

# 第一部分 一般线性与混合线性模型

## 第 1 章 征服一般线性模型

——General Linear Model 菜单详解（上）

## 第 2 章 征服一般线性模型

——General Linear Model 菜单详解（下）

## 第 3 章 混合线性模型入门

——Mixed Model 菜单详解


## 第 1 章 征服一般线性模型


### ——General Linear Model 菜单详解（上）

在我看来，学习统计软件时不去了解所用统计原理和模型，就像是学英语时只背单词而不记语法一样，是永远不可能说出一口流利的英文的。

——张文彤

请注意，本章的标题用了一些修辞手法，一般线性模型可不是三言两语就能说清楚的，因为它包括的内容实在太多了。那么，究竟什么是一般线性模型呢？用最通俗的话讲：方差分析模型和线性回归模型都是分析一个/多个自变量对一个/多个连续性应变量的影响，并且都假设应变量和自变量是线性数量关系，因此它们都算是一般线性模型的特例，凡是和它们沾边的都可以用 GLM 来做。包括成组设计的方差分析（即单因素方差分析）、配伍设计的方差分析（即两因素方差分析）、多因素方差分析、多元方差分析、重复测量的方差分析、协方差分析、多元线性回归分析等等。因此，能真正掌握 GLM 菜单的用法，会使大家的统计分析能力有极大的提高。

 什么？GLM 还可以完成线性回归分析！是的，它的确可以。只不过它作回归分析的功能没有专用的回归分析模块强，所以该功能较少被用到。

 有些书中将 General Linear Models 翻译成广义线性模型，实际上 GLM 模块的功能只限于一般线性模型。广义线性模型对应的英文是 Generalized Linear Models，指的是范围更广、功能更强大的另一大类模型，在 SPSS 中可以编程拟合，但并非 GLM 菜单提供的对话框就可以完成的，在本书的最后一章中对其有基本的介绍。

好了，本章要讲述的菜单名为 General Linear Model，既然一般线性模型的能力如此强大，那么下属的四个过程各自的功能是什么呢？请看：

- ✧ Univariate 过程：四个菜单中的大哥大，当应变变量只有一个时，我们所进行的分析就要用它来完成。显然，它是用的最多的一个。
- ✧ Multivariate 过程：当结果变量（应变变量）不止一个时，当然要用它来分析啦！
- ✧ Repeated Measures 过程：顾名思义，重复测量的数据就要用它来分析。这一点可能要强调一下，用前两个菜单似乎都可以分析出来结果，但在许多情况下该结果是不正确的。在相应章节我会详细讲述。
- ✧ Variance Components 过程：用于对层次数据拟合方差成份模型，它是普通线性模型向随机效应的进一步扩展，是一种可以考察各个层次因素的变异大小，从而为哪些层次上可能存在组内聚集性、如何可能减小数据变异提供信息的统计方法，也就是现在非常热门的多水平模型的最原始形式。

由于 Univariate 过程是最常用的一个过程，本章我们就围绕它来进行讲解，另外三个过程将放在下一章学习。

## 1.1 方差分析模型简介

现实世界中变量间的联系是错综复杂的,比如要研究性别对身高的影响,显然就要考虑到年龄、遗传、营养状况等因素的作用。这时t检验作为单因素分析方法就无能为力了,而方差分析可以在控制其他因素影响的同时研究两者之间的关系,分析的效率更高,适用范围更广。

同时,许多时候各自变量之间还会存在交互作用,如研究催化剂对化学反应的催化能力,如果该催化剂只在某个温度范围内效果最佳,则只单独研究该催化剂的催化作用是没有实际意义的,此时这种交互作用也成为了我们研究的重点,即必须要研究在什么温度条件下该催化剂的催化能力最佳。对交互作用的分析也是方差分析模型的特长。

### 1.1.1 模型入门

在均数的比较一章中我们实际上已经接触到了方差分析的基本思想,这里首先来复习一下单因素方差分析中变异的分解公式:

总变异=处理因素导致的变异+随机变异

实际上,该公式只是为了便于大家理解,标准的单因素方差分析模型如下:

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

其中  $X_{ij}$  表示第  $i$  组的第  $j$  个观察值;  $\mu$  表示总体的平均水平;  $\alpha_i$  表示影响因素在  $i$  水平下对应变量的附加效应,并假设所有  $\alpha_i$  之和应当为 0;  $\varepsilon_{ij}$  为一个服从正态分布  $N(0, \sigma^2)$  的随机变量,代表随机误差。一般情况下,我们做假设检验实际上就是检验各个  $\alpha_i$  是否均为 0,如都为 0,即各组总体均数都相等,则  $X_{ij}$  就会成为服从正态分布  $N(\mu, \sigma^2)$  的一个变量。



有的时候以上模型也被写为  $X_{ij} = \mu_i + \varepsilon_{ij}$ , 此时检验的含义就变成了各组均数  $\mu_i$  是否相同。显然,它和上面的公式是等价的,但这种写法应当更加容易理解。比如在上一章中我们比较三组石棉矿工的用力肺活量有无差别,那么石棉肺患者组中某一位观察对象的用力肺活量数值就等于石棉肺患者组的平均水平再加上一个随机误差项。而如果三组总体均数相同,则它就应当等于总体均数(平均水平)再加上一个随机误差项,实际上就变成了同一个变量分布中的某一点。

多因素方差分析模型只是对以上公式的进一步扩展,以两因素方差分析模型为例,公式如下:

$$X_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

其中  $\alpha_i$ 、 $\beta_j$  分别表示 A 因素  $i$  水平和 B 因素  $j$  水平的附加效应,  $\varepsilon_{ijk}$  仍为随机误差变量。此时如果要说明因素 A 有无影响,就是检验如下假设:

$H_0: \alpha_i = 0$ ,  $H_1$ : 至少有一个  $\alpha_i \neq 0$

如果要说明因素 B 有无影响,就是检验如下假设:

$H_0: \beta_i = 0$ ,  $H_1$ : 至少有一个  $\beta_i \neq 0$

而所说的模型无显著性就是指上面两个  $H_0$  同时成立（均不能被拒绝）。

更为复杂的是如考虑交互作用的情形，模型如下：

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijk}$$

其中  $\alpha_i$ 、 $\beta_j$  分别表示 A 因素 i 水平和 B 因素 j 水平的附加效应。 $\alpha_i\beta_j$  则为两者的交互效应。

可能有的朋友已经看得头痛了，但无论如何，只要记住方差分析的原理即可：根据资料类型以及研究目的，可将总变异分解为两个或多个部分，除一部分代表随机误差的作用外，每个部分的变异可由某因素的作用（或交互作用）来解释，通过比较可能由某因素所致的变异与随机误差的大小，借助 F 分布做出推断，即可了解该因素对结果变量的影响是否存在。

### 1.1.2 常用术语

方差分析中的常用术语有：

1. 因素（Factor）：因素是可能对应变量有影响的变量，一般来说，因素会有不止一个水平，而分析的目的就是考察或比较各个水平对应变量的影响是否相同。例如影响农作物产量的因素有气温、降雨量、日照时间等。在方差分析中，因素的取值范围不能无限，只能有若干个水平，即应当为分类变量。

2. 水平（Level）：因素的不同取值等级称作水平，例如性别有男、女两个水平。需要注意的是有时候水平是人为划分出来的，比如身高被分为高、中、低三个水平。

3. 单元（Cell）：指各因素水平之间的组合，我们所说的方差齐就是指的各个单元间的方差齐。注意有的试验设计并不会给出所有可能的水平组合，如拉丁方设计。

4. 元素（Element）：指用于测量应变量值的最小单位，比如研究石棉矿工用力肺活量，则肺活量是从每一位矿工身上测得，矿工就是试验的元素。一个单元格内可以有多个元素，也可以只有一个，甚至于没有元素。

5. 均衡（Balance）：如果一个实验设计中任一因素各水平在所有单元格中出现的次数相同，且每个单元格内的元素数相同，则该试验是均衡的，否则，就被称为不均衡。不均衡的实验设计在分析时较为复杂。

6. 固定因素（Fixed Factor）与随机因素（Random Factor）：两者都是因素的不同种类，固定因素指的是该因素在样本中所有可能的水平都出现了。换言之，该因素的所有可能水平仅此几种，针对该因素而言，从样本的分析结果中就可以得知所有水平的状况，无需进行外推。比如要研究糖尿病、IGT（糖耐量异常）和正常人的血糖有无差别，则按照糖尿病有无可将所有人分为糖尿病、IGT 和正常人三种，此时该因素就被认为是固定因素。另外，有些人设定的因素，比如下面引例中的 group，它被分为工前、工中、工后这三个值，如果我们所做的检验就是想弄清楚这三个水平对应变量有无影响，不需要外推到其他水平（如工后半小时），则它也可以被认为是固定因素。

和固定因素相对应的是随机因素，它指的是该因素所有可能的取值在样本中没有都出现，或不可能都出现。如下面引例中的 worker，它表示的是工人这个配伍因素，实际上总体中当然不可能只有这 10 个工人，他们只是所有工人的代表而已。因此要用

样本中 10 个工人对应变量的影响情况来推论总体中全体工人对应变量的影响情况, 包括未出现的那些工人, 这不可避免的存在误差(即随机效应), 需要估计该误差的大小, 因此被称为随机因素。又如我们要研究什么温度下催化剂的效果最好, 样本中取了 30、40、50℃三个水平, 如果我們希望在分析结果中能同时外推 35℃、45℃这些水平的情况, 此时温度也是随机因素。

一般来说固定因素和随机因素在分析时应分别指定, 如果将随机因素按固定因素来分析, 则可能得出错误的分析结果。但是如果所有单元格内都至多只有一个元素, 则随机效应无法被估计出来, 此时两种做法的统计分析结论完全相同。

7. 交互作用(Interaction): 如果一个因素的效应大小在另一个因素不同水平下明显不同, 则称为两因素间存在交互作用。当存在交互作用时, 单纯研究某个因素的作用是没有意义的, 必须分另一个因素的不同水平研究该因素的作用大小。

如果所有单元格内都至多只有一个元素, 则交互作用无法测量, 只能不予考虑, 最典型的例子就是配伍设计的方差分析。

### 1.1.3 方差分析模型的适用条件

方差分析并不是万金油, 它也有自己的适用条件(以  $H_0$  假设成立为前提), 具体说有以下几点:

- ✧ 各样本的独立性: 只有各样本为相互独立的随机样本, 才能保证变异的可加性(可分解性)。
- ✧ 正态性: 即所有观察值系从正态总体中抽样得出。
- ✧ 方差齐: 这里所说的方差齐是指假设总的模型无意义时方差齐, 亦即每一个单元格中的方差齐。

道理大家其实都懂, 但做起来就不一定都明白了, 为什么多因素方差分析一般都不提这些条件呢? 首先在以上条件中, 对独立性的要求是最严的, 但它一般都可以保证。其次为正态性和方差齐性, 具体来说就是:

- ✧ 单因素方差分析: 适用条件是必须要考虑的问题, 尤其是正态性和方差齐性一般都需要进行考察。
- ✧ 配伍设计的方差分析: 不考虑这两个问题, 这是由于正态性和方差齐性的考察是以单元格为基本单位的, 此时每个格子中只有一个元素, 当然没法分析了! 不止是配伍设计的方差分析, 只要是无重复数据的方差分析, 如交叉设计、正交设计等, 都不考虑这两个问题。
- ✧ 有重复数据的多因素方差分析: 由于方差齐性的考察是以单元格为基本单位, 此时单元格数目众多, 真正分到每个格子中的样本例数一般都只有 3~5 例, 此时很难检验出差别; 或者出现另一种极端情况, 因为极个别格子方差不齐而导致检验不能通过, 这种情况实际上对分析结果影响并不太严重。因此在多因素方差分析中, 方差齐性往往只限于理论探讨的程度。真正在应用时, 只要数据分布不是明显偏态, 不存在极端值即可, 这两种情况对结果的影响要比方差不齐严重得多。





根据 Box 的研究结果, 如果各组的例数相同 (即均衡), 或总体呈正态分布, 则方差分析模型对方差略微不齐有一定的耐受性, 只要最大/最小方差之比小于 3, 分析结果都是稳定的。



在实际操作中, 对数据正态性的考察有一个办法, 就是拟合完毕后作出残差分布图, 如果残差呈随机分布, 则可知 (单元格内) 原始数据满足正态条件。

## 1.2 Univariate 过程入门

如前所述, Univariate 过程是最为常用的一个过程, 几乎所有的实验设计都可以用它来进行分析, 它也是后面三个过程的基础。下面, 我们就以配伍组设计的方差分析为例来看一下该过程的使用。

### 1.2.1 引例

例 1.1 某厂医务室测定了 10 名氟作业工人工前、工中及工后 4 小时的尿氟浓度 ( $\mu\text{mol/L}$ )。问氟作业工人在这三个不同时间的尿氟浓度有无差别? (杨树勤, 《卫生统计学》第三版 P46)

工人编号	工前	工中	工后
1	90.53	142.12	87.38
2	88.43	163.17	65.27
3	47.37	63.16	68.43
4	175.80	166.33	210.54
5	100.01	144.75	194.75
6	46.32	126.33	65.27
7	73.69	138.96	200.02
8	105.27	126.33	100.01
9	86.32	121.06	105.27
10	60.01	73.69	58.95

	group	worker	fu
1	1.00	1.00	90.5
2	1.00	2.00	88.4
3	1.00	3.00	47.4
4	1.00	4.00	175.8
5	1.00	5.00	100.0
6	1.00	6.00	46.3

图 1.1 数据格式示意

解: 显然, 该数据的特点和以前学习过的配对数据非常相似, 也属于同一受试对象不同处理间的比较。但是, 这里的处理组成为了三组 (工前、工中和工后), 如果采用两两求出差值的方法来做配对  $t$  检验, 显然会犯和多个均数比较时用两两  $t$  检验一样的错误, 会扩大一类错误。这里我们应采用两因素的方差分析方法来分析。

根据统计分析的要求, 我们建立了三个变量来包括上述信息, time 表示测定时间为工前、工中或工后, worker 代表区组, weight 表示最终的体重增量, 如图 1.1 所示。建好的数据集存为文件 twoway.sav。

由于在配伍设计的数据中, 每个格子中只有一个元素 (无重复数据), 因此交互作用和方差齐性都无法检验, 下面只分析主效应即可。

Analyze → General Linear Model → Univariate

Dependent List 框: fu

要分析的应变量为 fu

Fixed Factor 框: group、worker

固定效应变量为 group、worker

Model:

☒ Custom

要求自定义方差分析模型

Build Terms 下拉列表: Main effects	模型中准备纳入主效应
Model 框: group、worker	模型中只纳入主效应 group、worker
Continue	
Post Hoc:	
Post Hoc Test For 框: group	要求对 group 作两两比较
<input checked="" type="checkbox"/> S-N-K	具体的两两比较方法为 SNK 法
Continue	
OK	

这里 worker 本应被选入随机因素框，此处将其也选入固定因素框是为了让输出结果和原书结果一致，同时也是基于下面的事实：在无重复数据的多因素方差分析中，随机效应无法被估计出来，两种做法在统计分析结论上完全相同。

### 1.2.2 界面说明

【主对话框】（见图 1.2）

1. Dependent Variable 框：选入需要分析的变量（应变量），只能选入一个。
2. Fixed Factors 框：选入自变量中的固定因素，对固定因素和随机因素的解释请参见前面的“常用术语”。
3. Random Factors 框：用于选入自变量中的随机因素变量，如果你看完前面的解释后还弄不明白，假装没看见这个框框就是了。

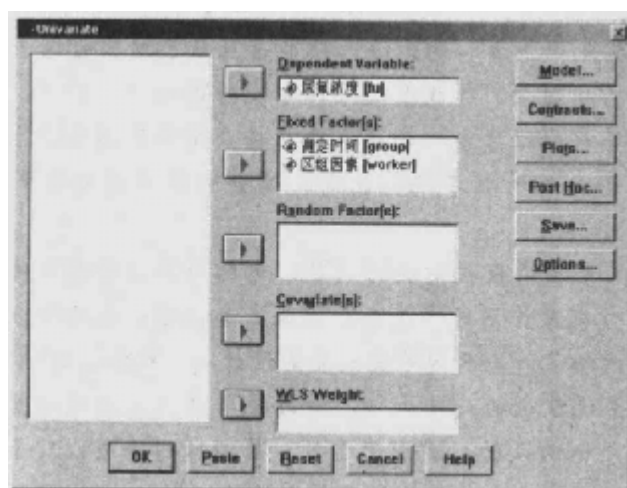


图 1.2 主对话框

4. Covariate 框：用于选入协变量，所谓协变量就是指一些与应变量、自变量可能都有关系的连续性变量，它们的存在可能会影响分析结果的正确性，从而不得不在分析中加以控制。这种控制了协变量的分析被称为协方差分析，后面有专门的一节介绍。

5. WLS Weight 框：即用于选入加权最小二乘法的权重系数。

此处进行的加权最小二乘法分析实际上和 Liner 过程主对话框中 WLS 框提供的功能相同，只不过在方差分析中该方法用得比较少，因此在这里我们不做

过多讲解，加权最小二乘法的详情请大家参见 Regression 菜单中的 Weight Estimation 过程。

从这里大家也可以看到方差分析模型和线性回归模型实际上是等价的。

【Model 子对话框】（见图 1.3）

该对话框是方差分析中最让初学者头痛和不理解的部分，往往做不出结果时毛病就在这里！好，现在让我们一起来看看。

1. Specify Model 组：用于对所用方差分析模型进行精确设定，可以规定模型中存在哪些主效应和交互效应。模型的默认情况为 Full factorial，即分析所有分类自变量的主效应和交互作用。这一般来说都用不着（从实用角度讲，三阶以上的交互作用都可以忽略），而且往往会导致模型无法拟合出结果来。比如此处的配伍设计的方差分析，由于每个格子中只有一个数值，不可能分析交互作用。强行分析的话会使得模型无法估计随机误差，从而无法进行检验。所以要改一下：将按钮切换到右侧的 custom，这时中部的 Build Term 下拉列表框就变黑可用，该框用于选择进入模型的因素交互作用级别，即是分析主效应、两阶交互、三阶交互，还是全部分析。这里我们只能分析主效应：选择 main，再用黑色箭头将 group 和 worker 选入右侧的 model 框中。

2. Sum of squares 下拉列表框：用于选择方差分析模型进行变异分解的方法，有 I 型到 IV 型四种，如果你搞不清它们之间的区别，使用默认的 III 型即可。



这里涉及到的统计理论实在太深，不过你要是真的敢看，我就敢讲。

I 型：分层处理平方和的方法，研究者往往已对因素的影响大小有了主次之分，该方法按因素引入模型的顺序依次对每项进行调整，因此，它的计算结果与因子的前后顺序密切相关。应当将最重要的因素放在前面，然后按二阶交互、三阶交互的顺序依次指定。该分解方法适用于平衡的模型和嵌套模型。

II 型：对其他所有效应均进行调整。它的计算会抑制其他参数的估计，所以不适用于有交互作用的方差分析以及嵌套模型。该分解方法适用范围较小，为完全平衡的设计，只牵涉主效应的设计以及纯粹的回归分析。

III 型：是系统默认的处理方法，对其他所有效应进行调整，但其计算方法也适用于不平衡的设计。适用于 I 型、II 型所列范围以及无缺失单元格的不平衡模型。对于含缺失单元格的不平衡设计，则应当使用下面的 IV 型。

IV 型：专门针对含有缺失单元格的数据而设计，它对任何效应计算平方和，如果效应存在嵌套，则只对效应的较高水平效应作对比。可用于 I 型、II 型所列模型，但更主要的是用于含缺失单元格的不平衡设计。

