下结果

Tests of Between-Subjects Effects

Dependent Variable: 棉花产量

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
C	Hypothesis	17.636	3	5.879	7.494	.066
	Error	2.353	3	.784 ^a		
A	Hypothesis	.193	3	6.448E-02	.319	.811
	Error	3.638	18	.202 ^b		
B	Hypothesis	.281	3	9.365E-02	.463	.711
	Error	3.638	18	.202 ^b		
D	Hypothesis	60.775	1	60.775	77.472	.003
	Error	2.353	3	.784 ^a		
C * D	Hypothesis	2.353	3	.784	3.881	.027
	Error	3.638	18	.202 ^b		

a MS(C * D)

b MS(Error)

Expected Mean Squares^{a, b}

Source	Variance Component					Quadratic Term
	Var(A)	Var(B)	Var(D)	Var(C * D)	Var(Error)	
C	.000	.000	.000	4.000	1.000	C
A	8.000	.000	.000	.000	1.000	
B	.000	8.000	.000	.000	1.000	
D	.000	.000	16.000	4.000	1.000	
C * D	.000	.000	.000	4.000	1.000	
Error	.000	.000	.000	.000	1.000	

a For each source, the expected mean square equals the sum of the coefficients in the cells times

the variance components, plus a quadratic term involving effects in the Quadratic Term cell.

b Expected Mean Squares are based on the Type III Sums of Squares.

棉花产量

	品 种	N	Subset			
			1	2	3	4
Duncan ^{a,b}	2	8	4.1625			
	3	8		4.7375		
	1	8			5.5750	
	4	8				6.0875
	Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed. Based on Type III Sum of Squares The error term is Mean Square(Error) = .202.

a Uses Harmonic Mean Sample Size = 8.000.

b Alpha = .05.

五、协方差分析

1. 有协变量试验的方差分析

要使所有试验单元或区组内的试验单元的条件完全一致,有时候是非常困难的.如果试验指标为 Y ,而影响试验单元不一致的因素 X 可以测量,则 X 为方差分析中的协变量,利用有协变量试验的方差分析可将其效应剔除.

例 9 为比较三种肥料对苹果树的增产效果,选了 24 株同龄苹果树,随机分为三组,并在第一年记下各树的产量 X .第二年每组随机施用一种肥料,并记录各组的产量 Y ,结果见下表,试进行方差分析.

肥 料		观 测 值							
A_1	X	47	58	53	46	49	56	54	44
	Y	54	66	63	51	56	66	61	50
A_2	X	52	53	64	58	59	61	63	66
	Y	54	53	67	62	62	63	64	69
A_3	X	44	48	46	50	59	57	58	53
	Y	52	58	54	61	70	64	69	66

(1) 在 SAS 中的分析程序如下

```
DATA EX09;
```

```
DO K = 1 TO 8;
```

```

DO A = 1 TO 3;
    INPUT X Y@; OUTPUT;
END;
END;
CARDS;
47 54 52 54 44 52
58 66 53 53 48 58
53 63 64 67 46 54
46 51 58 62 50 61
49 56 59 62 59 70
56 66 61 63 57 64
54 61 63 64 58 69
44 50 66 69 53 66
;
PROC GLM;
    CLASS A;
    MODEL Y = A X/SOLUTION;
    MEANS A/LSD;
RUN;

```

程序中 A 为分类变量, X 为协变量, Y 为因变量. 程序输出结果如下

General Linear Models Procedure

Class Level Information

Class	Levels	Values
A	3	1 2 3

Number of observations in data set = 24

General Linear Models Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	F Value	Pr > F
Model	3	842.79453688	115.06	0.0001
Error	20	48.83046312		
Corrected Total	23	891.62500000		
		R-Square	C. V.	Y Mean
		0.945234	2.577381	60.6250000

Dependent Variable: Y

Source	DF	Type I SS	F Value	Pr > F
--------	----	-----------	---------	--------

A	2	60.75000000	12.44	0.0003
X	1	782.04453688	320.31	0.0001
Source	DF	Type III SS	F Value	Pr > F
A	2	222.84063820	45.64	0.0001
X	1	782.04453688	320.31	0.0001

General Linear Models Procedure

T tests (LSD) for variable: Y

NOTE: This test controls the type I comparisonwise error rate

not the experimentwise error rate.