综上，除非很特殊的情况下必须要用到 IV 型，一般使用 III 型分解方法即可。

3. ☒ Include intercept in model：用于选择是否在模型中包括截距（即模型中的常数项  $\mu$ ），不用改动，默认即可。

【Contrast 子对话框】（见图 1.4）

用于对精细趋势检验和精确两两比较的选项进行定义。该对话框在这里比单因素



方差分析的时候还要专业,使用频率也更少。

1. **Factors** 框: 用于选择需要定义比较方法的因素。

2. **Change Contrast** 框组: 用于对比较方法进行定义。所用方法在 **Contrast** 下拉列表中, 然后单击 **Change** 按钮确定, 具体的比方法有:

- ✧ **None**: 不进行均数比较。
- ✧ **Deviation**: 偏对照, 将每个因子水平的均值与全部因子水平的均值作比较, 选中该项后需要在下方的 **Reference Category** 单选按钮组选择第一个或最后一个水平作为参考水平。
- ✧ **Simple**: 简单对照, 对因素每个水平的均值都与参考水平的均值进行比较, 需要选择一个参考水平。
- ✧ **Difference**: 亦被称为反 **Helmert** 对照, 除第一水平外, 因素的每个水平的均数都与该水平前各水平的总均数进行比较。
- ✧ **Helmert**: 与 **Difference** 刚好相反, 除最后一个水平外, 因素的每个水平的均数都与该水平后各水平的总均数进行比较。
- ✧ **Repeated**: 是对邻近水平的连续比较, 除第一水平外, 因素各水平的均数都与该水平前一个水平的均数作比较。
- ✧ **Polynomial**: 进行多项式比较, 如果该因素有  $n$  个水平, 则比较时会输出从线性到  $n-1$  次方曲线的比较结果。在该种比较中因素各水平被假设等间距。

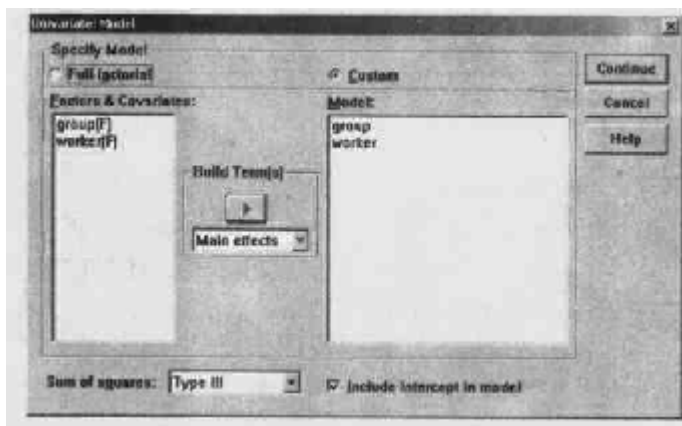


图 1.3 **Model** 子对话框

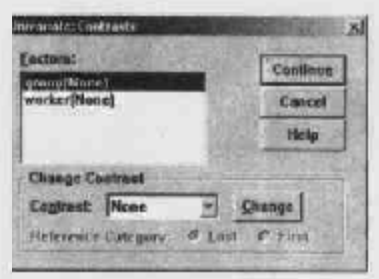



图 1.4 **Contrast** 子对话框

在 **Contrast** 框组中选好对比方法后单击 **Change** 按钮, 就可以看到选中的对比方法即参考水平显示在相应因素后的括号中。

 **Contrast** 子对话框的实质是进行哑变量的精细设置, 出于控制难度循序渐进, 使大家能逐步掌握的目的, 这里我只对比较方法作一简单介绍, 在 Logistic 模型和 Cox 比例风险模型中会更加深入的探讨该问题, 请大家参阅相应章节。

【**Plots** 子对话框】(见图 1.5)

用于指定用模型的某些参数作图。

1. **Factors** 框: 可用于作图的变量列表。

2. Horizontal Axis 框：做出选入该框中的因素的不同水平均数的曲线。
3. Separate Lines 框：按该框中因素的不同水平在同一张图中分层做出均数曲线。
4. Separate Plots 框：按该框中因素的不同水平分不同的图作均数曲线。
5. Plots 按钮组：该栏中有三个按钮：Add、Change 和 Remove。分别用于将上面的选择加入、更改和移出下方的绘图定义框。

#### 【Post Hoc】子对话框（见图 1.6）

该两两比较对话框和第 7 章单因素方差分析中的实际上一模一样，不再重复。

#### 【Save】子对话框（见图 1.7）

将模型拟合时产生的中间结果或参数保存为新变量供继续分析时用。

1. Predicted Values 复选框组：有标准化预测值、权重预测值以及预测值的标准误。其中权重的预测值只有在主对话框中选择了加权最小二乘变量（WLS）时才可用。
2. Diagnostics 复选框组：有 Cook's 距离和非中心化杠杆值。
3. Residuals 复选框组：有未标化、权重、标准化、史氏和剔除残差，其中史氏残差指的是将残差进行 t 变换，亦称学生化残差；剔除残差则是去掉当前记录时，当前模型对该记录应变量的预测值对应变量观测值的残差；权重的残差只有在主对话框中选择了权重最小二乘变量（WLS）时才可用。
4. Save to New File 组：选择一个新数据文件用于保存参数拟合的协方差矩阵。

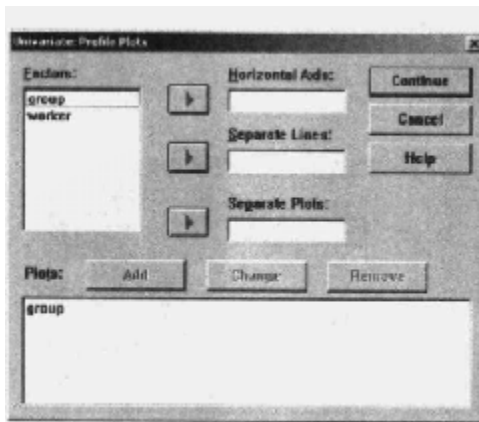


图 1.5 Plots 子对话框

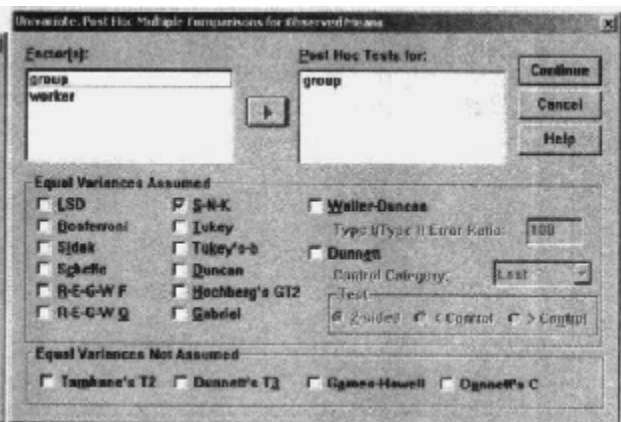


图 1.6 Post Hoc 子对话框



以上预测值、诊断指标和残差等的详细解释请参见线性回归部分。

#### 【Options】子对话框（见图 1.8）

当然是定义选项啦！

1. Estimated Marginal Means 组：用于定义输出哪些指标的估计均数、并选择所用的两两比较方法。
  - ✧ Factors and Factor Interactions 框：可用的因素及交互作用列表。
  - ✧ Display Means for 框：选入需要估计均数的因素和交互作用。
  - ✧ Compare main Effects：选中上方框内的因素或交互作用名称时才可用，要求

进行均数间的多重比较。

- ✧ **Confidence interval adjustment** 下拉列表：选中上方的多重比较复选框时才可用，用于定义具体的比较方法，共有 **LSD(none)**、**Bonferroni** 和 **Sidak** 三种方法，它们在实质上 and **Post Hoc** 子对话框中的同名方法是一样的。

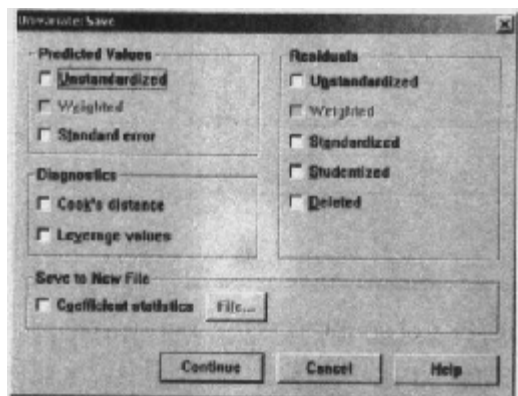


图 1.7 **Save** 子对话框

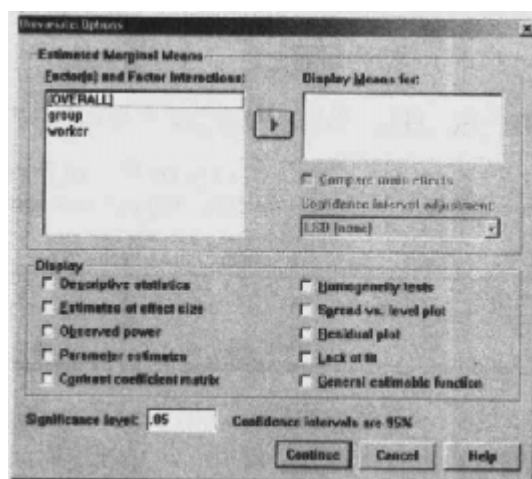


图 1.8 **Options** 子对话框

2. **Display** 复选框组：其他一些常用的、非常重要的输出指标，使用者往往会忽视里面的功能——其实这是个金矿。

- ✧ **Descriptive Statistics**：输出常用描述统计量。
- ✧ **Estimates of effects size**：对模型和各因素计算偏 eta 平方，它用于表示由该因素所导致的变异占应变量总变异的比例，其实在前面各章中我们已经见过这个指标了。
- ✧ **Observed power**：为模型和所有因素 / 交互项的检验计算检验效能，通过该数值我们可以得知试验设计的样本量是否充足，以及接近检验水准的因素有无必要继续研究。这是一个常被人忽视的选项。
- ✧ **Parameter estimates**：为所有因素各水平估计参数，等价于直接对各因素转换为哑变量使用多元线性回归拟合时计算出的系数值。
- ✧ **Contrast coefficient matrix**：列出计算系数时采用的变换矩阵（L 矩阵）。
- ✧ **Homogeneity**：进行方差齐性检验。
- ✧ **Spread vs. level plot**：绘出所有单元格的分布——水平图（该图的详情请参见 Explore 过程的相应内容）。
- ✧ **Residual plot**：绘制预测值、实测值与残差三者间两两散点图。这也是一个非常有用的功能，对检查模型拟合效果非常有帮助。
- ✧ **Lack of fit**：检查模型是否充分描述了应 / 自变量间的关系。如果无效假设被拒绝，则说明现有模型尚不能充分刻画应 / 自变量间的关系，可能还有交互作用未被发现，或尚有其他因素需要被引入模型。怎么样，这个选项非常有用吧。但是它的计算需要有一个或多个自变量的重复观测值。
- ✧ **General estimable function**：列出模型的设计矩阵。



再多解释一下 Lack of fit, 它指的是自变量为随机因素变量, 在重复测量时会有改变。如果变动太大, 则无效假设会被拒绝, 这说明现在的样本数据还不稳定, 需要增大样本量以补充信息。

3. Significance Level 框: 用于定义多重比较时采用的显著水平, 默认值为 0.05。

### 1.2.3 结果解释

## Univariate Analysis of Variance

Between-Subjects Factors

	测定时间			区组因素									
	1.00	2.00	3.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00
Value Label	工前	工中	工后										
N	10	10	10	3	3	3	3	3	3	3	3	3	3

这是一个所分析因素的取值情况列表, 没有什么不好懂的。这里为了节省空间, 我们将输出表格进行了转置。可见测定时间共分为三个水平, 各有 10 次测量; 区组因素共有 10 个水平, 各有三次测量。

Tests of Between-Subjects Effects

Dependent Variable: 尿氟浓度

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	47895.877 <sup>a</sup>	11	4354.171	4.513	.002
Intercept	362019.463	1	362019.463	375.246	.000
GROUP	8182.893	2	4091.447	4.241	.031
WORKER	39712.984	9	4412.554	4.574	.003
Error	17365.561	18	964.753		
Total	427280.901	30			
Corrected Total	65261.438	29			

<sup>a</sup>. R Squared = .734 (Adjusted R Squared = .571)

现在大家看到的是一个典型的方差分析表, 只不过是两因素的而已, 我来解释一下:

1. 首先是所用方差分析模型的检验, F 值为 4.513,  $P=0.02$ , 因此所用的模型有统计学意义, 可以用它来继续判断模型中系数有无统计学意义;

2. 第二行是截距, 它在我们的分析中没有实际意义, 忽略即可;

3. 第三行是变量 GROUP, 可见  $P=0.031$ , 它也有统计学意义;

4. 第四行是配伍因素 worker, 它的 P 值为 0.003, 也有统计学意义。

因此我们的结论是可以认为氟作业工人工前、工中及工后三个时间的尿氟浓度有差别, 但具体是哪两个或几个间存在差别, 还需要用后面的两两比较来回答。

上表的标题内容翻译如下:

变异来源	III 型方差 SS	自由度	均方 MS	统计量 F	P 值
校正的模型	47895.877	11	4354.171	4.513	.002
截距	362019.463	1	362019.463	375.246	.000
GROUP	8182.893	2	4091.447	4.241	.031
WORKER	39712.984	9	4412.554	4.574	.003
误差	17365.561	18	964.753		
合计	427280.901	30			
校正的合计	65261.438	29			

## Post Hoc Tests

### 测定时间

### Homogeneous Subsets

#### 尿氟浓度

Student-Newman-Keuls <sup>a,b</sup>			
测定时间	N	Subset	
		1	2
工前	10	87.375	
工后	10	115.589	115.589
工中	10		126.590
Sig.		.057	.439

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 964.753.

a. Uses Harmonic Mean Sample Size = 10.000.

b. Alpha = .05

现在是两两比较的结果, 方法为 SNK 法, 从中可见工前和工中的尿氟浓度是有差异的, 其余两两无差异。

#### 1.2.4 对引例的进一步分析

##### 【均数图与残差图】

图形可以帮助我们更加直观和清晰的了解分析结果, 使用 Plot 子对话框中的相应功能, 我们可以做出均数图与残差图如图 1.9、图 1.10 所示。

图 1.9 是用 Plots 子对话框做出的三个时间点的均数图, 可见氟作业开始后尿氟浓度迅速上升, 作业结束后呈现下降趋势。结合前面的分析结果, 均数图可以提供更多的信息, 从而更加直观的解释变动趋势。

图 1.10 是利用 Options 子对话框中的 Residual Plots 复选框做出的残差分布图, 图中分别给出了观察值、模型预测值和标准化残差间两两对应的散点图, 如果模型拟合得非常理想, 则观察值和预测值应当有非常密切的直线趋势, 而且标准化残差和他们各自的散点图都应当呈没有任何趋势的随机分布。显然, 本例的残差分布非常的理想。



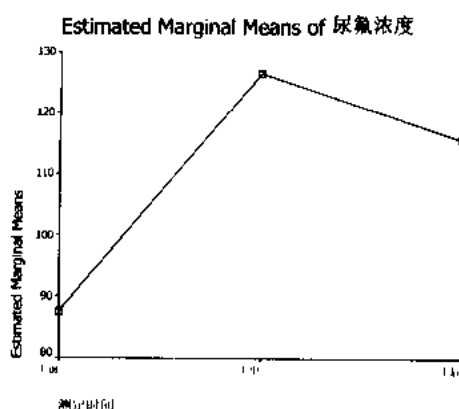


图 1.9 均数图

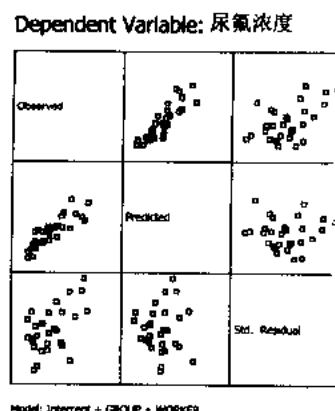


图 1.10 残差图

### 【检验效能的计算】

#### Tests of Between-Subjects Effects

Dependent Variable: 尿氮浓度							
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power <sup>a</sup>
Corrected Model	47895.877 <sup>b</sup>	11	4354.171	4.513	.002	49.646	.981
Intercept	362019.463	1	362019.463	375.246	.000	375.246	1.000
GROUP	8182.893	2	4091.447	4.241	.031	8.482	.666
WORKER	39712.984	9	4412.554	4.574	.003	41.164	.972
Error	17365.561	18	964.753				
Total	427280.901	30					
Corrected Total	65261.438	29					

<sup>a</sup> Computed using alpha = .05

<sup>b</sup> R Squared = .734 (Adjusted R Squared = .671)

检验效能在统计分析中非常重要，许多时候检验结果无差异是因为效能不够。本次结果已检验出统计学差异，不存在此问题。但我们可以看看这次分析的检验效能有多大。选中 Options 子对话框中的 Observed Power 复选框，则方差分析表最右侧会多出一列结果如上表所示，可见对因素 group 的检验效能仅有 0.666，显然比较低，这次我们能检验出差异，应当说运气还可以，但下一次就不敢说了。

### 【按随机因素分析 worker 变量】

前面我们将工人按固定因素变量来分析，如果按照正确的随机因素变量分析，则结果中方差分析表如下：

#### Tests of Between-Subjects Effects

Dependent Variable: 尿氮浓度						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	362019.463	1	362019.463	82.043	.000
	Error	39712.984	9	4412.554 <sup>a</sup>		
GROUP	Hypothesis	8182.893	2	4091.447	4.241	.031
	Error	17365.561	18	964.753 <sup>b</sup>		
WORKER	Hypothesis	39712.984	9	4412.554	4.574	.003
	Error	17365.561	18	964.753 <sup>b</sup>		

<sup>a</sup> MS(WORKER)

<sup>b</sup> MS(Error)

上表中对 group 和 worker 的统计分析结果和前面完全相同，截距的分析结果虽然

不同,但它是我们并不关心的东西。模型还对每一部分计算了均方,此处没有实际意义,后面会结合实例具体讲解。

### 1.3 常用试验设计及分析方法详解

现实世界是错综复杂的,各种因素的影响交织在一起,事物间真实的联系被掩盖在重重迷雾之下,难以一眼看穿。正是因为如此,统计学才成为了迷雾中的一盏明灯,指引、帮助着我们发现这些客观规律。然而,一个设计粗糙,质量控制混乱的实验所带来的误差足可以完全淹没掉真实的信息;另一方面,研究所需费用也并非无限,如何能以尽量少的耗费完成相同的研究任务是一个非常现实的问题。一个周密的试验设计能合理的安排各种试验因素,正确的估计样本含量大小,严格的控制试验误差,从而用较少的人力、物力和时间,最大限度的获得丰富而可靠的资料。因此,人们发展了一系列的试验设计方法。由于这些方法非常常用,而它们虽然几乎都可以采用方差分析模型来分析,但具体操作上又各有特点,本节将介绍一些常用的试验设计方法,并重点介绍在 SPSS 中应当如何对这种数据进行分析。

#### 1.3.1 完全随机设计 (Completely Random Design)

亦被称为单因素设计。该设计方案仅涉及一个处理因素,它是将受试对象按随机化的方法分配到各个处理组中,观察实验效应。各个处理组样本含量可以相等,也可以不等,但相等时效率较高。本设计的优点是简单易行,缺点是只能分析一个因素,和其他试验设计相比效率不太高。

在统计分析时,完全随机设计可以采用单因素方差分析来进行。在分析时应考虑到数据的正态性、方差齐性等问题。如果不满足这些条件,则应当进行变量变换以达到要求;若无法找到合适的变换方法,则应当采用非参数方法来检验。

如果分析结果显示该因素有统计学意义,则应当继续进行各组均数间的两两比较,比较次数较少的时候各种方法差异不大,但较多时就应当根据以下几点来确定具体的两两比较方法:1. 是证实性研究还是探索性研究,前者往往只需比较少数几组,而后者一般要进行两两比较;2. 有无明确的对照组;3. 重点是希望控制一类错误还是二类错误。