Alpha = 0.05 df = 20 MSE = 2.441523

Critical Value of T = 2.09

Least Significant Difference = 1.6297

Means with the same letter are not significantly different.

T Grouping	Mean	N	A
A	61.7500	8	3
A			
A	61.7500	8	2
B	58.3750	8	1

(2) 在 SPSS 中可用 GLM-General Factorial 过程进行有协变量试验的方差分析. 首先定义变量 A(肥料)、R(重复)、X(原产量)和 Y(新产量), 输入数据后, 执行“GLM-General Factorial”过程, 打开主对话框. 在主对话框中选择 Y 为因变量, A 为固定因素变量, R 为随机因素变量, X 为协变量, 多重比较选择 LSD 和 DUNCAN 方法, 按“OK”按钮得主要结果如下

Tests of Between-Subjects Effects

Dependent Variable: 新产量

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	60.750 ^a	2	30.375	.768	.477
Intercept	88209.375	1	88209.375	2229.453	.000
A	60.750	2	30.375	.768	.477
Error	830.875	21	39.565		
Total	89101.000	24			
Corrected Total	891.625	23			

a. R Squared = .068 (Adjusted R Squared = -.021)

2. 随机模型试验的协方差分析

随机模型试验分析中,往往需要进行多指标间的协方差分析,因为这些指标之间是相互影响的.

例 10 下表为玉米品种试验的每区株数(x)和产量(y),试进行协方差分析.

区组	品 种					区组	品 种				
	A	B	C	D	E		A	B	C	D	E
1 _x	10	12	17	14	12	3 _x	6	8	13	17	10
1 _y	18	36	40	21	42	3 _y	14	28	35	24	38
2 _x	8	13	15	14	10	4 _x	8	11	11	15	16
2 _y	17	38	36	23	36	4 _y	15	30	29	20	52

(1) 在 SAS 中的分析程序如下

```
DATA EX10;
  DO K = 1 TO 4;
    DO A = 1 TO 5;
      INPUT X Y@; OUTPUT;
    END;
  END;
CARDS;
10 18 12 36 17 40 14 21 12 42
8 17 13 38 15 36 14 23 10 36
6 14 8 28 13 35 17 24 10 38
8 15 11 30 11 29 15 20 16 32
;

PROC GLM;
  CLASS K A;
  MODEL Y = K A X/SOLUTION;
  MEANS A/LSD;
RUN;
```

程序中 K 为区组, A 为品种, X 为每区株数, 是协变量; Y 为产量, 是因变量. 程序输出主要结果如下

```
General Linear Models Procedure
Class Level Information
Class      Levels      Values
K           4          1 2 3 4
A           5          1 2 3 4 5
```


Number of observations in data set = 20					
General Linear Models Procedure					
Dependent Variable: Y					
		Sum of		Mean	
Source	DF	Squares	Square	F Value	Pr > F
Model	8	1467.7993289	183.4749161	30.12	0.0001
Error	11	67.0006711	6.0909701		
Corrected Total	19	1534.8000000			
	R-Square	C. V.	Root MSE		Y Mean
	0.956346	8.629332	2.4679891		28.600000
Source	DF	Type I SS	Mean Square	F Value	Pr > F
K	3	110.0000000	36.6666667	6.02	0.0111
A	4	1332.8000000	333.2000000	54.70	0.0001
X	1	24.9993289	24.9993289	4.10	0.0677
Source	DF	Type III SS	Mean Square	F Value	Pr > F
K	3	94.4993289	31.4997763	5.17	0.0180
A	4	1146.9465449	286.7366362	47.08	0.0001
X	1	24.9993289	24.9993289	4.10	0.0677
T for H0: Pr > T Std Error of					
Parameter		Estimate	Parameter = 0		Estimate
INTERCEPT		25.69865772 B	6.12	0.0001	4.20088820
K	1	5.68187919 B	3.59	0.0042	1.58170618
	2	4.92953020 B	3.16	0.0092	1.56220227
	3	3.50671141 B	2.16	0.0537	1.62379035
	4	0.00000000 B			
A	1	-18.40939597 B	-8.51	0.0001	2.16347979
	2	-3.35234899 B	-1.89	0.0855	1.77417094
	3	-3.29530201 B	-1.77	0.1039	1.85856799
	4	-16.94295302 B	-8.51	0.0001	1.99129667
	5	0.00000000 B			
X		0.64765101	2.03	0.0677	0.31968341