根据对相关研究的检索结果,除了参照所研究领域的惯例外,两两比较方法的选择可以参照如下标准:如果不存在明确的对照组,要进行的是验证性研究,即计划好的某两个或几个组间(和对照组)的比较,宜用 Bonferroni (LSD) 法;若需要进行的是多个均数间的两两比较(探索性研究),且各组人数相等,适宜用 Tukey 法;其他情况宜用 Scheffe 法。该标准仅供大家参考。

#### 1.3.2 配伍设计 (Randomized Block Design)

亦称随机区组设计,或双因素无重复试验设计。可以分为两种情况:1. 对同一个受试对象在同一处理不同水平间的比较;2. 将几个受试对象按一定条件划分成配伍组,再将每一配伍组的各受试者随机分配到各个处理组中去。每个配伍组的例数等于

处理组个数。用于配伍的因素应当是影响实验效应的主要非处理因素。例如，在动物实验中，常将窝别、性别相同，体重相近的两个动物配成对子；人群试验中，常将性别相同、年龄相近的两个人配成对子，这样可提高各处理组间的均衡性。

当只有两个配伍组时，该设计就变成了常用的配对设计。



**Randomized Block** 这个怪怪的英文名称来自 R.A.Fisher，他老人家在伦敦附近的 Rothamsted 农业实验站创立的随机化实验设计和方差分析的理论和方法在全世界得到了广泛运用。Fisher 做的是农田试验，土地都是被划分成一块一块的 Block 来分配的，所以该设计就起了这么个名字。

配伍设计资料的分析方法为两因素方差分析。由于单元格内无重复数据，交互作用和方差齐性均无法考察。

需要特别指出的是：如果配伍因素也是希望研究的因素，或者配伍因素与研究因素间的交互作用不能忽略，则不应当采用配伍设计，而应当采用下面的各种设计方法。

### 1.3.3 交叉设计 (Cross-over Design)

交叉设计是一种特殊的自身对照设计。它克服了试验前后自身对照由于观察期间由于各种非试验因素对试验结果的影响所造成的偏移。以平衡的两阶段交叉设计为例，若研究 A、B 两种药物的疗效是否不同，则设两种顺序，即 AB 和 BA，首先将条件相近的观察对象配对，再将 2n 名实验对象随机分配到这两种顺序组中，AB 组先接受 A 药，而 BA 组先接受 B 药，经过一段“效应去除时间” (washout time) 后，AB 组病人改用 B 药，BA 组病人改用 A 药。这样每个病人都均匀的接受了两个阶段、两种药物的处理，在最大限度上平衡了各种非研究因素对结果的影响。显然，这些病人间唯一不同的因素是处理顺序，其中一半病人为 AB 顺序，而另一半病人为 BA 顺序，但在整个样本中这一因素是平衡的。

交叉设计的优点是：1. 节约样本含量；2. 能够控制时间因素及个体差异对处理方式的影响；3. 每一个试验对象同时接受试验因素和对照（如安慰剂），从医德的观点出发，均等地考虑了每一个患者的利益。

交叉设计的缺点也非常明显：它不允许有病人失访，否则会造成该个体已有数据的完全浪费；不适用于病程较短的急性病治疗效果的研究，如大叶肺炎、急性扁桃腺炎等，因为如果在第一阶段给予试验措施该病便已治愈，第二阶段的措施则不可能反映出来。因此，它只适用于某些病程相对较长，且不大可能完全治愈的疾病。

交叉设计的数据可以采用方差分析法来进行统计，所观察到的数据的变异有以下几个来源：处理（药物）效应、阶段效应、顺序效应和个体差异。其中处理效应是希望研究的因素，而顺序效应则在目前常用的统计分析中被忽略，因为这是交叉设计能够实施的前提条件。保证顺序效应能够被忽略的办法是给予足够长的效应去除时间，以使前一种药物的效果完全消失，不至于影响后面药物的疗效观察。

例 1.2 为研究 12 名高血压病人用 A、B 两方案疗效的差别，随机地让其中 6 名病人先以 A 法治疗，后以 B 法治疗；另外 6 名病人先 B 法，后 A 法。记录治疗后血压的下降值(kPa)，请分析 A、B 两方案疗效有无差别。（倪宗瓚《医学统计学》第二版 P88）

阶段	病人编号											
	1	2	3	4	5	6	7	8	9	10	11	12
I	B	B	A	B	A	A	A	A	B	B	B	A
	3.07	1.33	4.40	1.87	3.20	3.73	4.13	1.07	1.07	2.27	3.47	2.40
II	A	A	B	A	B	B	B	B	A	A	A	B
	2.80	1.47	3.73	3.60	2.67	1.60	2.67	1.73	1.47	1.87	3.47	1.73

解：该数据已输入为文件 crossover.sav，变量 bp 为血压值，treat、stage、patient 分别为治疗方案、治疗阶段和病人编号。分析时应注意不要引入交互作用，操作如下：

Analyze→General Linear Model→Univariate

Dependent List 框：bp

要分析的应变量为 bp

Fixed Factor 框：treat、stage

固定因素为 treat、stage

Random Factor 框：patient

随机因素为 patient

Model:

☒ Custom

要求自定义方差分析模型

Build Terms 下拉列表：Main effects

模型中准备纳入主效应

Model 框：treat、stage、patient

只纳入主效应 treat、stage、patient

Continue

OK

结果中的方差分析表如下：

Tests of Between-Subjects Effects


Dependent Variable: 血压下降值

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	154.128	1	154.128	94.673	.000
	Error	17.908	11	1.628 <sup>a</sup>		
TREAT	Hypothesis	1.707	1	1.707	4.599	.058
	Error	3.711	10	.371 <sup>b</sup>		
STAGE	Hypothesis	.427	1	.427	1.150	.309
	Error	3.711	10	.371 <sup>b</sup>		
PATIENT	Hypothesis	17.908	11	1.628	4.387	.014
	Error	3.711	10	.371 <sup>b</sup>		

<sup>a</sup> MS(PATIENT)

<sup>b</sup> MS(Error)

由上表可见，变量 treat 的  $F=4.599$ ， $P=0.058$ ，因此尚不能认为两种治疗方案的疗效有差别。

 由于数据为自身配对，单元格中没有重复数据，因此上面分析中也可以将 patient 就按固定因素来分析，结果完全相同。

这里可能有朋友会提出，既然 stage 没有统计学意义，能否将该变量剔除出模型重新分析？我认为这样是不合适的：交叉设计最常用于新药临床试验，此时并非探索性分析，而是证实性分析。证实性分析在试验设计时就充分考虑到了所有因素，从而决定了所用的统计方法，绝对不能根据统计结果来更改方法，这种做法是统计分析，尤其是证实性分析的大忌。未能检出统计学差异也许只是因为检验效能不够高（本例中

stage 的检验效能仅有 0.163)，期次实际上反映的是时间的影响，该试验设计中考虑到了期次这个因素，实际上就已经默认该因素有影响，不宜将其忽略，这样做等于是更改了试验设计。打个比方，如果配伍设计方差分析中配伍因素无统计学意义，我们能认为同一个体的不同次测量数据间无联系，配伍因素不存在，从而按成组设计来分析吗？当然不能。

### 1.3.4 析因设计 (Factorial Design)

在临床研究中，许多试验因素之间往往是相互联系、相互制约的。当一种因素的质和量改变时另一种现象的质和量也随之改变，几个各因素间存在交互作用时，析因设计是一种非常理想的设计。以  $2 \times 2$  析因设计为例，它不仅检验两因素各水平之间的差异有无统计学意义，而且可以同时检验两因素间的交互作用。

除可以分析交互作用外，析因设计的优点还在于可以节约样本含量，和将两种药物分别进行随机对照试验相比，析因设计将节约样本含量的  $1/2$ ，若用两种药物相互对比的设计，可节约  $1/3$  的样本含量。

在析因设计资料的分析中，应当先重点考察各因素间是否存在交互作用，如果存在交互作用，此时各因子的主效应检验结果已无实际意义，应当按各因素各种水平的组合来研究。

例 1.3 治疗缺铁性贫血病人 12 例，分为 4 组给予不同治疗，一个月后观察红细胞增加数（百万/ $\text{mm}^3$ ）。现欲研究甲乙药单独使用的疗效如何？同时使用的疗效又如何（倪宗瓚，《医学统计学》第二版 P92）？

解：按照统计分析的要求，本例应建立数据格式如图 1.11 所示。

		甲 药			druga	drugb	rbc
		用	不用				
乙 药	用	2.1	0.9	1	1.00	1.00	2.10
		2.2	1.1	2	1.00	1.00	2.20
	不用	2.0	1.0	3	1.00	1.00	2.00
		1.3	0.8	4	1.00	.00	1.30
		1.2	0.9				
		1.1	0.7	5	1.00	.00	1.20

图 1.11 数据格式示意

druga 表示是否服用甲药，drugb 表示是否服用乙药，rbc 为红细胞增加数。见数据文件 xiyin.sav。分析步骤如下：

Analyze → General Linear Model → Univariate

Dependent List 框: rbc

要分析的应变量为 rbc

Fixed Factor 框: drug a、drug b

自变量为 drug a、drug b

Plots:

Horizontal axis 框: drug a

Separate lines 框: drug b: Add

做出因素各种水平组合下的效应图

Continue

OK

主要的分析结果如下：



## Tests of Between-Subjects Effects

Dependent Variable: 红细胞增加数

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.963 <sup>a</sup>	3	.988	98.750	.000
Intercept	19.507	1	19.507	1950.750	.000
DRUGA	1.688	1	1.688	168.750	.000
DRUGB	.907	1	.907	90.750	.000
DRUGA * DRUGB	.367	1	.367	36.750	.000
Error	8.000E-02	8	1.000E-02		
Total	22.550	12			
Corrected Total	3.043	11			

<sup>a</sup>. R Squared = .974 (Adjusted R Squared = .964)

从上面的方差分析表可见两种药物的交互作用有统计学意义, 因此它们各自的主效应有无统计学意义已经没有实用价值, 上面的分析结果仅供参考。

此处的均数图也显示, 无论是甲药还是乙药, 当与另一种药物一起使用时其效果都明显优于单独使用(见图 1.12)。

由于前面分析出交互作用有统计学意义, 因此下面应当继续按各种不同的组合来分析, 即:

1. 不使用乙药时甲药有无作用 (a0b0 和 a1b0 比)。
2. 使用乙药时甲药有无作用 (a0b1 和 a1b1 比)。
3. 不使用甲药时乙药有无作用 (a0b0 和 a0b1 比)。
4. 使用甲药时乙药有无作用 (a1b0 和 a1b1 比)。

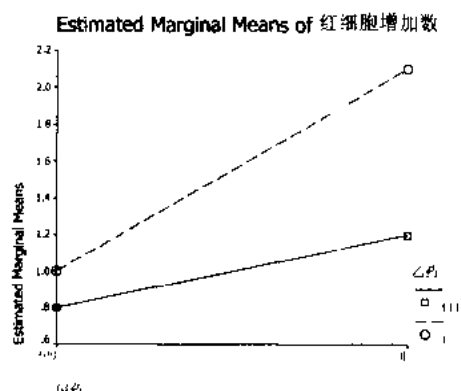


图 1.12 均数图

这样才能真正准确的回答两种药物疗效如何的问题。这种精细的两两比较在 SPSS 中是不能用 GLM 过程的对话框界面完成的, 但可以在编程窗口中进行。因篇幅所限, 此处不再详述。

## 1.3.5 拉丁方设计 (Latin Square Design)

如果在临床实验过程中涉及到三个因素(一般为一个试验因素, 两个区组因素); 此时按照普通的多因素设计方法需要的样本例数就较多, 从而所需费用太高, 对试验质量的控制也比较麻烦。如果各个因素间无交互作用且水平数相等, 则可采用拉丁方设计。拉丁方设计是将三个因素按水平数  $r$  排列成一个  $r \times r$  随机方阵, 如  $3 \times 3$  拉丁方,  $4 \times 4$  拉丁方, 在同一行或同一列中任何一个因素的水平均无重复。由于这些方阵最早是用拉丁字母来排列的, 故得名拉丁方设计。

拉丁方设计可看成纵横两向皆为配伍组，且可以用较少的重复次数获得较多的信息，因而比配伍组设计更优越。但它要求各因素的水平数相等且无交互作用，并要尽量避免数据缺失，在实际应用中有一定局限性。

例 1.4 为研究不同剂量甲状腺素注射液对甲状腺体的影响，以豚鼠 5 个种系，每个种系各 5 只，分养于 5 个笼子，每笼内放置各种系豚鼠 1 只，并以甲状腺素的 5 个不同剂量分别注射，测得甲状腺体的重量见下表，试分析不同剂量甲状腺素组均数间的差别（郭祖超，《医用数理统计方法》P274）。

种系	笼号				
	I	II	III	IV	V
甲	C(65)	E(85)	A(57)	B(49)	D(79)
乙	E(82)	A(73)	D(92)	B(81)	B(63)
丙	D(68)	C(67)	A(56)	D(77)	C(51)
丁	B(63)	E(99)	C(70)	E(76)	A(41)
戊	D(75)	A(46)	B(52)	E(68)	C(66)

解：按照一个观察对象一条记录的原则，该数据集的录入格式如图 1.13 所示。

	jiliang	zhongxi	longzi	xianti
1	3.00	1.00	1.00	65.00
2	5.00	2.00	1.00	82.00
3	1.00	3.00	1.00	73.00
4	4.00	4.00	1.00	92.00
5	2.00	5.00	1.00	81.00

图 1.13 数据格式示意图

见数据文件“latin.say”，以上四个变量分别表示甲状腺素剂量、种系、笼子和甲状腺体重量。分析时按三因素方差分析操作即可，但应注意由于拉丁方设计中不能考虑交互作用，模型设置时应只引入主效应。

Analyze→General Linear Model→Univariate

Dependent List 框: xianti

要分析的应变量为 xianti

Fixed Factor 框: jiliang、zhongxi、longzi

Model:

☒ Custom

要求自定义方差分析模型

Build Terms 下拉列表: Main effects

模型中准备纳入主效应

Model 框: jiliang、zhongxi、longzi

只纳入主效应 jiliang、zhongxi、longzi

Continue

OK

结果中的方差分析表如下：

## Tests of Between-Subjects Effects

Dependent Variable: 甲状腺体重量

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3974.880 <sup>a</sup>	12	331.240	3.943	.012
Intercept	115736.040	1	115736.040	1377.701	.000
JILIANG	2690.960	4	672.740	8.008	.002
ZHONGXI	375.760	4	93.940	1.118	.393
LONGZI	908.160	4	227.040	2.703	.081
Error	1008.080	12	84.007		
Total	120719.000	25			
Corrected Total	4982.960	24			

<sup>a</sup> R Squared = .798 (Adjusted R Squared = .595)

从中可见种系和笼子号对结果无影响,但注射剂量对腺体重量是有影响的。为此继续用 SNK 法对剂量组作两两比较,结果如下:

## 甲状腺体重量

Student-Newman-Keuls<sup>a,b</sup>

甲状腺素剂量	N	Subset	
		1	2
A	5	54.6000	
B	5	61.6000	
C	5	63.8000	
D	5		78.2000
E	5		82.0000
Sig.		.288	.524

Means for groups in homogeneous subsets are displayed.

Based on Type III Sum of Squares

The error term is Mean Square(Error) = 84.007.

<sup>a</sup>. Uses Harmonic Mean Sample Size = 5.000.<sup>b</sup>. Alpha = .05.

可见 D、E 两个剂量组的腺体重量明显高于 A、B、C 三组。

## 1.3.6 正交设计 (Orthogonal Design)

当实验涉及的因素在三个或三个以上,且因素间可能存在交互作用时,此时使用拉丁方设计就不合适了。这种情况下可使用正交设计。它利用统计学家专门设计的一套规格化的正交表将各试验因素、各水平之间的组合进行均匀搭配,从而可以用较少的、有代表性的处理组合数,提供充分有用的信息。与普通的因素试验设计相比,它在仍然可以分析主效应与交互作用的同时大大减少了样本量,因而是一种高效、快速的多因素试验设计方法。

正交设计的步骤是:

1. 根据实验目的, 确定实验因素及其水平数, 重要因素的水平数可多些, 次要因素的水平数可少些。

2. 确定因素间是否存在交互作用及观察的交互作用项。

3. 确定试验次数。可根据需要与可能而定。

4. 综合上述情况选取正交表, 然后按表上安排进行实验。在以前人们必须查阅专用统计书籍来找到正交表格的排列方法, 现在有了 SPSS, 它已经在 Data 菜单中为我们提供了正交设计模块, 只需要按需要选好试验因素数、水平数以及样本例数, 系统就会自动生成相应格式的数据文件。

 有关 Orthogonal Design 过程的详细介绍请参阅联合分析一章中的相应内容。

例 1.5 过氧乙酸是广泛应用的消毒剂, 但其有效成份极不稳定, 以致影响其消毒效果。现欲通过试验找出有关因素对其稳定性的影响, 选出各因素的一个最佳组合, 组成保持过氧乙酸稳定性的最优条件。已知的可能影响因素及水平有:

	水平	
	1	2
A: 稳定剂	加磷酸 0.3%	不加磷酸
B: 水浴温度	25~30℃	35~40℃
C: 浸泡口表	浸泡口表 10 支	不浸口表
D: 加盖与否	加盖	不加盖

同时稳定剂与水浴温度、稳定剂与加盖与否间可能存在交互作用, 现已根据 L8 正交设计表进行了试验, 每种组合重复两次, 测得数据如下:

A	B	C	D	第一次测量	第二次测量
1	1	1	1	7	4.11
1	1	2	2	6.05	3.5
1	2	1	2	1.1	0.8
1	2	2	1	1.9	0.96
2	1	1	2	2.4	1.65
2	1	2	1	4	1.5
2	2	1	1	0.35	0.3
2	2	2	2	0.3	0.9

请根据以上数据找出最佳条件 (郭祖超, 《医用数理统计方法》第三版 P793)。

解: 根据题意, 该试验共有四个因素, 每个因素均为两个水平, 可以直接使用 Orthogonal Design 过程进行设计, 但应注意此处实际上是将分析主效应用的 L8 表重复了两次, 而不是采用设计出的 L16 正交表, 我们生成表格的时候也应如此照办, 否则结果不正确, 操作如下:

Data → Orthogonal Design → Generate

Factor Name 框: 输入 A:  输入 B:  输入 C:  输入 D:

选中变量 A:  Value 列: 头三行分别输入 1、2:

选中变量 B:  Value 列: 头两行分别输入 1、2:

选中变量 C:  Value 列: 头两行分别输入 1、2:

选中变量 D:  Value 列: 头两行分别输入 1、2:

☒ Replace Working data file

系统生成的数据文件如图 1.14 所示。

	a	b	c	d	status_	card_
1	1.00	2.00	2.00	2.00	0	1
2	2.00	1.00	1.00	2.00	0	2
3	2.00	2.00	1.00	1.00	0	3
4	2.00	1.00	2.00	2.00	0	4
5	1.00	1.00	2.00	1.00	0	5
6	1.00	2.00	1.00	2.00	0	6
7	1.00	1.00	1.00	1.00	0	7
8	2.00	2.00	2.00	1.00	0	8

图 1.14 系统生成的数据文件

现在数据的基本结构已经形成，只需要用拷贝、粘贴的办法重复这 8 条记录，然后建立一个用于记录过氧化乙酸残存量的结果变量（yisuan，具体数据见 orthol.sav），下面就可以进行分析了，操作如下：

Analyze→General Lineal Model→Univariate

Dependent List 框: yisuan

要分析的应变量为 yisuan

Fixed Factor 框: a、b、c、d

Model:

☒ Custom

要求自定义方差分析模型

Model 框: a、b、c、d

纳入主效应

Model 框: a\*b、a\*d

按设计要求纳入交互效应

Continue

OK

分析结果如下：

#### Tests of Between-Subjects Effects

Dependent Variable: YISUAN

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	52.315 <sup>a</sup>	6	8.719	6.581	.007
Intercept	84.732	1	84.732	63.954	.000
A	12.285	1	12.285	9.272	.014
B	34.810	1	34.810	26.274	.001
C	.122	1	.122	.092	.768
D	.731	1	.731	.552	.477
A * B	4.202	1	4.202	3.172	.109
A * D	.164	1	.164	.124	.733
Error	11.924	9	1.325		
Total	148.971	16			
Corrected Total	64.239	15			

<sup>a</sup>. R Squared = .814 (Adjusted R Squared = .691)



上表显示：稳定剂与温度是影响过氧乙酸稳定性的主要因素，另两个对稳定性影响不大，稳定剂于其他变量的交互作用也不显著。因此，保持过氧乙酸稳定性的最佳条件为加 0.3% 磷酸、水浴温度 25~30 度。

### 1.3.7 星点设计 (Central Composite Design)

星点设计是近年来国外药学研究人员常用的实验设计方法，它比起析因设计等方法来有着自身的优越性，这里我们对其作一简介，希望能对大家有所帮助。

正交设计和析因设计都是同时考察多个因素，从中筛选最佳条件的优秀设计方法，其应用已比较普及。但是，它们存在以下弊端：

1. 两者均是基于线性模型的设计，实际上有的时候试验因素和结果变量间可能是曲线关系，基于这两种设计的分析可能会给出错误的信息。比如一般效应值在最佳实验条件区域附近变化会比较灵敏，其效应面多为曲面。如果用线性模型分析，很可能分析结果为该因素对结果无影响！

2. 一般来说我们不可能将效应变量的取值范围全部取完，只可能取其中的一部分，如温度取 20 度、30 度等，也许这些选择的取值仅仅是接近最佳取值（比如催化效果在 23 度最佳），用传统的设计和分析方法就无法精确找到最佳点。

3. 对交互作用的分析也是如此，也许分析结果没有发现交互作用，仅仅是因为我们取的数值刚好在交互曲面的底部，从而未能发现。

针对以上问题，国外近年来提出并采用了星点设计。它在正交或析因设计的基础上将自变量与应变量的关系扩展到了曲面，其操作步骤大致为：

1. 确立考察因素的水平范围：用预实验的办法大致确定该因素的最佳取值范围。

2. 效应面设计：可以采用析因设计或正交设计，但此处最常用的为星点设计，它是在二水平析因设计的基础上加上极值点和中心点构成的，由于加入了这些成分，使得自变量与应变量的关系可以扩展到曲面。根据不同的目的，星点设计也有系列的设计表格供设计者使用。

3. 结果分析：结果的分析可以用方差分析来进行，但一般推荐采用二次以上的多元非线性拟合，在 SPSS 中可以使用 Linear 过程或 Nonlinear 过程来操作。通过不断的剔除无意义的交互项和高次项，最终得到一个最佳的简化方程。如果结果为线性模型，则与普通的设计无太大差别，分析到此结束。

4. 模型优化与预测：根据所建立的数学模型描绘出三维效应面，从效应面的较优区域直接读取较佳工艺范围。通过对每一个因素划定较佳工艺范围，最后将所有因素结合起来，选择出最佳工艺条件，并将该条件代入模型考察预测效果。如效果好，则分析完成。

以上是对星点设计的简单介绍，希望详细了解相应内容的朋友可以参阅相关药学文献。

### 1.3.8 嵌套设计 (Nested Design) 与裂区设计 (Split-plot Design)

这里需要指出的是嵌套设计与裂区设计是完全不同的两种设计类型，只是因为它们在分析方法上的相似而放到一起来讲述。

在嵌套设计中,各个试验因素的影响有主次之分,这在设计时就应当根据专业知识确定。次要因素的各个水平是嵌套在主要因素的水平下的,因而在统计分析时不能分析它们之间的交互作用。具体来说有两种情况:

- ✧ 最终的实验条件是各因素各水平的全面组合,只是专业上有主次之分。这种情况和其他试验设计区分时主要是看各因素有无主次差别。
- ✧ 最终的实验条件并非各因素各水平的全面组合,而是在主因素不同水平下次要因素有不同的水平。比如说我们要比较几种抗菌素的抑菌作用,抗菌素的用药浓度也应当是一个考虑的因素,浓度的作用显然比抗菌素种类要次要一些,而且不同的药物所用的浓度也完全不同,如 A 药物的浓度是 50、100、200, B 药物的是 100、200、300。在这里,浓度的作用就嵌套在药物中。

裂区设计的特点是实验因素并非一次安排完毕,而是分为两次甚至多次,因此被称为裂区设计。首先安排的是影响最为重要的,或者必须最先安排,或者材料消耗较大、工序较难改变的因素,经过一段时间后,影响较小,或者精确度要求较高的次要因素才加入主要因素的不同水平中。

无论是嵌套设计与裂区设计,它们的特点都是在设计时就充分利用了已知的信息,将主要影响因素优先安排,从而提高了分析效率,避免造成不必要的样本量浪费。这样在分析时也应考虑到影响因素的主次之分,按照嵌套的方差分析模型来进行分析。如果按照普通的方差分析模型来分析,则在变异的计算上和其他设计没有任何区别,不能利用设计中提供的影响因素的主次信息,完全不能体现出这两种设计的优越性。

例 1.6 某单位研究三种饲料的营养价值,用雄性大白鼠做实验,先将 12 只大白鼠随机分为三组,每组饲喂一种饲料。于 10 周、11 周、12 周和 13 周分别测量体重增加量,数据格式如图 1.15 所示,全部的数据见文件 nested.sav,请作统计分析。

	siliao	mouse	week	weight
1	1.00	1.00	.00	39.20
2	1.00	1.00	1.00	41.10
3	1.00	1.00	2.00	41.30
4	1.00	1.00	3.00	49.40
5	1.00	4.00	.00	37.60

图 1.15 数据格式

解:显然,饲料是主要的研究因素,它对体重增加的影响是最大的,大白鼠的个体因素是嵌套在各个饲料组中,对结果的影响较为次要。最后观察时间(周次)也属于次要影响因素,分析时应根据以上原则来处理,采用嵌套的方差分析模型,但是,该模型在 Univariate 对话框中无法直接实现,需要在编程窗口中对生成的程序作进一步修改。

Analyze→General Lineal Model→Univariate

Dependent List 框: weight

要分析的应变量为体重增量

Fixed Factor 框: siliao、week

Random Factor 框: mouse

Model:

☒ Custom

Model 框: siliao、week、mouse

Continue

Paste

要求自定义方差分析模型  
分析中只纳入主效应

以上操作生成的程序如左下所示, 请将其改为右下方的程序, 即将 Design 子句更改为我们所需的嵌套模型:

UNIANOVA

weight BY siliao week mouse  
/RANDOM = mouse  
/METHOD = SSTYPE(3)  
/INTERCEPT = INCLUDE  
/CRITERIA = ALPHA(.05)  
/DESIGN = siliao week mouse .

UNIANOVA

weight BY siliao week mouse  
/RANDOM = mouse  
/METHOD = SSTYPE(3)  
/INTERCEPT = INCLUDE  
/CRITERIA = ALPHA(.05)  
/DESIGN = siliao week mouse(siliao) .

运行右侧的程序, 结果如下:

#### Tests of Between-Subjects Effects

Dependent Variable: WEIGHT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	80876.710	1	80876.710	5199.147	.000
	Error	140.002	9	15.556 <sup>a</sup>		
SILIAO	Hypothesis	709.790	2	354.895	22.814	.000
	Error	140.002	9	15.556 <sup>a</sup>		
WEEK	Hypothesis	856.994	3	285.665	12.277	.000
	Error	767.874	33	23.269 <sup>b</sup>		
MOUSE(SILIAO)	Hypothesis	140.002	9	15.556	.689	.731
	Error	767.874	33	23.269 <sup>b</sup>		

<sup>a</sup> MS(MOUSE(SILIAO))

<sup>b</sup> MS(Error)

结果显示饲料种类和测量时间有统计学意义, 由上表的注解可见, 饲料种类在检验时采用的是二级单位误差 (mouse), 而周次则采用的是随机误差, 这正是我们所需要的。

由于上面的结果显示饲料种类有差异, 我们需要用两两比较的方法进一步分析, 这里仍然要用程序, 但已经不能通过简单的改动来实现, 限于本书范围, 此处不再详述。



有兴趣的朋友可以用对话框方式作一下上面的分析, 会发现结果完全不相同, 这主要是饲料种类检验时所用的误差不同所致。

## 1.4 协方差分析

### 1.4.1 概述

试验设计的重要任务之一就是尽力排除非处理因素的干扰和影响,从而准确地获得处理因素的试验效应。但是,许多需要控制的因素并不能简单的按分类变量来对待。比如两种药物治疗高血压的疗效,如果两组患者年龄分布不同,则必须考虑年龄对舒张压的影响。此时年龄是需要控制、排除影响的因素,它在统计分析中又称协变量。若忽视协变量的影响直接对这些资料进行分析,就可能会得出错误的结论。

协方差分析(analysis of covariance)是将直线回归和方差分析结合应用的一种统计方法,用来消除混杂因素对分析指标的影响。它的基本思想是在作两组或多组均数  $\bar{Y}_1, \bar{Y}_2, \dots, \bar{Y}_k$  之间的比较前,用直线回归方法找出各组  $Y$  与协变量  $X$  之间的数量关系,求得在假定  $X$  相等时的修正均数  $\hat{Y}'_1, \hat{Y}'_2, \dots, \hat{Y}'_k$ ,然后用方差分析比较修正均数间的差别。协方差分析是定量变量分析中控制混杂因素的重要手段之一,是医务工作者应掌握的一种重要统计分析方法。

进行协方差分析时应具备以下两个条件:

- ◇  $X$  与  $Y$  的线性关系在各组均成立,而且不同组间的总体回归系数应相等。
- ◇ 各比较组间协变量  $X$  的取值范围不宜相差太大,否则修正均数的差值在回归直线的延长线上,此时不知道回归线外推后是否仍然满足平行性和线性关系的条件,协方差分析的结论可能不正确。

例 1.7 某医生欲了解成年人体重正常者与超重者的血清胆固醇是否不同。而胆固醇含量与年龄有关,资料见下表:

正常组		超重组	
年龄	胆固醇	年龄	胆固醇
48	3.5	58	7.3
33	4.6	41	4.7
51	5.8	71	8.4
43	5.8	76	8.8
44	4.9	49	5.1
63	8.7	33	4.9
49	3.6	54	6.7
42	5.5	65	6.4
40	4.9	39	6.0
47	5.1	52	7.5
41	4.1	45	6.4
41	4.6	58	6.8
56	5.1	67	9.2

	group	age	chol
1	正常组	48.00	3.50
2	正常组	33.00	4.60
3	正常组	51.00	5.80
4	正常组	43.00	5.80
5	正常组	44.00	4.90
6	正常组	63.00	8.70
7	正常组	49.00	3.60
8	正常组	42.00	5.50
9	正常组	40.00	4.90
10	正常组	47.00	5.10

图 1.16 数据格式

该题选自《医学统计学》第二版第 7 章。考虑到统计分析对数据格式(见图 1.16)的要求,我们这里建立三个变量:group 表示组别,age 代表年龄,chol 则表示胆固醇,见数据文件 coanova.sav。

由于协方差分析涉及到一些较深的统计理论,为节省篇幅,我这里只好采用照本

宜科的方法，告诉大家如何作，而不作过多解释，欲进一步了解原理的朋友请参考有关统计学著作。

首先应进行预分析，了解资料是否符合协方差分析的要求，其中最重要的一点就是看 age 的影响在两组中是否相同，这可以用 age 与 group 是否存在交互作用来表示。对该问题，首先应当作分组散点图，肉眼观察两组直线趋势是否近似，然后进行预分析，看交互作用有无统计学意义，当交互作用无统计学意义时，才进行正式分析，计算出修正均数，并比较它们有无统计学差异。这里我们按照最为精确的分析步骤来讲解。

#### 1.4.2 预分析：线性趋势的判断

首先应当进行的是两组间线性趋势的初步判断。

Graph→Scatter→Simple

Y Axis 框: chol

X Axis 框: age

Set Markers by 框: group

OK

散点图中 Y 轴代表胆固醇

X 轴代表年龄

用不同的标记区分正常人和超重者的散点

做出的散点图如图 1.17 所示。

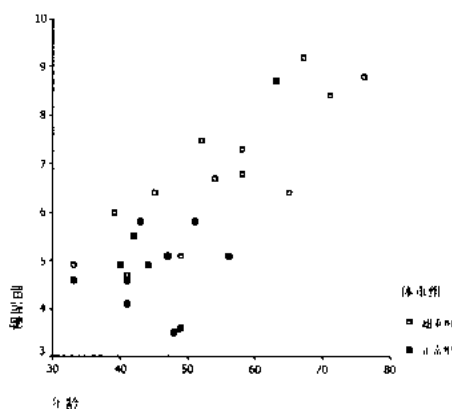


图 1.17 散点图

从图中我们可以得到如下信息：

- ◇ 正常组和超重组的年龄分布基本相同，没有明显的偏差。
- ◇ 两组中年龄和胆固醇都有明显的直线趋势。
- ◇ 两组中直线趋势的斜率相近。

从以上几点，我们初步判断资料符合协方差分析的要求，可以继续分析。

#### 1.4.3 预分析：检验各组总体斜率是否相等

在初步判定资料符合协方差分析的要求后，进一步检验各组总体斜率是否相等。

Analyze→General Linear Model→Univariate

Dependent List 框: chol

Fixed Factor 框: group

Covariates 框: age

Model:

要分析的应变量为胆固醇

固定因素自变量为 group

协变量为 age



<input checked="" type="radio"/> Custom Model 框: group、age、group*age Sum of squares 下拉列表框: Type I <input type="button" value="Continue"/> <input type="button" value="OK"/>	要求自定义方差分析模型 强行纳入交互效应 group*age 采用 I 类方差分析模型
---	---

I 型方差分析模型的结果和变量纳入的顺序有关, 即上面将 group、age 换为 age、group, 分析结果就会改变。此处将 group 放在前面, 即模型重点分析 group 单独对 chol 的影响大小, 类似于单因素分析。模型中交互项的目的是检验 group 处于不同水平时, chol 随 age 变化的斜率是否相等, 因为各组总体斜率相等是协方差分析的条件之一。

Tests of Between-Subjects Effects

Dependent Variable: 胆固醇					
Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	43.002 <sup>a</sup>	3	14.334	14.988	.000
Intercept	916.898	1	916.898	958.733	.000
GROUP	18.615	1	18.615	19.465	.000
AGE	24.380	1	24.380	25.492	.000
GROUP * AGE	6.499E-03	1	6.499E-03	.007	.935
Error	21.040	22	.956		
Total	980.940	26			
Corrected Total	64.042	25			

a. R Squared = .671 (Adjusted R Squared = .627)

上表显示交互作用无统计学意义, 而且 P 值非常大, 因此两组的斜率可以认为相同。Group 具有统计学意义, 但由于此处使用的是 I 型方差分析模型, 该结果仅供参考。

若交互作用项有统计学意义, 需对资料作一定处理, 再作协方差分析, 或者选用其他方法分析。如果还希望重点了解 age 对 chol 的影响, 还可以按 age、group、group\*age 的顺序分析, 此处略。

#### 1.4.4 正式分析: 比较修正均数有无差异

在前面各步均得出资料符合协方差分析条件的结论后, 下面就可以进行正式分析来比较两组的修正均数有无差异。由于前面已得出两组斜率相同的结论, 下面正式分析中就可以将交互项从模型中去除, 然后采用模型 III 给出 group 各水平下 chol 对 age 的修正均数, 并比较两组修正均数有无统计学差异。

Analyze → General Linear Model → Univariate	
Dependent List 框: chol	要分析的应变量为胆固醇
Fixed Factor 框: group	固定因素自变量为 group
Covariates 框: age	协变量为 age
Model:	
<input checked="" type="radio"/> Custom	要求自定义方差分析模型

Model 框: group、age

模型中只分析主效应

Sum of squares 下拉列表框: Type III

采用 III 型方差分析模型

Continue

Options:

Display means for 框: group

☒ Compare mean effect

对两组的修正均数进行比较

Continue

OK

分析结果如下:

#### Tests of Between-Subjects Effects

Dependent Variable: 胆固醇

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	42.995 <sup>a</sup>	2	21.498	23.493	.000
Intercept	1.527	1	1.527	1.668	.209
GROUP	4.458	1	4.458	4.872	.038
AGE	24.380	1	24.380	26.642	.000
Error	21.047	23	.915		
Total	980.940	26			
Corrected Total	64.042	25			

<sup>a</sup>. R Squared = .671 (Adjusted R Squared = .643)

这是正式的统计分析结果, 显示 group 和 age 都对胆固醇含量有影响, P 值分别为 0.038 和小于 0.001。

### Estimated Marginal Means

#### 体重组

##### Estimates

Dependent Variable: 胆固醇

体重组	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
正常组	5.491 <sup>a</sup>	.276	4.919	6.062
超重组	6.386 <sup>a</sup>	.276	5.815	6.958

<sup>a</sup>. Evaluated at covariates appeared in the model: 年龄 = 50.2308.

这是两组的修正均数及相应的可信区间, 显然超重组的胆固醇均值较高。下方的提示表明该修正均数是按年龄为 50.2308 岁的情形计算的。

##### Pairwise Comparisons

Dependent Variable: 胆固醇

(I) 体重组	(J) 体重组	Mean Difference (I-J)	Std. Error	Sig. <sup>a</sup>	95% Confidence Interval for Difference <sup>a</sup>	
					Lower Bound	Upper Bound
正常组	超重组	-.895 <sup>a</sup>	.406	.038	-1.735	-.5.619E-02
超重组	正常组	.895 <sup>a</sup>	.406	.038	5.619E-02	1.735

Based on estimated marginal means

<sup>a</sup>. The mean difference is significant at the .05 level.