(2) 在 SPSS 中可用 GLM-General Factorial 过程进行随机模型试验的协方差分析. 首先定义变量 A(品种), R(区组), X(每区株数)和 Y(产量), 输入数据后, 执行“GLM-General Factorial”过程, 打开主对话框. 在主对话框中选择 Y 为因变量, A 为固定因素变量, R 为随机因素变量, X 为协变量. 在“Model”选项中选择 A, R, X 主效应, 平方和使用“Type I”, 包含截距. “Options”中选择显示 A 并进行效应比较, 最后按“OK”按钮得如下结果

Tests of Between-Subjects Effects

Dependent Variable: 产量

Source		Type I Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	17523.200	1	17523.200	1546.165	.000
	Error	34.000	3	11.333 ^a		
A	Hypothesis	1748.800	4	437.200	100.938	.000
	Error	47.645	11	4.331 ^b		
R	Hypothesis	34.000	3	11.333	2.617	.104
	Error	47.645	11	4.331 ^b		
X	Hypothesis	220.355	1	220.355	50.874	.000
	Error	47.645	11	4.331 ^b		

a MS(R)

b MS(Error)

Expected Mean Squares^{a,b}

Source	Variance Component		
	Var(R)	Var(Error)	Quadratic Term
Intercept	5.000	1.000	Intercept, A, X
A	.000	1.000	A, X
R	5.000	1.000	X
X	.000	1.000	X
Error	.000	1.000	

a For each source, the expected mean square equals the sum of the coefficients in the cells times the variance components, plus a quadratic term involving effects in the Quadratic Term cell.

b Expected Mean Squares are based on the Type I Sums of Squares.

Estimates				
Dependent Variable: 产量				
品种	Mean	Std. Error	95 % Confidence Interval	
			Lower Bound	Upper Bound
1	23.691 ^a	1.499	20.393	26.990
2	34.923 ^a	1.075	32.557	37.289
3	31.154 ^a	1.172	28.575	33.734
4	16.232 ^a	1.318	13.331	19.132
5	42.000 ^a	1.041	39.710	44.290

a Evaluated at covariates appeared in the model: 每区株数 = 12.0000.

Pairwise Comparisons						
Dependent Variable: 产量						
(I) 品种	(J) 品种	Mean Difference (I-J)	Std. Error	Sig. ^a	95 % Confidence Interval for Difference	
					Lower Bound	Upper Bound
1	2	-11.232 [*]	1.679	.000	-14.927	-7.536
	3	-7.463 [*]	2.187	.006	-12.276	-2.650
	4	7.460 [*]	2.393	.010	2.193	12.727
	5	-18.309 [*]	1.824	.000	-22.324	-14.293
2	1	11.232 [*]	1.679	.000	7.536	14.927
	3	3.768 [*]	1.679	.046	7.254E-02	7.464
	4	18.691 [*]	1.824	.000	14.676	22.707
	5	-7.077 [*]	1.496	.001	-10.370	-3.784
3	1	7.463 [*]	2.187	.006	2.650	12.276
	2	-3.768 [*]	1.679	.046	-7.464	-7.254E-02
	4	14.923 [*]	1.496	.000	11.630	18.216
	5	-10.846 [*]	1.567	.000	-14.295	-7.396
4	1	-7.460 [*]	2.393	.010	-12.727	-2.193
	2	-18.691 [*]	1.824	.000	-22.707	-14.676
	3	-14.923 [*]	1.496	.000	-18.216	-11.630
	5	-25.768 [*]	1.679	.000	-29.464	-22.073

续表

(I) 品种	(J) 品种	Mean Difference (I - J)	Std. Error	Sig. ^a	95 % Confidence Interval for Difference	
					Lower Bound	Upper Bound
5	1	18.309 [*]	1.824	.000	14.293	22.324
	2	7.077 [*]	1.496	.001	3.784	10.370
	3	10.846 [*]	1.567	.000	7.396	14.295
	4	25.768 [*]	1.679	.000	22.073	29.464

Based on estimated marginal means

* The mean difference is significant at the .05 level.

a Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Univariate Tests

Dependent Variable: 产量

	Sum of Squares	df	Mean Square	F	Sig.
Contrast	1561.447	4	390.362	90.125	.000
Error	47.645	11	4.331		

The F tests the effect of 品种. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.