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

以上是把年龄转化为相等后,不同组胆固醇值 Y 的修正均数、标准误以及各组修正均数是否相等的假设检验的结果。可见,在扣除了 X 对 Y 的影响之后,两组的胆固醇值的差别确有统计学意义。此处采用的方法是 LSD 法,即所有组均和对照组相比,大家也可根据分析目的在对话框中选用其他两两比较方法。

#### Univariate Tests

Dependent Variable: 胆固醇					
	Sum of Squares	df	Mean Square	F	Sig.
Contrast	4.458	1	4.458	4.872	.038
Error	21.047	23	.915		

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

上表为对修正均数按方差分析法进行的检验,结论和上面完全相同(两种方法等价)。

最后需要指出的是,如果在模型中保留交互项,则方差分析表格的结果如下:

#### Tests of Between-Subjects Effects

Dependent Variable: 胆固醇					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	43.002 <sup>a</sup>	3	14.334	14.988	.000
Intercept	1.202	1	1.202	1.257	.274
AGE	19.053	1	19.053	19.922	.000
GROUP	.247	1	.247	.259	.616
GROUP * AGE	6.499E-03	1	6.499E-03	.007	.935
Error	21.040	22	.956		
Total	980.940	26			
Corrected Total	64.042	25			

<sup>a</sup>. R Squared = .671 (Adjusted R Squared = .627)

可见由于计算了协变量和自变量的交互项,导致 group 的检验结论发生较大偏差,亦即存在和协变量的交互项时,主效应的方差分析结果不能用。但用该模型计算的修正均数比较结论是正确的,有兴趣的朋友可以自己做一下,检验结论和前面相同。



聪明的朋友可能已经想到:这道题用回归过程一样可以完成呀。而且一步就做完了!的确如此,但协方差分析的重点是求出修正均数,其次才是比较,这一目的只能用上面的方法得到。而且用回归过程实际上是省略了预分析的步骤,这有时候可能会导致错误的结论。

## 第2章 征服一般线性模型

### ——General Linear Model 菜单详解（下）

所谓模型，就是用统计语言简化了的现实世界：越贴近实际情况，效果越佳。

——张文彤

在上一章中我们已经系统学习了 Univariate 过程，本章将要学习的是它分别在多个应变量、单个应变量的重复测量以及详细估计随机效应大小这几种情况下的推广模型。

### 2.1 Multivariate 过程

Multivariate 过程用于进行多元方差分析模型的分析，所谓多元方差分析，就是说应变量不止一个，而是两个或两个以上的应变量同时受到自变量的影响。比如要研究某些因素对儿童生长的影响程度，则身高、体重等都可以作为生长程度的测量因子，即都应作为应变量。此时当然可以分别采用单应变量的方差分析模型来一一分析，但是这样做存在两个问题：

1. 如果结果互相矛盾，即有的分析结果有意义，有的无意义时，无法准确判断究竟自变量有无影响。
2. 单应变量的分析结果不能简单的叠加起来向多应变量推广，就如同我们在地面上（二维）认为大地是平的，但实际上在太空中（三维）一看才发现是个球面一样，仅仅进行单应变量的分析会损失相当多的信息，甚至于得出错误的结论。



如果用数学上的术语来描述，在多元方差分析中各指标是否存在差异的检验是同时完成的，它涉及到各应变量的多元联合分布。而多个单指标方差分析时分别采用各自的分布曲线，并不需要同时完成。因此后者使用的信息较少，不能代替前者。许多时候单因素方差分析检验不出差异，但放在多元空间中却能检验出来。

应用多元方差分析的另外一种情形是：所希望研究的特征根本就找不到最佳的测量方式，比如研究人类心理、分析公司财务健康状况，此时根本就不可能找到一个指标来准确测量该特征，而只能得到一系列与之相关的、各自反应某一方面的变量，这样在进行分析时就必须要将它们均作为结果变量纳入分析，以完整、全面地回答研究问题。

除了大家都能想到的多元正态、方差齐以外，多元方差分析还有其他的一些适用条件：首先，它要求各应变量间的确存在一定的关系，这可以从专业或分析目的上来判断，如测量全身多处的骨骼指标都是为了反映骨龄的大小；其次，它在样本规模上有一定的要求，不仅总样本量要较大，各单元格中样本数量也应较大，否则所进行的

多元分析检验效能太低,不容易得出阳性结论。

如果在多元方差分析中只有一个自变量,所作的分析就成为了 Hotelling  $T^2$  检验,它是  $t$  检验向多元情况的推广,由于在操作上和多个自变量时没有任何区别,此处不再单独讲述。

### 2.1.1 引例与界面说明

例 2.1 某药学工艺流程可以生产丹参素,原儿茶醛和没药得碱,影响生产效率的因素有温度和反应时间,现测得各种条件组合下的产量,共 18 条记录(每种条件重复试验两次),请从中筛选出最佳生产条件。

解:该工艺流程同时生产所需的三种产品,因此我们所筛选的最佳生产条件是指的三种产品产量都较高的条件,并非其中一种。显然三者应同时进行检验,即进行多元方差分析。所用数据集已按要求录入为文件 manova.sav,见图 2.1 所示。

	wendu	time	dancen	ercha	meiyude		wendu	time	dancen	ercha	meiyude
1	100	0.5	.91	.06	28.40	10	100	0.5	.93	.06	28.90
2	100	1	3.70	.50	45.40	11	100	1	3.62	.44	45.40
3	100	1.5	5.40	.75	56.40	12	100	1.5	5.40	.75	56.90
4	110	0.5	3.70	.51	52.10	13	110	0.5	3.70	.54	52.50
5	110	1	6.10	.96	54.50	14	110	1	6.20	1.00	54.70
6	110	1.5	4.75	.90	45.20	15	110	1.5	4.64	.8	46.00
7	120	0.5	5.40	.90	54.90	16	120	0.5	5.40	.8	55.40
8	120	1	4.36	.83	46.60	17	120	1	4.42	.83	46.80
9	120	1.5	9.20	1.70	53.90	18	120	1.5	9.10	1.66	54.60

图 2.1 引例所用的数据

由于每个单元格内有两条重复记录,此处亦可以分析交互作用,模型设置用默认的分析全部因素即可。

Analyze→General Lineal model→Multivariate

Dependent List 框: dancen、ercha、meiyude 要分析的一共有三个应变变量

Fixed Factor 框: wendu、time 固定因素为 wendu、time

OK

可见操作和 Univariate 过程并无什么不同。

Multivariate 过程的对话框界面和上一章学过的 Univariate 过程几乎完全相同,只是 Dependent 框可以选入多个应变变量,同时取消了 Random Factor 框而已,所以这里我们不再重复。

### 2.1.2 结果解释

多元方差分析的结果从表面上看起来比较复杂,但实际上它是首先给出多元方差分析的结果,然后分别给出各应变变量单独的方差分析结果,仅此而已。下面我们就具体来看看。



## General Linear Model

Between-Subjects Factors

	Value	Label	N
温度	1.00	100	6
	2.00	110	6
	3.00	120	6
时间(h)	1.00	0.5	6
	2.00	1	6
	3.00	1.5	6

上表列出了所有因素的各个水平及相应的测量次数。可见温度和时间都有三个水平，且测量次数均为 6。

Multivariate Tests<sup>c</sup>

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	1.000	241421.858 <sup>a</sup>	3.000	7.000	.000
	Wilks' Lambda	.000	241421.858 <sup>a</sup>	3.000	7.000	.000
	Hotelling's Trace	103466.511	241421.858 <sup>a</sup>	3.000	7.000	.000
	Roy's Largest Root	103466.511	241421.858 <sup>a</sup>	3.000	7.000	.000
WENDU	Pillai's Trace	1.962	138.591	6.000	16.000	.000
	Wilks' Lambda	.000	583.536 <sup>a</sup>	6.000	14.000	.000
	Hotelling's Trace	2378.320	2378.320	6.000	12.000	.000
	Roy's Largest Root	2352.533	6273.421 <sup>b</sup>	3.000	8.000	.000
TIME	Pillai's Trace	1.368	5.776	6.000	16.000	.002
	Wilks' Lambda	.000	136.162 <sup>a</sup>	6.000	14.000	.000
	Hotelling's Trace	2223.608	2223.608	6.000	12.000	.000
	Roy's Largest Root	2223.024	5928.065 <sup>b</sup>	3.000	8.000	.000
WENDU * TIME	Pillai's Trace	2.355	8.211	12.000	27.000	.000
	Wilks' Lambda	.000	278.949	12.000	18.812	.000
	Hotelling's Trace	3010.693	1421.716	12.000	17.000	.000
	Roy's Largest Root	2802.883	6306.487 <sup>b</sup>	4.000	9.000	.000

<sup>a</sup>. Exact statistic

<sup>b</sup> The statistic is an upper bound on F that yields a lower bound on the significance level.

<sup>c</sup>. Design: Intercept+WENDU+TIME+WENDU \* TIME

上表是针对模型及模型中各因素所作的检验，采用的是四种多元检验方法：Pillai 轨迹、Wilks'  $\lambda$ 、Hotelling 轨迹和 Roy 最大根统计量。一般它们的结果都是相同的，如果不同，一般应当以 Pillai 轨迹的结果为准，因为它就像 t 检验一样，对数据不满足检验假设的前提条件具有一定的耐受性，结果较稳定。从表中可见所用的模型中两个因素及交互作用都有统计学意义，即总的来说它们都对产量有影响。但究竟是它们对三种产品的产量都有影响，还是只对其中某几种有影响，则需要用后面单因素方差分析的结果来说明。

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	丹参素(mg/g)	78.381 <sup>a</sup>	8	9.798	4120.474	.000
	原儿茶醛(mg/g)	3.092 <sup>b</sup>	8	.387	828.262	.000
	没得得产出率%	1219.538 <sup>c</sup>	8	152.442	1294.321	.000
Intercept	丹参素(mg/g)	420.113	1	420.113	176683.215	.000
	原儿茶醛(mg/g)	10.962	1	10.962	23533.762	.000
	没得得产出率%	42885.442	1	42885.442	364121.679	.000
WENDU	丹参素(mg/g)	26.824	2	13.412	5640.458	.000
	原儿茶醛(mg/g)	1.470	2	.735	1575.190	.000
	没得得产出率%	251.858	2	125.929	1069.208	.000
TIME	丹参素(mg/g)	28.478	2	14.239	5988.346	.000
	原儿茶醛(mg/g)	1.059	2	.529	1134.619	.000
	没得得产出率%	138.791	2	69.396	589.208	.000
WENDU * TIME	丹参素(mg/g)	23.079	4	5.770	2426.547	.000
	原儿茶醛(mg/g)	.563	4	.141	301.619	.000
	没得得产出率%	828.889	4	207.222	1759.434	.000
Error	丹参素(mg/g)	2.140E-02	9	2.378E-03		
	原儿茶醛(mg/g)	4.200E-03	9	4.667E-04		
	没得得产出率%	1.060	9	.118		
Total	丹参素(mg/g)	498.515	18			
	原儿茶醛(mg/g)	14.079	18			
	没得得产出率%	44106.040	18			
Corrected Total	丹参素(mg/g)	78.402	17			
	原儿茶醛(mg/g)	3.096	17			
	没得得产出率%	1220.588	17			

<sup>a</sup>. R Squared = 1.000 (Adjusted R Squared = .999)

<sup>b</sup>. R Squared = .999 (Adjusted R Squared = .987)

<sup>c</sup>. R Squared = .999 (Adjusted R Squared = .998)

上表实际上是三个一元方差分析表的合并,即分别考虑三个应变量时的单因素方差分析结果。前面的多元方差分析已经得知两自变量对应变量有影响,从现在的分析表就可以更清楚的知道是对哪些自变量影响较大。对照可知,温度及时间这两个因素以及它们的交互作用对所有三个产品的产量都有影响。



如果在上面的分析中不引入交互项,就会得出没得得碱产量与以上两因素无关的错误结论,这也算是交互作用重要性的又一个实例吧。

### 2.1.3 对引例的进一步分析

上面仅仅分析出温度与时间都对产量有影响,但究竟是哪种水平的组合产量最高?如果交互作用无统计意义,则此处可以直接进行两两比较。但是,当交互作用存

在时，两两比较可能会得出错误结论。此处可以使用 **Options** 子对话框中的功能直接给出各组合条件下的均数表，如下所示。

温度 * 时间(h)						
Dependent Variable	温度	时间(h)	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
丹参素(mg/g)	100	0.5	.920	.034	.842	.998
		1	3.660	.034	3.582	3.738
		1.5	5.400	.034	5.322	5.478
	110	0.5	3.700	.034	3.622	3.778
		1	6.150	.034	6.072	6.228
		1.5	4.700	.034	4.622	4.778
	120	0.5	5.400	.034	5.322	5.478
		1	4.400	.034	4.322	4.478
		1.5	9.150	.034	9.072	9.228
原儿茶醛(mg/g)	100	0.5	8.000E-02	.015	4.544E-02	.115
		1	.470	.015	.435	.505
		1.5	.750	.015	.715	.785
	110	0.5	.525	.015	.490	.560
		1	.990	.015	.955	1.025
		1.5	.815	.015	.780	.850
	120	0.5	.875	.015	.840	.910
		1	.830	.015	.795	.865
		1.5	1.695	.015	1.660	1.730
没食得产出率%	100	0.5	28.650	.243	28.101	29.199
		1	45.400	.243	44.851	45.949
		1.5	56.650	.243	56.101	57.199
	110	0.5	52.300	.243	51.751	52.849
		1	54.600	.243	54.051	55.149
		1.5	45.600	.243	45.051	46.149
	120	0.5	55.150	.243	54.601	55.699
		1	46.700	.243	46.151	47.249
		1.5	54.250	.243	53.701	54.799

可见丹参素和原儿茶醛都是在 120℃、反应 1.5 小时的时候产量最高；没食得碱虽并非如此，但也离最佳产量不远，因此上述条件这就是所寻找的最佳生产条件。

上述分析中也可以采用两两比较进行选择，但应注意其结果只有参考价值，并且此处所做的实际上是单因素方差分析的两两比较，用单因素分析结果外推到多元时不一定准确，需要相当的慎重。尤其是当多元检验结果有差异，而所有单因素分析结果均无差异时，应以多元的结果为准，并通过在三维空间做出均数图的办法来寻找究竟可能的差异来自那些变量，图形化分析在这里又一次显示出了它的强大威力。

## 2.2 Repeated Measures 过程


Repeated Measures 过程是用来进行重复测量的方差分析的，重复测量指的是一个

应变量被重复测量好几次，比如说研究某种药物的疗效，在患者用药后两周、四周、六周、八周时分别测量相应疗效指标，这种情况就是重复测量。由于结果变量在同一个个体内被测量多次，在分析时将更容易发现影响因素的作用。


实际上，如果对普通方差分析模型做出正确的设置，即在自变量中引入时间因素，好像这种数据也可以用单因素模型来做，那为什么还要专门用多元的方法来检验呢？有以下两点原因：

1. 同一个体的不同次测量间往往存在相关性，比如张三第二周的舒张压是 90，那么第四周舒张压估计也就在 90 附近，不会差得太远。但如果去测李四，它的舒张压就可以是 70 或 110 之类所有可能的取值，因为这个数值和张三第二周的舒张压值没有任何关系！而我们进行方差分析时一个非常严格的条件就是各数据应独立，显然，该条件在重复测量时往往被违反。此时用单因素模型来分析，就会得出错误结论，需要按照一定标准对结论进行校正。

如果重复测量数据之间实际上不存在相关性，则多元分析和一元分析的结果应当一致，这种情况被称为数据符合 Huynh-Feldt 条件，而最常用的判断数据是否满足该条件的检验就是球形检验。它的结果在 Repeated Measures 过程的输出结果中非常重要。

 所谓 Huynh-Feldt 条件，就是指同一个体的各次重复测量结果间实际不存在相关性，资料的协方差矩阵成为了一种被称为 H 型协方差的结构。

2. 由于实验设计的要求或数据本身的特点，有时即使数据符合 Huynh-Feldt 条件，我们需要研究的因素采用单因素方差分析模型来研究也是非常困难，甚至是不可能的。下面我们所用引例中的 anxiety 和 tension 两个变量就是典型的例子。

 在重复测量的数据中，除了影响因素外，研究者往往只关心随着时间的流逝，所测量指标的变动趋势如何。而不是各次测量间两两比较的结果如何。因此对不同次测量间它只提供趋势检验，而不给出两两比较结果。

### 2.2.1 引例

例 2.2 在 SPSS 自带数据集 anxiety2.sav 中判断：anxiety 和 tension 对实验结果（即 trial1~trial4）有无影响；四次试验间有无差异；试验次数和两个变量有无交互作用。

解：anxiety2.sav 和 anxiety.sav 实际上是同一个数据，但根据不同的分析方法采用了不同的数据排列方式。如果采用 anxiety.sav 进行分析，我们可以分析四次试验间有无差异的问题，但对另两个问题就无能为力了，因为用普通的方差分析模型，anxiety 和 tension 的影响被合并到了 subject 中，分解起来非常困难。况且四次测量间是否有相关，即资料是否满足 Huynh-Feldt 条件也不清楚，因此，这里就使用重复测量的方差分析模型。

Analyze→General Lineal model→Repeated measures

Within-subject factor name 框：改为 trial

定义重复测量的变量名为 trial  
重复测量的次数为四次

Number of levels 框：键入 4:

Within-subject variables (trial)框：trial1~trial4

trial1~trial4 代表四次测量结果

Between subjects factor 框: anxiety、tension

Model:

☒ Custom

Within-subject Model 框: trial

Between subjects Model 框: anxiety、tension

Continue

OK

分析四次重复测量间有无趋势  
只分析主效应

⚠ 请注意，这里没有选入变量 subject，因为它实际上在这里成为了一个记录 ID，要是将它选入，则什么都检验不了了。

## 2.2.2 界面说明

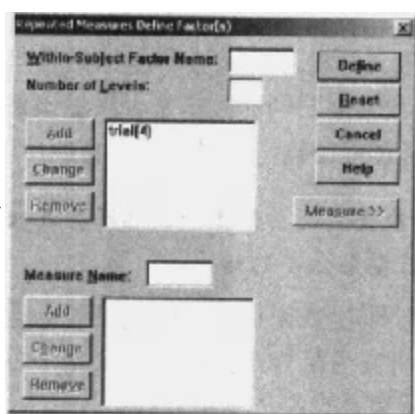


图 2.2 预定义对话框

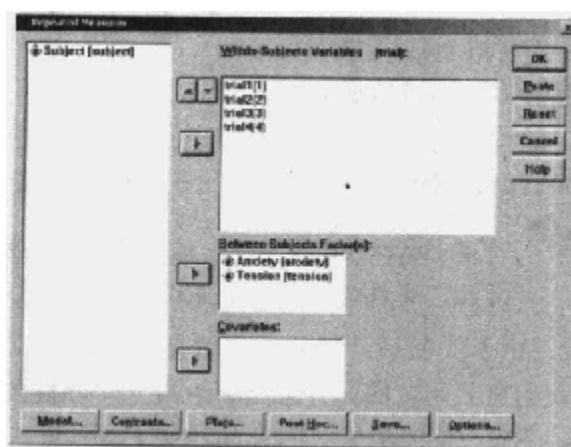


图 2.3 主对话框

【预定义对话框】（见图 2.2）

该对话框用于设定数次重复测量都记录的是哪个变量的数值，以及共重复测量了多少次。

1. Within-subject factor name 框：定义重复测量的变量名，默认为 factor1。

在重复测量的方差分析模型中，同一个变量的所有重复测量被看成是一个组，因此代表重复测量的变量被称为组内因素。相应的，其他因素就被称为组间因素。

2. Number of levels 框：定义上述指标被重复测量了几次。

3. Add 按钮组：用于添加、修改和删除重复测量变量的定义。此处我们单击 Add 按钮，则变量 trial 被加入右侧大框中，括号里的“4”表示被重复测量了四次。可以在大框中加入多个重复测量变量的定义。

4. Measure>>：如果所进行的重复测量有嵌套现象，如对每个病人共重复测量四周的血压，但每次测量均重复三次，此时每次测量的三次重复就嵌套在四周的重复测量中，需要单击该按钮扩展预定义框，在新出现的下部部分定义所嵌套的重复测量变量。



## 【主对话框】(见图 2.3)

这个对话框和我们以前看到的方差分析对话框不太一样：它没有应变量表，而是改为了组内效应框，但实际上是一回事，前面定义了 trial 有四次测量，此处就给出了四个空让你填入相应代表四次测量的变量，选中 trial1~trial4，将其选入；然后要选择自变量了（这里又将其称为 between subjects factor），将剩下的三个都选入即可。

其他子对话框和前面介绍过的内容无甚差异，不再重复。

## 2.2.3 结果解释

## General Linear Model

## Within-Subjects Factors

Measure: MEASURE_1	
TRIAL	Dependent Variable
1	TRIAL1
2	TRIAL2
3	TRIAL3
4	TRIAL4

上表给出了所定义的 4 次重复测量的变量名，在模型中它们都代表一个应变变量 trial，只是重复测量的次数不同而已。

## Between-Subjects Factors

	N	
Anxiety	1	6
	2	6
Tension	1	6
	2	6

这是引入模型的其他自变量的情况列表。

Multivariate Tests<sup>b</sup>

Effect		Value	F	Hypothesis df	Error df	Sig.
TRIAL	Pillai's Trace	.970	75.136 <sup>a</sup>	3.000	7.000	.000
	Wilks' Lambda	.030	75.136 <sup>a</sup>	3.000	7.000	.000
	Hotelling's Trace	32.201	75.136 <sup>a</sup>	3.000	7.000	.000
	Roy's Largest Root	32.201	75.136 <sup>a</sup>	3.000	7.000	.000
TRIAL * ANXIETY	Pillai's Trace	.558	2.949 <sup>a</sup>	3.000	7.000	.108
	Wilks' Lambda	.442	2.949 <sup>a</sup>	3.000	7.000	.108
	Hotelling's Trace	1.264	2.949 <sup>a</sup>	3.000	7.000	.108
	Roy's Largest Root	1.264	2.949 <sup>a</sup>	3.000	7.000	.108
TRIAL * TENSION	Pillai's Trace	.374	1.393 <sup>a</sup>	3.000	7.000	.322
	Wilks' Lambda	.626	1.393 <sup>a</sup>	3.000	7.000	.322
	Hotelling's Trace	.597	1.393 <sup>a</sup>	3.000	7.000	.322
	Roy's Largest Root	.597	1.393 <sup>a</sup>	3.000	7.000	.322

<sup>a</sup> Exact statistic

<sup>b</sup>

Design: Intercept+ANXIETY+TENSION  
Within Subjects Design: TRIAL

上表是针对组内变量 trial, 以及它和另两个引入模型的自变量间的交互作用是否存在统计学意义进行的多元方差分析检验, 此处的结果阅读可以完全参照 Multivariate 过程中的输出来进行。具体采用的是四种多元检验方法: Pillai 轨迹、Wilks'  $\lambda$ 、Hotelling 轨迹和 Roy 最大根统计量。一般它们的结果都是相同的, 如果不同, 应当以 Pillai 轨迹的结果为准。可见在所用的模型中, trial 的四次测量间的确是存在着统计学差异的, 它和另两个变量间的交互作用无统计学意义。不过这里的分析结论究竟算不算数还要看了下面的球形检验结果才能知道。

Mauchly's Test of Sphericity<sup>b</sup>

Measure: MEASURE_1							
Within Subjects Effect	Mauchly's W	Approx Chi-Square	df	Sig.	Epsilon <sup>a</sup>		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
TRIAL	.297	9.383	5	.097	.557	.821	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

<sup>a</sup> May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

<sup>b</sup> Design: Intercepts\*ANXIETY\*TENSION  
Within Subjects Design: TRIAL

上表是非常重要的球形检验结果, 可见近似卡方为 9.383, 自由度为 5, P 值为 0.097, 因此 trial 四次重复测量的数据满足直接进行一元方差分析的 Feldt-Huynh 条件, 在下面的分析中可以不进行校正。而且一旦多元分析结果和一元结果发生了冲突, 应当以一元的为准。

表格右侧则给出了如果需进行校正时所采用的三种校正方法各自所需的校正系数 Epsilon 的大小, 这些数值仅供参考, 因为后面的检验结果已经自动使用它们进行了校正, 无需我们再费心了。

Tests of Within-Subjects Effects

Measure: MEASURE_1						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Sphericity Assumed	991.500	3	330.500	137.481	.000
	Greenhouse-Geisser	991.500	1.671	593.254	137.481	.000
	Huynh-Feldt	991.500	2.464	402.449	137.481	.000
	Lower-bound	991.500	1.000	991.500	137.481	.000
TRIAL * ANXIETY	Sphericity Assumed	8.417	3	2.806	1.167	.341
	Greenhouse-Geisser	8.417	1.671	5.036	1.167	.329
	Huynh-Feldt	8.417	2.464	3.416	1.167	.338
	Lower-bound	8.417	1.000	8.417	1.167	.308
TRIAL * TENSION	Sphericity Assumed	12.167	3	4.056	1.687	.193
	Greenhouse-Geisser	12.167	1.671	7.280	1.687	.219
	Huynh-Feldt	12.167	2.464	4.938	1.687	.204
	Lower-bound	12.167	1.000	12.167	1.687	.226
Error(TRIAL)	Sphericity Assumed	84.917	27	2.404		
	Greenhouse-Geisser	84.917	15.04	4.316		
	Huynh-Feldt	84.917	22.17	2.928		
	Lower-bound	84.917	8.000	7.213		

Tests of Within-Subjects Effects 表是采用一元方差分析的方法对组内因素进行了检验, 此处的结果应当和前面多元检验的结果结合起来看, 互为补充。注意第一种为球形分布假设成立时的结果, 就是在本例中所要看的。如果该假设不成立, 则应当以前面多元检验的结果为准, 并采用下面三种校正方法的结果加以参考, 一般推荐采用 Greenhouse-Geisser 的校正结果。



Huynh-Feldt Epsilon 是在 Greenhouse-Geisser 的基础上又进行了修正, 但它对于小样本则太过保守。Lower-bound Epsilon 值则是校正系数可信区间的最小值, 属于极端保守的结果。

Tests of Within-Subjects Contrasts 表输出的是非常重要的一部分信息: 各次重复测量间变化趋势的分析。这里要求检验没有统计学意义, 否则说明变化趋势不服从该曲线。以 trial 为例, 对 Linear 的检验 P 值小于千分之一, Quadratic 的 P 值略大于 0.05。只有 Cubic 的 P 值在 0.5 附近, 因此最佳的拟合曲线应为 Cubic (三次方曲线); 但由于一共才四次测量, 三次方曲线显然太奢侈了, 因此如果没有任何其他提示或专业的知识, 最终的拟合曲线应为 Quadratic (二次方曲线)。

Tests of Within-Subjects Contrasts

Measure: MEASURE_1						
Source	TRIAL	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Linear	984.150	1	984.150	214.118	.000
	Quadratic	6.750	1	6.750	3.758	.085
	Cubic	.600	1	.600	.731	.415
TRIAL * ANXIETY	Linear	1.667	1	1.667	.363	.562
	Quadratic	3.000	1	3.000	1.670	.228
	Cubic	3.750	1	3.750	4.571	.061
TRIAL * TENSION	Linear	10.417	1	10.417	2.266	.166
	Quadratic	8.333E-02	1	8.333E-02	.046	.834
	Cubic	1.667	1	1.667	2.032	.188
Error(TRIAL)	Linear	41.367	9	4.596		
	Quadratic	16.167	9	1.796		
	Cubic	7.383	9	.820		

Tests of Between-Subjects Effects

Measure: MEASURE\_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	4800.000	1	4800.000	265.710	.000
ANXIETY	10.083	1	10.083	.558	.474
TENSION	8.333	1	8.333	.461	.514
Error	162.583	9	18.065		

上表为组间效应的方差分析结果，可见 anxiety 和 tension 均无统计学意义。

## 2.2.4 对引例的进一步分析

### 【重复测量均数的变化趋势】

前面我们在趋势检验中得出结论，认为四次测量间呈二次曲线趋势，但是，统计检验的结果在变化趋势的确定中只能起辅助作用，更加重要的方法是做出均数图，直接观察均数的变化情况如何。因此我们另外用 plots 子菜单做出模型估计的四次测量均数值如图 2.4 所示。

可见四次测量均数实际上还是近似于直线趋势的，因此前面的均数变动趋势应以线性为最佳。

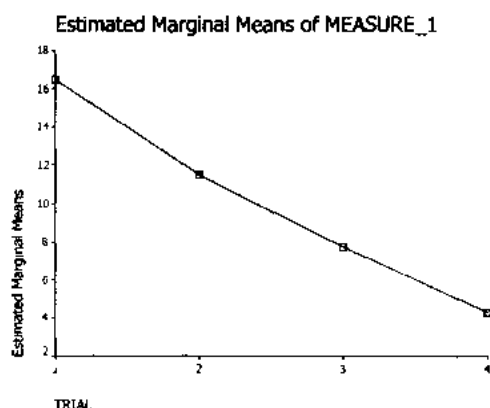


图 2.4 四次测量的均数图

### 【和直接应用一元方差分析模型的结果比较】


前面我们得知数据满足 Feldt-Huynh 条件，可以直接用一元方差分析模型来分析。那么本例直接分析的结果如何呢？文件 anxity.sav 就是按这种分析格式排列好的数据。对它进行分析，结果见下表：

Tests of Between-Subjects Effects

Dependent Variable: Score					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1193.083 <sup>a</sup>	20	59.654	24.811	.000
Intercept	4800.000	1	4800.000	1996.406	.000
SUBJECT	162.583	9	18.065	7.513	.000
TRIAL	991.500	3	330.500	137.461	.000
ANXIETY	.000	0			
TENSION	.000	0			
TRIAL * ANXIETY	8.417	3	2.806	1.167	.341
TRIAL * TENSION	12.167	3	4.056	1.687	.193
Error	64.917	27	2.404		
Total	6058.000	48			
Corrected Total	1258.000	47			

<sup>a</sup> R Squared = .948 (Adjusted R Squared = .910)

可见由于 Anxiety 和 Tension 是个体的特征值, 当引入变量 Subject 后它们的自由度为 0, 无法进行分析。但 trail 及交互作用的分析结果和前面完全一致。在本例中, 重复测量方差分析模型的优势一目了然。

 当然在上面的分析中可以不引入变量 Subject, 此时就可以对 Anxiety 和 Tension 进行检验了, 但这样明显没有充分利用样本信息, 结论不准确。

## 2.3 Variance Components 过程

前面我们多次遇到过随机因素, 但模型在拟合时仅仅估计出了它的变异大小, 并未进一步深入分析。实际上在许多时候这样做不太合适, 比如要研究某个城市中小学五年级学生的学习成绩和哪些因素有关, 随机在几百所小学中抽取了十几所作为样本。显然, 不同小学的教育质量不同, 同一个学校的学生在学习成绩上有更多的相似性, 不同学校的成绩变异程度也不同, 传统模型没有对这些问题进行考虑, 在数据组内聚集性较强时可能会得出错误结论。

Variance Components 过程用于计算方差成分, 它可以将总变异精细的分解到不同层次的不同变量上, 是一种可以考察各个层次因素的变异大小, 从而为哪些层次上可能存在组内聚集性、如何可能减小数据变异提供信息的统计方法。但是, 该过程实际上只是个预处理过程, 本身的功能非常有限, 只能做初步的分析, 为进一步处理提供线索。

实际上, 方差成分模型是近十余年来发展起来的专门用于分析层次结构数据的多水平统计模型(混合效应模型)的最简形式, 该模型涉及到非常专业的统计知识, 这里我们不再深入。在下一章中我们将对它进行专门讲解。

### 2.3.1 引例

例 2.3 现测量了四个家庭共 18 个个体的高度以及性别, 请分析不同家庭、性别间身高的变异情况(见图 2.5)。

	family	gender	height		family	gender	height
1	1	F	67	10	2	M	68
2	1	F	66	11	2	M	70
3	1	F	64	12	3	F	63
4	1	M	71	13	3	M	64
5	1	M	72	14	4	F	67
6	2	F	63	15	4	F	66
7	2	F	63	16	4	M	67
8	2	F	67	17	4	M	67
9	2	M	69	18	4	M	69

图 2.5 引例数据

解: 该数据见文件 mixed.sav, 来自高慧璇《SAS/STAT 软件使用手册》342 页, 之所以要采用该数据, 是为了便于大家和书上 SAS 的分析结果相对照。此处 18 个人的身高分别来自四个家庭, 显然家庭是随机因素, 并且身高在不同的家庭中可能有聚集



性，下面就用 Variance Components 过程来估计相应的方差大小。

Analyze→General Linear model→Variance Components

Dependent Variable 框: height

Fixed Factor 框: gender

Random Factor 框: family

Options:

☒ Restricted maximum likelihood

Continue

OK

Gender 为固定因素

Family 为随机因素

要求用 REML 法估计

### 2.3.2 界面说明

#### 【主对话框】

Variance Components 过程的主对话框和 Univariate 过程没有什么区别，这里不再详细介绍，只是有一点需要指出：由于需要估计方差大小，模型中必须要指定至少一个随机因素变量方可运算。

#### 【Model】子对话框

同样和 Univariate 过程没有什么区别，这里不再重复。



图 2.6 Options 子对话框

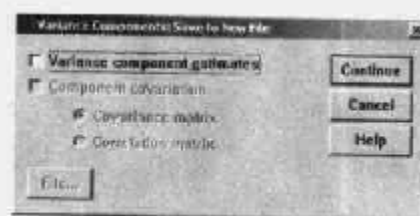


图 2.7 Save 子对话框

#### 【Options】子对话框（见图 2.6）

1. Method 单选框组：用于指定拟合方差成分模型时所用的方差估计方法。

✧ MINQUE：默认值，进行最小正态二乘无偏估计（minimum normal quadratic unbiased estimator）。在变量为正态分布时，使用此方法作方差估计要比其他方法得到的方差小。

✧ ANOVA：用方差分析法估计方差大小，该方法可使用 I 或 III 型方差分析模型进行估计，实际上和我们用普通的方差分析模型来估计是一回事。如果数据组内聚集性明显，则误差较大，甚至会出现方差估计值为负的现象。

✧ Maximum likelihood：用最大似然法对方差做估计，对固定效应做估计时不考虑自由度。

✧ Restricted maximum likelihood：采用限制最大似然法（REML）对方差做无偏估计，前面的三种估计方法实际上都是单水平模型的传统估计方法，REML 是真正按照多水平模型的情况开发的算法，大多数情况下都更精确。在 SAS 的 Mixed 过程中，默认的拟合方法也是 REML。我推荐大家尽量使用此法。

2. Random Effect Priors 单选框组：只在选中 MINQUE 方法时可用，用于对该方

法进行精确设定，可以定义所有残差项均相等（Uniform），也可以假定随机效应方差都为零（Zero）。

3. Sum of Square 单选框组：只在选中 ANOVA 方法时可用，用于选择是使用 I 型还是 III 型方差分析模型进行估计。

4. Criteria 框组：只在选中 ML 或 REML 方法时可用，在其中指定收敛标准（Convergence）和最大重复迭代次数（Maximum iterations）。

5. Display 复选框组：用于设置需要输出的其余一些统计量。

✧ Sum of square：输出方差。

✧ Expected mean squares：输出期望均方，它和上面的 Sum of square 只在分析方法为 ANOVA 时可用。

✧ Iteration history：给出每 n 次迭代时方差成分估计的结果，下方框内用于定义 n 的大小。该复选框只在选中 ML 或 REML 方法时可用。

【Save】子对话框（见图 2.7）

在该对话框中可以将方差成分的估计值以及协方差矩阵存储到指定的数据文件中，以便做进一步的统计分析。

### 2.3.3 结果解释

## Variance Components Estimation

首先仍然应当是各因素取值水平的列表，此处略去。

Variance Estimates	
Component	Estimate
Var(FAMILY)	2.401
Var(FAMILY * GENDER)	1.766
Var(Error)	2.167

Dependent Variable: HEIGHT  
Method: Restricted Maximum Likelihood Estimation

以上表格中给出了各个因素及交互作用的方差分量估计，由于在 SPSS 中该表格不给出各方差分量的标准误以及检验结果，无法判断这些变异有无统计学意义，只能根据大小来认为家庭导致的变异和随机误差相比相差不大，因此可以估计数据在家庭层次上可能不存在聚集性。交互作用导致的变异也不明显。

Asymptotic Covariance Matrix			
	Var(FAMILY)	Var(FAMILY * GENDER)	Var(Error)
Var(FAMILY)	11.220	-2.684	-1.478E-02
Var(FAMILY * GENDER)	-2.684	5.591	-.448
Var(Error)	-1.478E-02	-.448	.932

Dependent Variable: HEIGHT  
Method: Restricted Maximum Likelihood Estimation

上表为各因素的对称协方差矩阵。主对角线上即各因素的 MS（均方，也就是方差），可见 family、family\*gender 和随机误差的 MS 分别为 11.22、5.591 和 0.932，如果采用 Univariate 过程分析，它们的值分别为 12.392、6.059 和 2.100。可见由于普通模型将原本应当分别分配入家庭层次和个体层次的变异一律都归并入个体层次，导致了随机误差不能被精确的分解出来。同时在固定效应参数的估计上也略有偏差。

## 第3章 混合线性模型入门

### ——Mixed Model 菜单详解

Multilevel models are tools to be used with care and understanding

——H. Goldstein

混合线性模型是混合效应模型中最简单的一种，而后者又有许多的别名，根据所从事的领域和师承关系，它又可能被称为多水平模型（Multilevel models，我就是这一派的）、广义估计方程（GEEs）等。那么，这个东东究竟有什么用呢？

我们工作中遇到的许多资料都具有层次结构。例如在市场研究的抽样调查中，受访者会来自不同的城市，这就形成了一个层次结构，高层为城市，低层为受访者。显然，同一城市内的受访者在各方面的特征应当更加相似。又如在几个随机选择的中心或组进行的临床试验，中心间的医疗水平不同，在相同中心的病人也往往比从一般总体中随机抽取的个体趋向于更相似。换言之，所谓层次是指基本的观察单位聚集在更高层次的不同单位中，如同一城市的受访者特征间具有相关性，同一中心的病人数据具有相关性等。传统模型没有对这些问题进行考虑，在数据组内聚集性较强时可能会得出错误结论。

另一方面，在传统的统计分析方法中，对集中趋势（均数）的分析方法已经发展到了比较完善的地步，但对于离散趋势的分析则还处于正在起步的阶段。即我们可以准确的推断是哪些因素对应变量的均数有影响；却无法分析是哪些因素对应变量的变异程度有影响。这一问题现在越来越受到重视，已成为统计理论的一个重要发展方向。

混合效应模型是八十年代初针对资料的层次结构而发展起来的一类模型，它充分考虑到了数据聚集性的问题，可以在数据存在聚集性的时候对影响因素进行正确的估计和假设检验；不仅如此，它还可以对变异的影响因素加以分析，即哪些因素导致了数据间聚集性的出现，哪些又会导致个体间变异的增大。由于该模型成功解决了长期困扰统计学界的数据聚集性问题，二十年来得到了飞速的发展。各大权威统计软件纷纷开始提供这一分析方法，SPSS 11.0 版中也新增了这一模块，借助 SPSS 简洁的操作界面，无疑将使得多水平模型的操作性大为改观，真正做到“飞入寻常百姓家”。

由于该模型非常专业，本章仅仅是入门介绍，只涉及到了最简单的混合线性模型中最简单的内容。希望深入学习的朋友可以参考 H. Goldstein 教授所著的《多水平统计模型》（第二版）一书，该书现已译成中文出版。

### 3.1 模型简介

混合效应模型要比大多数统计模型都复杂的多，为了能够使人家易于入门，这里我们将使用该领域内极为经典的 JSP 数据作为引子，向大家展示该模型的基本结构。

JSP.sav 是“初级学校项目”(Junior School Project, JSP)的一部分,它共包含了伦敦 65 所初级学校中共 4059 名学生的数据,文件中包括了如下变量:

- ◇ School: 学生所在学校的代码,取值为 1~65。
- ◇ Student: 每个学生在学校内的 ID 号。
- ◇ Nomexam: 学生在 16 岁时的考试成绩,已进行了标准正态变换。
- ◇ Standlrt: 学生在 11 岁时的考试成绩,已进行了标准正态变换。
- ◇ Gender: 学生的性别。
- ◇ Schgend: 学校的类型,1 为男女混合,2 为男校,3 为女校。
- ◇ Avslrt: 各个学校学生 11 岁考试成绩的平均数,已进行了标准正态变换。

⚠ 由于混合效应模型非常的复杂,任何“轻微”的干扰都可能对结果变得面目全非,因此在拟和前对变量进行标准正态变换几乎是必需的工作。

### 3.1.1 模型入门

现在假设我们的分析目的是想以 11 岁的成绩为自变量建立针对 16 岁时成绩的回归方程,则在考虑到学校间差异的影响时,使用一般线性模型建立的回归模型如下:

$$y_{ij} = \alpha + \beta_1 \text{standlrt}_{ij} + \sum \beta_j \text{school}_j + e_{ij}$$

$\beta_1$  代表了 11 岁成绩的影响,后面的  $\beta_j$  则表示了第  $j$  个学校的效应,  $e_{ij}$  为第  $j$  所学校第  $i$  个学生的随机误差,被假定为服从均数为 0 的正态分布。

⚠ 请大家注意,在混合效应模型中,下标的使用顺序和一般线性模型恰恰相反!以前我们说  $y_{ij}$  代表了第  $i$  所学校第  $j$  个学生的数值,现在则为第  $j$  所学校第  $i$  个学生的数值,即  $i$  代表了最小的观察单位(学生), $j$  代表高一级的观察单位(学校),如果有更高层次(如城市),则会以  $k$  来代表,以此类推。为统一起见,我们在本章中都会这样定义模型。

上面这个方程看起来没有什么问题,但如果换一个角度来思考,就会发现它忽略了许多深层次的信息。首先我们来看分单独一所学校时的情况,以第一所学校为例,其散点图 3.1 所示,如果对其拟合回归方程,则模型如下:

$$y_i = \alpha + \beta_1 \text{standlrt}_i + e_i$$

其中下标  $i$  代表第  $i$  个学生。在单独考虑这一所学校时,该模型是非常完善的,但当同时考虑多所学校时问题就出现了。显然,各个学校的教学水平是有差异的,也就是说同一所学校学生的成绩之间实际上并不独立,好学校的学生成绩会普遍好一些,差学校的成绩会普遍差一些。

现在再加入其他几所学校,以前三所学校为例,其散点图及相应的回归线如图 3.2 所示。显然,三条回归线的截距不同,这种差异实际上反映了学校间教学水平的差

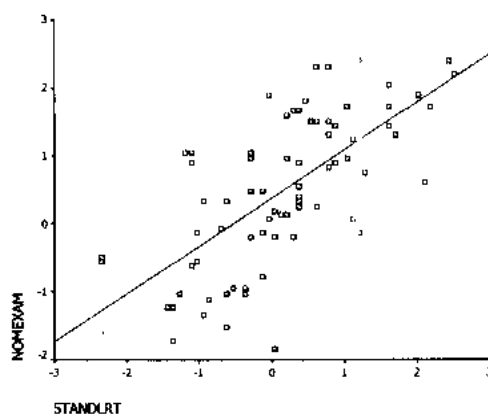


图 3.1 第一所学校的散点图

异, 即 11 岁成绩相同的学生, 在不同的学校中学习后其 16 岁成绩的平均估计值可能是不同的。如果考虑到该变异, 则刚才的模型应扩展为如下形式:

$$y_{ij} = (\alpha_0 + u_{0j}) + \beta_1 s \tan dlrt_{ij} + e_{ij}$$

$y_{ij}$  代表了第  $j$  所学校第  $i$  个学生 16 岁成绩的大小。其中的  $u_{0j}$  就表示了不同学校间截距的变异。如果样本中学校个数  $m$  不多, 且研究者的兴趣就在样本中的这几所学校, 则可以将学校看成是一个固定因素, 用一组  $m-1$  个哑变量表示, 拟合  $m+1$  个参数来进行分析, 此时  $u_{0j}$  就是对应于那  $m-1$  个哑变量的参数, 该模型实际上就是刚才列出的最常用的模型公式:

$$y_{ij} = \alpha + \beta_1 s \tan dlrt_{ij} + \sum \beta_j school_j + e_{ij}$$

但是, 如果我们不仅只关注这儿所学校, 而是关注更广泛的学校总体, 那么就需要估计在学校总体中截距的变异有多大, 此时实际上是将原来的  $\alpha$  真正当作一个随机变量来看待, 令其为  $\alpha = \alpha_0 + u_{0j}$ , 这样, 通过检验  $u_{0j}$  是否为 0 (具体方法为检验其方差是否大于 0), 我们就可以得知这种截距的变异在学校总体中是真的存在, 还是仅仅因为抽样导致的假象而已。此时模型需要同时估计随机项  $u_{0j}$  的大小, 就成为了一个真正的混合效应模型。

下面我们来观察更多的数据, 图 3.3 是前 10 所学校各自的回归线, 从中我们可以看到除了截距以外, 各回归线的斜率间也不尽相同。也就是说, 成绩在学校间的聚集性除了表现为成绩的平均水平不同外, 还表现在不同学校中成绩的离散度上! 斜率高的学校其 16 岁成绩离散度较高, 斜率低的则成绩比较集中。同上, 模型将被继续扩展如下:

$$y_{ij} = (\alpha_0 + u_{0j}) + (\beta_1 + u_{1j}) s \tan dlrt_{ij} + e_{ij}$$

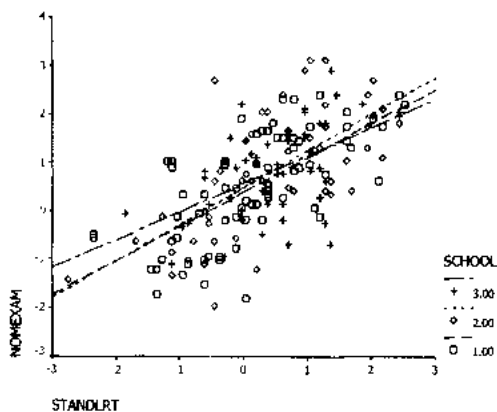


图 3.2 前三所学校的散点图和回归线

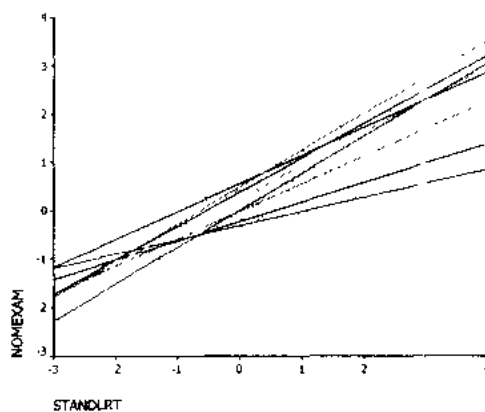


图 3.3 前十所学校的回归线

同理, 通过检验随机项  $u_{1j}$  是否等于 0, 我们就可以得知是否各个学校的成绩离散度不同。更一般的, 模型中的随机项常常被写在一起, 如下所示:

$$y_{ij} = (\alpha_0 + \beta_1 s \tan dlrt_{ij}) + (u_{0j} + u_{1j} s \tan dlrt_{ij} + e_{ij})$$

上式中的两部分分别被称为固定部分和随机部分, 可见和普通的线型模型相比, 混合线性模型主要是对原先的随机误差进行了更加精细的分解。但正因如此, 该模型



就可以正确估计并分析数据在高水平单位内聚集的问题，同时可以为研究者提供更加丰富的信息。

### 3.1.2 混合效应模型的用途

前几年在杂志上看到过这样一件事：大家都知道皮鞋比较捂脚，如果能解决这一问题，则产品一定会非常受欢迎。一位日本人就做了一项发明，并申请了专利，技术改进的核心就是将皮鞋鞋底的前半部分去掉，这样脚不是在密闭环境中，自然也不会捂了。该发明好是好，就是有一个非技术性的问题没能解决——如果一脚踩到了狗屎上可该怎么办？

我说这件事的目的是什么呢？再好的发明创造，如果它不具有实用性，那么就和垃圾（文字垃圾或精神垃圾）没有什么区别。混合效应模型这么的复杂，那么它具有实用性吗？当然是有的，具体来说有以下几点：

1. 对固定效应参数进行更准确的估计：由于在模型的设置上就考虑到了数据的聚集性问题，并采用了相应的迭代方法加以拟和，混合效应模型可以获得回归系数的有效估计，并且可以提供正确的标准误，从而假设检验的结果也更加准确。一般来说，它比传统方法更“保守”，后者的标准误是通过简单地忽略聚集的存在而获得，往往并不准确。这一问题在聚集性较强的时候更为明显。

2. 对变异的影响因素加以研究：传统模型对离散趋势的估计、推断及影响因素研究几乎是无能为力的，而这正是多水平模型的特长。通过对模型的精细设置，研究者可以探讨数据的变异究竟是否在高层次中存在聚集性，而这种变异间的差异又是由哪些变量的不同所导致的。从而可为控制某些数据的离散度提供线索。

3. 重复测量资料的分析及规律探讨：传统模型也可以对重复测量资料加以分析，但是，一般而言，这些方法需要数据是平衡的，即要求每一个体有相同次数的重复测量值。但在实践上，测量次数常是不规则的，有的个体有很多测量值，而有的可能只有一个。此时传统模型的估计可能有误。而混合效应模型通过将这种数据看成为一般的两水平结构（单次测量为低水平，个体为高水平），从而轻车熟路的应用标准的多水平模型技术处理任何测量模式的数据，并提供无偏的参数估计。另一方面，如果重复测量的是生长数据类型的数据，则在多水平分析框架中，每一个体可以具有它们自身的生长曲线，从而可以在更精细的程度上探讨生长发育的一般规律及个体变动情况。

## 3.2 Linear 过程

Linear 项是现在 Mixed Model 菜单中提供的唯一一个过程，用于拟和混合线性模型。它是混合效应模型中最简单的一个，但仍然非常的复杂，因此本节中我们仍将使用 JSP 数据来演示该过程的基本用法。

### 3.2.1 引例与界面说明

现在我们开始对 JSP 数据进行分析，和普通模型不同，混合线性模型中可以只纳入常数项，即模型最简式可以如下：

$$nomexam_{ij} = (\alpha_0) + (u_{0j} + e_{ij})$$

通过检验随机项  $u_{0j}$  是否等于 0, 我们就可以得知是否成绩在学校内存在聚集性。

对话框操作如下:

Analyze → Mixed Models → Linear

Subjects 框: school

选入可认为相互独立的高水平指示变量

Continue

Dependent Variable 框: nomexam

应变量为 nomexam

Statistics:

☒ Parameter estimates

要求给出固定部分参数的估计值

☒ Tests for covariance parameters

要求对随机部分参数进行检验

Continue

Random:

Random Effects: ☒ Include Intercept

要求在随机效应中包含常数项

Subject Groupings: Combinations 框:

指定学校间的随机效应是大小不同的

school

Continue

OK

上面操作中用到了 Linear 过程的大多数对话框, 下面对其加以详细解释:

【预定义对话框】(见图 3.4)

前面提到过混合效应模型的最大特点就是可以正确分析观察单位间数据不独立的问题。预定义对话框就是用于设定哪些观察单位间的数据是相关的, 具体方式是设定相应的指示变量, 该变量取值相同时, 则认为这些个体间数据不独立。

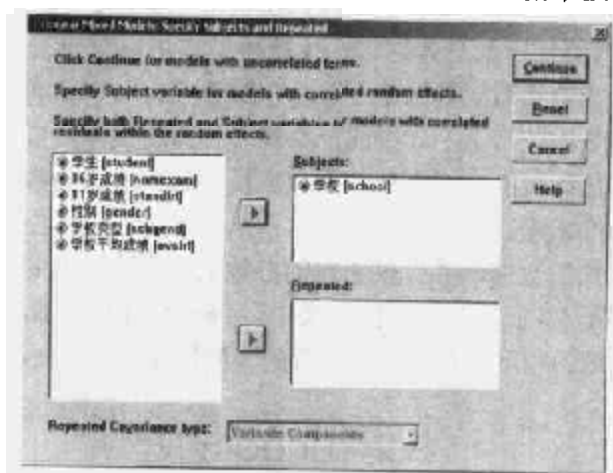


图 3.4 预定义对话框

1. **Subjects** 框: 选入可被认为相互独立的观察单位的指示变量。例如在血压的重复测量中, 同一个体的血压测量值之间是不独立的, 而不同个体的测量值相互独立, 因此个体的指示变量就应被选入。在引例中, 由于处在相同的教育环境中, 同一学校学生的成绩是相关的, 而不同学校间可认为相互独立, 因此这里选入 school。

2. **Repeated** 框: 用于选入重复测量的指示变量。例如对血压重复进行了四次测量, 则指示第几次测量的变量应选入该框, 在引例中不存在这种情况, 所以未加选择。

3. Repeated Covariance type 下拉列表：用于定义残差的协方差结构。可选的有等尺度、一阶自回归、成份对称、Huynh-Feldt 球形阵、无结构、方差成份六种，一般选择默认的方差成份即可，更复杂的情况请参考相关的统计专著。

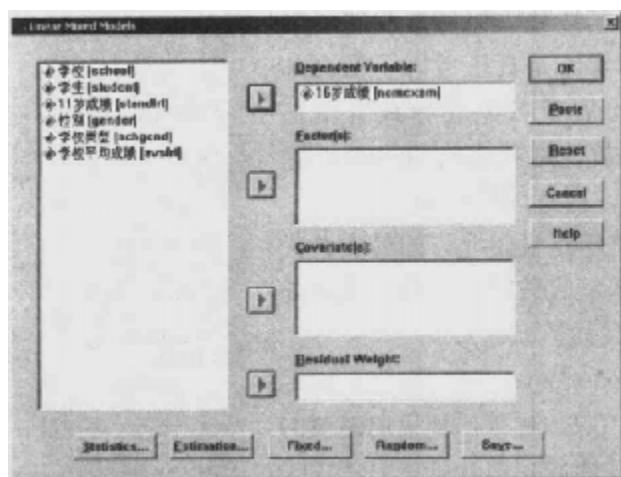


图 3.5 主对话框

#### 【主对话框】（见图 3.5）

主对话框和一般线性模型的非常相似，实际上就是直接的扩展。应变量框、因素框、协变量框都是相同的，注意这里不再区分固定因素和随机因素，这是因为随机效应需要在 Random 子对话框中被详细定义。下方增加了一个 Residual Weight 框，用于选入残差权重变量，类似于以前的 WLS Weight 框。

#### 【Statistics 子对话框】（见图 3.6）

用于选择可以输出的统计量。

1. Summary Statistics 复选框组：可给出常用的描述统计量和纪录处理情况汇总，后者包括所有因素、重复测量变量、重复测量个体、随机效应排序的频数。

2. Model Statistics 复选框组：可列出下列统计指标。

- ✧ Parameter estimates: 固定部分的参数估计值及近似的标准误。
  - ✧ Tests for covariance parameters: 随机部分中协方差参数的近似标准误和 Wald 检验结果，注意这里的检验结果都是近似的，仅供参考，详后。
  - ✧ Correlations of parameter estimates: 固定效应参数估计值的渐进相关阵。
  - ✧ Covariances of parameter estimates: 固定效应参数估计值的渐进协方差阵。
  - ✧ Covariances of random effects: 随机效应参数估计值的渐进协方差阵。
  - ✧ Covariances of residuals: 残差的协方差阵。
  - ✧ Contrast coefficient matrix: 列出用于检验固定效应和自定义假设的设计矩阵。
3. Confidence interval 框：用于设定可信区间的范围，默认为 95%。

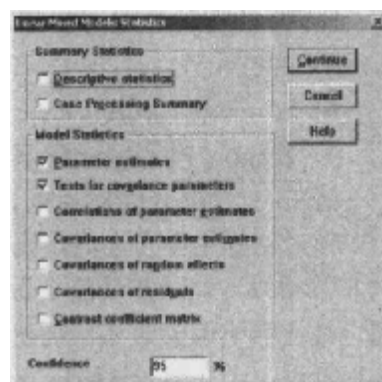


图 3.6 Statistics 子对话框

## 【Estimation】子对话框 (见图 3.7)

用于选择迭代方法, 设定收敛标准等。

1. Method 单选框组: 选择拟合时所采用的迭代方法, 有限制性最大似然法和普通最大似然法两种, 前者才是真正考虑了混合效应的迭代方法, 而后者可能会导致有偏估计, 因此请大家尽量使用默认的 REML, 除非 REML 未能收敛, 再考虑使用 ML。

2. Iterations 框组: 设定允许的最大迭代次数, 当 REML 未能收敛时, 可考虑在此处加以修改, 增加迭代次数。

3. Log-likelihood Convergence 框组: 设定迭代时的对数似然值收敛标准, 如果似然值的绝对或相对改变量小于设定值, 则认为迭代收敛。如果设定值为 0, 则该标准无效。

4. Parameter Convergence 框组: 设定迭代时的参数估计值收敛标准, 如果参数估计值的绝对或相对改变量小于设定值, 则认为迭代收敛。如果设定值为 0, 则该标准无效。

5. Hessian Convergence 框组: 设定迭代时的 Hessian 阵收敛标准, 如果基于 Hessian 阵的统计量绝对或相对改变量小于设定值, 则认为迭代收敛。如果设定值为 0, 则该标准无效。

6. Maximum scoring steps 框: 设定最多在多少次迭代中使用 Fisher 得分算法。

7. Singularity tolerance 下拉列表: 设定检验奇异值的标准。

## 【Fixed Effects】子对话框 (见图 3.8)

该对话框用于设定混合线性模型中的固定效应部分, 内容和以前的 【Model】子对话框非常相似, 只是在功能上做了一些扩展。请注意这里不提供缺省设置, 因此当存在固定效应影响因素时必须要进行正确设定。在引例中由于没有任何影响因素, 所以没有加以设定 (只纳入常数项)。

1. Fixed Effects 框组: 用于对模型中的固定效应部分进行精确设定, 上方的单选框组用于选择将要纳入的是普通效应还是嵌套效应; 中部的框组大家非常熟悉, 用于固定效应的具体设定, 只是现在需要使用下部的 【Add】按钮来加入; 下方的按钮组和 Build Term 框用于效应的精确设置, 其中 【By\*】按钮用于设定交互作用, 而 【Within】按钮用于设定嵌套效应, 在上部选择了 Build nested terms 单选框时才可用。我们将在以后的实例中演示其用法。

2. ☒ Include intercept: 用于选择是否在模型固定效应中包括截距 (即固定效应中的常数项), 不用改动, 默认即可。

3. Sum of squares 下拉列表框: 用于选择方差分析模型进行变异分解的方法, 有 I 型到 IV 型四种, 如果你搞不清它们之间的区别, 请参见相关内容, 或干脆只使用默认的 III 型即可。

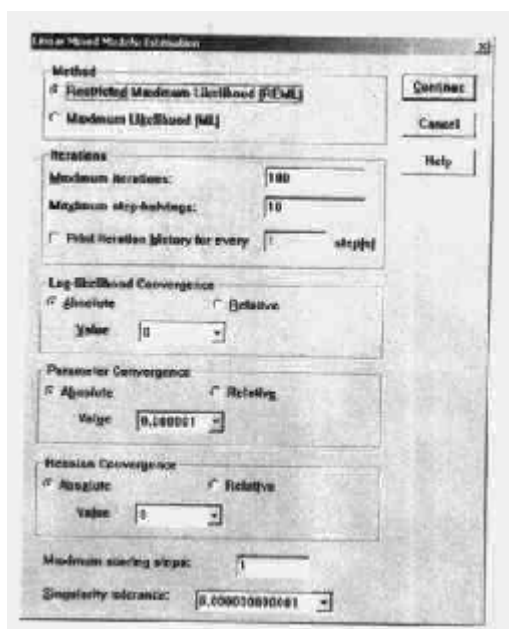


图 3.7 【Estimation】子对话框



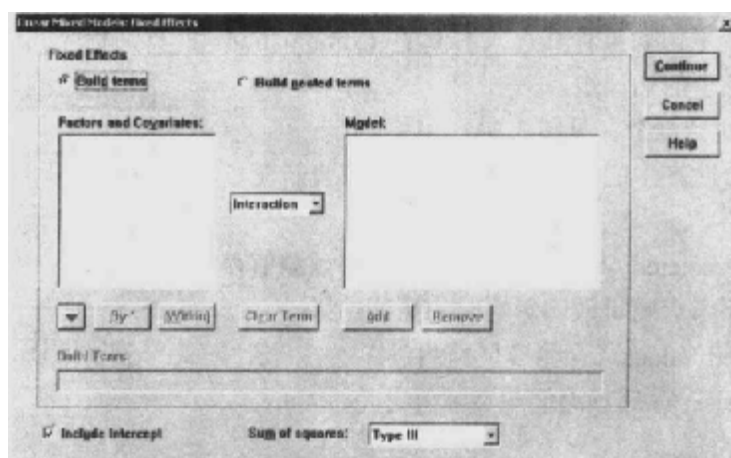


图 3.8 Fixed Effects 子对话框

【Random Effects 子对话框】（见图 3.9）

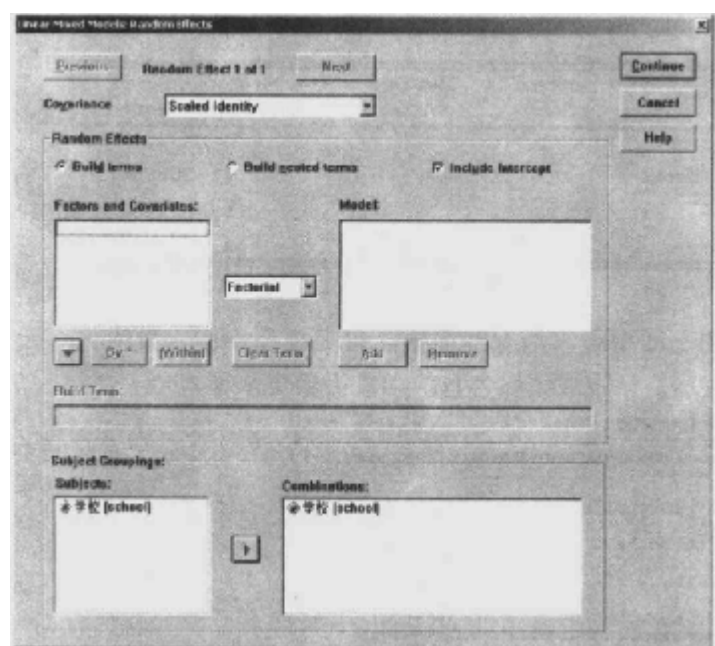


图 3.9 Random Effects 子对话框

用于设定混合线性模型中的随机效应部分，注意这里同样不提供缺省设置。

1. Previous/Next 按钮组：在该对话框中可以同时设置多个随机效应模型，此时需要使用该按钮组进行分别设置。

2. Covariance 下拉列表：对于随机效应的任何一个水平都可以单独进行协方差矩阵的估计，可以在此设置模型中每个协方差阵的结构。默认为各协方差阵完全相同。

3. Random Effects 框组：内容和 Fixed Effects 子对话框中的 Fixed Effects 框组非常相似，无需过多解释。但是在右上方多出了一个 Include Intercept 单选框，用于在随机效应中包含常数项，默认不选中。

4. Subject Groupings 框组：左侧候选变量为在预定义框中选入的个体指示变量或



重复测量指示变量，如果将其选入右侧的 Combinations 框，则相应的个体差异将会被纳入随机效应模型的估计中。

【Save】子对话框（见图 3.10）

将模型拟合时产生的中间结果或参数保存为新变量供继续分析时用。

1. Fixed Predicted Values: 将固定效应的预测值存为新变量，即不考虑随机效应影响时的预测结果。
2. Predicted Values: 存储模型预测值。
3. Residuals: 存储预测值与实际值之差。



图 3.10 Save 子对话框

### 3.2.2 结果解释

下面我们来看看引例中所拟和模型的分析结果。

## Mixed Model Analysis

Model Dimension <sup>a</sup>				
		Number of Levels	Covariance Structure	Number of Parameters
Fixed Effects	Intercept	1		1
Random Effects	Intercept	1	Identity	1
Residual				1
Total		2		3

<sup>a</sup> Dependent Variable: 18岁成绩

首先给出的是模型设置情况的简报，可见模型中固定效应只有一个常数项，随机效应也只有常数项，学生成绩以各自的学校为聚集水平。

Information Criteria <sup>a</sup>	
-2 Restricted Log Likelihood	11014.624
Akaike's Information Criterion (AIC)	11018.624
Hurvich and Tsai's Criterion (AICC)	11018.627
Bozdogan's Criterion (CAIC)	11033.241
Schwarz's Bayesian Criterion (BIC)	11031.241

The information criteria are displayed in smaller-is-better forms.

<sup>a</sup> Dependent Variable: 18岁成绩

上面给出的是模型拟合信息，包括对数似然值和其他一些信息准则。它们可以用于判断模型中引入的因素是否有统计学意义，作用要远大于后面 Walds 检验的近似结果。

## Fixed Effects

现在开始给出对模型固定部分的分析结果。

Type III Tests of Fixed Effects <sup>a</sup>				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	62.495	.060	.807

<sup>a</sup> Dependent Variable: 18岁成绩

上表为固定效应的方差分析表格，现在只有一个常数项，可见其无统计学意义。

Estimates of Fixed Effects <sup>a</sup>

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	-1.325610E-02	5.4054E-02	62.495	-.245	.807	-.1212924	9.478016E-02

<sup>a</sup>. Dependent Variable: 16岁成绩

上表为对常数项的估计值，可见大小为-0.0132。

## Covariance Parameters

下面给出的是对模型随机部分的分析结果。

Estimates of Covariance Parameters <sup>a</sup>

Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	.8477514	1.897097E-02	44.687	.000
Intercept (subject = SCHOOL)	.1715981	3.363002E-02	5.103	.000

<sup>a</sup>. Dependent Variable: 16岁成绩

上表为随机效应的估计值和检验结果，右侧还应当给出可信区间，因太长，此处截去。可见代表学校差异的常数项估计值（也就是  $u_{0j}$  的方差！）为 0.1716，具有统计学意义。这说明 16 岁学生成绩的变异在学校水平上的确存在聚集性，也就是说，不同学校间的教学水平的确有差异。上面的 Residual 则是对  $e_{ij}$  的检验结果，可见其方差大于 0，也就是学生间的确存在着个体差异。



根据上面的分析结果，实际上可以计算出同一学校内任两个学生成绩的相关系数，公式为  $r = \sigma_{u0}^2 / (\sigma_{e0}^2 + \sigma_{u0}^2)$ ，本例中为  $r = 0.1716 / (0.8476 + 0.1716) = 0.1684$ 。

请注意表格中的 Wald 检验给出的只是近似的结果，只能作为参考，当 P 值接近 0.05 时，需要使用前面的对数似然值做正确的判断。关于该问题的详细情况请大家参见本书 Logistic 回归一章的相应内容。

### 3.2.3 对引例的进一步分析

前面拟和的模型非常简单，更多的是为了演示用法，下面我们将对 JSP 数据进行深入分析，从而向大家展示混合效应模型的强大威力。

【对 11 岁成绩的影响进行分析】

考虑到学生的入学成绩可能对 16 岁时的成绩有影响，下面我们将变量 standlrt 纳入，拟和的模型为：

$$nomexam_{ij} = (\alpha_0 + \beta_1 s \tan dlrt) + (u_{0j} + u_{1j} s \tan dlrt + e_{ij})$$

请注意 standlrt 也被引入了随机部分中，即我们将同时分析它对学校内成绩的离散程度有无影响。对话框界面的操作比刚才复杂一些，具体如下：

Analyze → Mixed Models → Linear

Subjects 框: school

选入可被认为独立的高水平指示变量

Continue

Dependent Variable 框: nomexam

应变量为 nomexam

Covariate 框: standlrt

协变量(连续自变量)为 standlrt

Statistics:

☒ Parameter estimates

要求给出固定部分参数的估计值

☒ Tests for covariance parameters

要求对随机部分参数进行检验

Continue

Fixed:

选中 standlrt: Add

将 standlrt 选入固定部分

Continue

Random:

选中 standlrt: Add

将 standlrt 选入随机部分

Random Effects: ☒ Include Intercept

要求在随机部分中包含常数项

Subject Groupings: Combinations 框:

指定学校间的随机效应是大小不同的

school

Continue

OK

分析结果中的主要部分如下:

## Fixed Effects

Type III Tests of Fixed Effects <sup>a</sup>

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	60.408	.041	.841
STANDLRT	1	56.348	768.207	.000

<sup>a</sup> Dependent Variable: 16岁成绩

上表给出了固定部分中各变量的检验结果,可见 Standlrt 具有统计学意义。

Estimates of Fixed Effects <sup>a</sup>

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	-8.094E-03	4.008E-02	60.408	-.202	.841	-8.82635E-02	7.207471E-02
STANDLRT	.5570213	2.010E-02	56.348	27.717	.000	.5167675	.5972751

<sup>a</sup> Dependent Variable: 16岁成绩

上表为对固定部分中各变量参数的估计值。

## Covariance Parameters

Estimates of Covariance Parameters <sup>a</sup>

Parameter		Estimate	Std. Error	Wald Z	Sig.
Residual		5536260	1.249E-02	44.319	.000
Intercept [subject = SCHOOL]	ID diagonal	9.195E-02	1.877E-02	4.900	.000
STANDLRT [subject = SCHOOL]	ID diagonal	1.474E-02	4.640E-03	3.177	.001

<sup>a</sup> Dependent Variable: 16岁成绩

上表为随机效应的估计值和检验结果,可见 standlrt 对学校内成绩的变异的确实是有

影响的,即对于不同的11岁成绩水平,16岁成绩的离散度并不相同。显然,这种分析结论是普通线性模型所无法得到的。



连续性变量在随机部分有意义可能不太好理解,那么举个简单些的例子吧:如果性别在随机部分有统计学意义,则表明男学生和女学生16岁成绩的离散度不同。现在再回过头来看看上面的结果,脑袋里清楚些了吧。

表格中也给出了对  $u_{0j}$  和  $e_{ij}$  的检验结果,可见由于变量 `standlrt` 的引入,随机误差和学校聚集性所能解释的变异都小了许多。在理论上,分析者可以通过在随机部分中不断引入有关的解释变量使学校聚集性越来越弱,最后消失,也就是说,可以最终找到究竟是学校间的哪些特征不同导致了成绩聚集性的出现,这也是普通线性模型所无法完成的任务。

#### 【在固定部分中引入更多解释变量】

下面我们将在模型中引入更多的解释变量,来考察它们是否对16岁成绩有影响,这些变量为性别、学校类型和学校平均成绩,相应使用的模型如下:

$$\text{nomexam}_{ij} = (\alpha_0 + \beta_1 \text{standlrt} + \beta_2 \text{gender} + \sum \beta_{3k} \text{schgend}) + (u_{0j} + u_{1j} \text{standlrt} + e_{ij})$$

注意这里的学校类型为三分类变量,所以会使用两个哑变量系数来纳入模型。对话框操作时注意将 `gender` 和 `schgend` 选入 Fixed 框,而 `avslrt` 选入 Covariate 框,拟和结果如下:

### Fixed Effects

Type III Tests of Fixed Effects<sup>a</sup>

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	58.771	.464	.499
GENDER	1	3945.404	24.597	.000
SCHGEN	2	61.821	2.573	.084
STANDLRT	1	56.587	726.996	.000
AVSLRT	1	61.041	11.578	.001

<sup>a</sup> Dependent Variable: 16岁成绩

Estimates of Fixed Effects<sup>b</sup>

Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.1453729	6.375E-02	54.301	2.280	.027
[GENDER=.00]	-.1679034	3.385E-02	3945.404	-4.959	.000
[GENDER=1.00]	0 <sup>a</sup>	0	.	.	.
[SCHGEN=1.00]	-.1445218	8.175E-02	60.303	-1.768	.082
[SCHGEN=2.00]	4.289E-02	.1162241	67.871	.369	.713
[SCHGEN=3.00]	0 <sup>a</sup>	0	.	.	.
STANDLRT	5484455	2.034E-02	56.587	26.963	.000
AVSLRT	.3632101	.1067457	61.041	3.403	.001

<sup>a</sup> This parameter is set to zero because it is redundant

<sup>b</sup> Dependent Variable: 16岁成绩

Estimates of Fixed Effects <sup>a</sup>

Parameter	Estimate	Std. Error	df	t	Sig.
Intercept	.1453729	6.375E-02	54.301	2.280	.027
[GENDER=.00]	-.1679034	3.385E-02	3945.404	-4.959	.000
[GENDER=1.00]	0 <sup>a</sup>	0			
[SCHGEND=1.00]	-.1445218	8.175E-02	60.303	-1.768	.082
[SCHGEND=2.00]	4.289E-02	.1162241	67.871	.369	.713
[SCHGEND=3.00]	0 <sup>a</sup>	0			
STANDLRT	.5484455	2.034E-02	56.587	26.963	.000
AVSLRT	.3632101	.1067457	61.041	3.403	.001

<sup>a</sup> This parameter is set to zero because it is redundant

<sup>b</sup> Dependent Variable: 16岁成绩

由上面的分析结果可见，性别和学校平均成绩的确是有影响的，其中性别也采用的是哑变量的形式，具体为女生作为参照水平，[GENDER=1.00]一行为男生与之相比的结果，可见男学生的成绩较低。Schgend的P值大于0.05，此时最好使用似然值得到更为精确的P值，最终结论为Schgend的确无统计学意义，此处不再详述。

## ▼ Covariance Parameters

Estimates of Covariance Parameters <sup>a</sup>

Parameter		Estimate	Std. Error	Wald Z	Sig.
Residual		.5501983	1.242E-02	44.310	.000
Intercept (subject = SCHOOL)	ID diagonal	6.911E-02	1.510E-02	4.575	.000
STANDLRT (subject = SCHOOL)	ID diagonal	1.523E-02	4.742E-03	3.211	.001

<sup>a</sup> Dependent Variable: 16岁成绩

最后为随机效应的检验结果，由于没有引入新的变量，各参数方差的估计值变动不大。

下面还可以对JSP数据做进一步的分析，如研究性别与学校类型间有无交互作用，学校平均成绩对离散程度有无影响等，因篇幅所限，这里不再详述，请大家自行操作。

## 3.3 混合线性模型分析实例

上面的引例是一个样本量较大的数据，实际上，对于一些常见的数据结构，混合线性模型仍然有其用武之地，下面就给出一些分析实例，这些例子我们在前面两章中曾经使用一般线性模型加以分析，现在给出使用混合线性模型的分析结果，以进行比较，帮助大家开阔思路。



### 3.3.1 家庭聚集性数据

这里的数据就是上一章中的 Mixed.sav，我们已经使用 Variance Components 过程对聚集性进行了初步分析，现在将对它进行正式的分析。根据前面的结果，得到的信息如下：

1. 个体为低水平单位，家庭为高水平单位。
2. 身高为应变量，性别为自变量，研究目的为考察性别对身高有无影响。
3. 根据 Variance Components 过程的结果，家庭所导致的变异和随机变异差不多，因此聚集性可能并不明显。

据此，我们可以使用混合线性模型对它加以分析，在随机部分中指定常数项以分析家庭有无聚集性，主要的分析结果如下：

Type III Tests of Fixed Effects <sup>a</sup>

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	2.090	5339.484	.000
GENDER	1	12.253	18.293	.001

<sup>a</sup> Dependent Variable: HEIGHT

Estimates of Covariance Parameters <sup>a</sup>

Parameter	Estimate	Std. Error	Wald Z	Sig.
Residual	3.0961227	1.2523242	2.472	.013
Intercept [subject = FAMILY]	2.5346016	3.1257039	.811	.417

<sup>a</sup> Dependent Variable: HEIGHT.

首先给出的是固定部分因素的检验结果，可见性别对身高有影响。随后为随机部分的检验结果，可见常数项无统计学意义，即身高在家庭内部的聚集性未能被证实。在本例中，这是由于数据量太少，检验效能不足所致。

大家可以使用 GLM→Univariate 过程分析该数据，并将对性别的检验结果和参数估计与这里相对照，就会发现基本一致。这说明当数据间聚集性不太明显时，混合线性模型的优势不能被体现出来，它对固定部分的分析结果和一般线性模型是一致的。

### 3.3.2 重复测量数据

上一章中我们使用重复测量的方差分析模型对 Anxiety 2.sav 进行了分析，得出结论认为四次重复测量间实际上无相关，并且 Anxiety 和 tension 对评分无影响。现在我们使用混合线性模型来加以分析。但是，由于数据格式要求不同，这里使用的应当是长型的 Anxiety.sav。操作如下：

Analyze→Mixed Models→Linear

Subjects 框: subject

选入可被认为独立的高水平指示变量

Repeated 框: trial

Trial 为重复测量次数指示变量

Repeated Covariance type: Scaled Identity

限定各次重复测量间的变异大小相同

Continue

Dependent Variable 框: score

应变量为 score

Factors 框: anxiety、tension、trial

希望分析的因素为 anxiety、tension、trial

**Fixed:**选中 anxiety: **Add**; 选中 tension: **Add**  
选中 trial: **Add**; **Continue**在固定部分中加入因素 anxiety、tension  
在固定部分中加入因素 trial**Random:**Random Effects: ☒ Include Intercept

要求在随机部分中包含常数项

Subject Groupings: Combinations 框: 指定个体间的随机效应是大小不同的  
subject**Continue****OK**

在上面的操作中, 我们指定了 Trial 为重复测量次数指示变量, 但是限制其各次重复测量间的变异大小相同, 同时在随机部分中加入对个体变异的分析, 主要分析结果如下:

Type III Tests of Fixed Effects <sup>a</sup>

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	8.802	265.710	.000
ANXIETY	1	8.802	.558	.474
TENSION	1	8.802	.461	.514
TRIAL	3	33	127.561	.000

<sup>a</sup>. Dependent Variable: Score

上表为固定部分因素的检验结果, 大家可以看到该结果和上一章中使用重复测量的方差分析模型的分析结果完全相同。

Estimates of Covariance Parameters <sup>a</sup>

Parameter		Estimate	Std. Error
Repeated Measures	ID diagonal	2.5909091	.6378377
Intercept [subject = SUBJECT]	ID diagonal	3.8684764	2.1349223

<sup>a</sup>. Dependent Variable: Score

上表为随机部分因素的估计, 从估计值和标准差可知个体差异的大小实际上和随机误差大小差不多, 且无统计学意义。

本例我们限定了各次重复测量间的变异大小相同, 主要是为了得到和前面相同的分析结果, 但这样做实际上使得重复测量变量的指定失去了意义, 大家可以将 trial 不指定为重复测量变量, 会发现分析结果完全相同。在真正的混合线性模型分析中, 我们可以指定各次测量间独立且方差不同 (记得 Huynh-Feldt 条件吗), 从而得到四次测量单独的方差估计, 限于篇幅, 请大家自行操作, 此处不再详述。

### 3.3.3 嵌套设计数据

这里要分析的数据是曾经使用过的 nested.sav, 在 Univariate 过程中我们无法使用对话框直接分析, 而必须要对程序加以修改。现在混合线性模型的对话框中可以直接设定变量嵌套, 对不喜欢编程的朋友来说无疑是一大福音。这里就来演示其用法, 由于该数据实质上不涉及随机效应的分解, 此处应当只对固定部分加以设置, 操作如下:

**Analyze**→**Mixed Models**→**Linear**

Subjects 框: mouse

选入可被认为独立的高水平指示变量

Continue

Dependent Variable 框: weight

应变量为 weight

Factors 框: siliao、week、mouse

希望分析的因素为 siliao、week、mouse

Fixed:

选中 siliao: Add; 选中 week: Add

在固定部分中加入因素 siliao、week

☒ Build nested terms

要求建立嵌套因素表达式

选中 mouse: : (Within):

指定 mouse 为嵌套因素

选中 siliao: : Add

生成表达式 mouse(siliao), 并加入因素列表

Continue

OK

在上面的操作中主要是嵌套因素的设置比较复杂, 请注意不要操作错误。分析结果中的主要部分如下:

Type III Tests of Fixed Effects <sup>a</sup>

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	33	3475.743	.000
SILIAO	2	33	15.252	.000
WEEK	3	33	12.277	.000
MOUSE(SILIAO)	9	33	.669	.731

<sup>a</sup> Dependent Variable: WEIGHT.

上表为固定部分的检验结果, 可见三种饲料间是有差异的, 该结果可以和前面使用 Univariate 过程的分析结果相对照, 会发现基本完全相同。

Estimates of Covariance Parameters <sup>a</sup>

Parameter	Estimate	Std. Error
Residual	23.2688952	5.7284060

<sup>a</sup> Dependent Variable: WEIGHT.

上表为随机部分的参数估计, 由于没有做进一步的分解, 所以结果没有什么实际意义。