

kaobook 类

使用此文档作为模板

示例及说明文档 of the kaobook 类

根据自己需要定制本页

Federico Marotta *

November 25, 2019

奥色姆曼出版社

* A L^AT_EX lover

kaobook 类

Disclaimer

你可以编辑这个页面来满足你的需要。例如，这里有一个无版权的声明、一个版权页标记和其他一些信息。这一页是基于肯·阿罗约·奥赫里的论文的相应页面，改动很小。

No copyright

© This book is released into the public domain using the CC0 code. To the extent possible under law, I waive all copyright and related or neighbouring rights to this work.

To view a copy of the CC0 code, visit:

<http://creativecommons.org/publicdomain/zero/1.0/>

Colophon

This document was typeset with the help of KOMA-Script and L^AT_EX using the kaobook class.

The source code of this book is available at:

<https://github.com/fmarotta/kaobook>

(You are welcome to contribute!)

Publisher

First printed in Jan 2019 by 奥色姆曼出版社

世界的和谐体现在形式和数量上，自然哲学的心和灵魂以及一切诗歌都体现在数学美的概念上。

– D'Arcy Wentworth Thompson

前言

我的观点是，每个 \LaTeX 极客，至少在他的一生中有一次，都觉得有必要创建自己的类：这就是发生在我身上的事情，这就是结果，但是，这应该被看作是一个仍在进行中的工作。实际上，这个类并不完全是原创的，但它是我在许多指南、教程、博客和 `text.stackexchange.com` 文章中发现的所有最佳思想的混合体。特别是，主要的想法来自两个来源：

- ▶ [Ken Arroyo Ohori's Doctoral Thesis](#), which served, with the author's permission, as a backbone for the implementation of this class;
- ▶ The [Tufte-Latex Class](#), which was a model for the style.

本书第一章是导论，涵盖了本课程最基本的特点。接下来，有一堆章节专门讨论所有的命令和环境，你可以用来写一本书；特别地，它将解释如何添加注释、图形和表以及引用。第二部分讨论页面布局和设计，以及其他特性，如彩色框和定理环境。

我开始写这门课是作为一个实验，因此它应该被视为。由于它一直是我个人使用的缩进，它可能不是完美的，但我发现它很满意的用途，我想让它。我分享这篇文章，希望有人能从这里找到写作的灵感。

Federico Marotta

详细目录

前言	vii
详细目录	viii
因素实验设计与方差分析计算原理	1
1 两因素被试间实验设计	3
1.1 两因素交互作用的进一步探讨	3
交互作用含义	3
交互作用图解	3
简单效应检验	5
1.2 两因素随机机区组实验设计	7
2 两因素重复测量实验设计	9
基本特点	9
实验设计与计算举例	9
一些解释	10
基本特点	12
实验设计与计算举例	13
一些解释	13

插图目录

1.1 两因素交互作用可能的 8 种情况，中间蓝色的加粗线为 A 因素的简单主效应，可以看到只考虑一种因素会掩盖很多信息。(1) 和 (5) 中三条线本是重合的，但是为了可以看得清楚故将其稍微拉开了一点，(8) 中绿色的虚线和实线表示 b_1 这两种状况都属于这种情况	4
1.2 AB 交互作用两个方向的图	5

表格目录

1.1 两因素完全随机设计 ABS 表	5
1.2 两因素完全随机设计 AB 平均数表	5
1.3 两因素完全随机方差分析与两个方向简单效应检验	7
2.1 Add caption	9
2.2 两因素混合实验设计方差分析表	10
2.3 忽略 B 因素 (被试内变量) 的 AS 表	10
2.4 在 a_1 水平的 BS 表	11
2.5 在 a_2 水平的 BS 表	12
2.6 Add caption	12
2.7 两因素被试内实验设计方差分析表	13

定理目录

1.1.1 定义 (交互作用)	3
2.0.1 定义 (两因素混合实验设计模型)	9
2.0.2 定义 (两因素混合实验设计平方和分解)	9
2.0.3 定义 (两因素被试内实验设计模型)	12
2.0.4 定义 (两因素被试内实验设计平方和分解)	13

因素实验设计与方差分析计算原理

1 单因素实验设计

1.1 两因素交互作用的进一步探讨	3
交互作用含义	3
交互作用图解	3
简单效应检验	5
1.2 两因素随机机区组实验设计	7

良好实验在统计上的体现

我们知道这样一个事实, 随机区组设计可以带走区组变量引起的无关变异, 重复测量设计可以带走被试个体带来的无关变异, 这样实验的结果就更加精确了. 本书的目的是从数据的角度上展示, 用了好的实验设计后, 更精确的实验的结果的体现. 我们将实验设计与方差分析结合, 从数据上看看为什么好的实验设计可以使 F 增大. 我们知道:

$$F = \frac{MSA}{MSE}$$

想让 F 变大有两个手段, 让处理效应 (MSA) 变大, 或者让误差效应 (MSE) 变小, 这件事情在前面已经介绍过. 心理学家感兴趣的那些现象往往都不太明显, 人为加大处理效应也没有意义. 那么, 减小误差才是关键.

这就好比用天文望远镜观测外太空, 我想发现星星并用我自己的名子命名. 我透过天文望远镜看到了一些没有规律发亮的光点, 也就是噪音. 它们是附近一些已知的行星发出的, 也可能是从隔着很远的地方的某个星体传过来的, 总之它们像背景一样衬托在我感兴趣的那个星星所发出的光下面. 我感兴趣的那颗星星也在发光, 不过我现在还不知道这个星星是否真的存在, 但我知道, 如果这个星星真的存在, 它一定会努力发光, 并且会和周围那样噪音有显著区别. 事情不是那么容易, 星星发的光和噪音总是同时出现, 只有星星发出的光足够大, 大到和周围的噪音是那么的与众不同, 我才敢认定那是星星. 我们清楚地知道, 想要更加清楚地观察到星星无非有两种手段, 其一是让星星在原有基础上多发光电; 其二是让周围的噪音降低点. 不过好像第一种手段听起来有点滑稽.

方差分析模型中的 MSA 和 MSE 是相互独立的, 它们俩可以独立地变化. 随机误差指的是由随机因素导致数据在理想值周围随机晃动的偏差, 处理效应是人为控制自变量在不同水平上的系统变异, 由完全不相干的两个东西导致当然它们是独立的了. 下面是一道很重要的思考题:

思考

假定我们有一组分数, 分别计算出 MSA 和 MSE .

(1) 我们可以改变 AS^1 分数, 保持处理平均数恒定. 即保持 MSA 不变而改变 MSE .

(2) 如果我们改变组平均数, 处理组内分数关系不就, MSA 会变化, 而 MSE 不就. 表明两个均方是独立变化的.

我们看一组数据

$a_1 : 1, 2, 3, 4, 5$

$a_2 : 5, 6, 7, 8, 9$

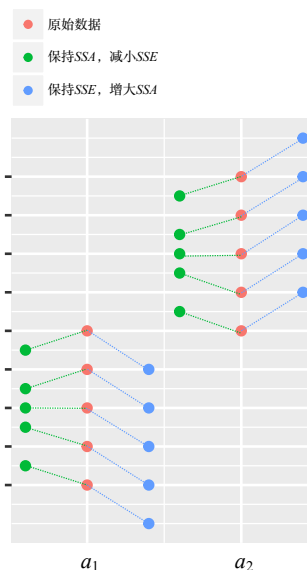


Figure 1.1: 红色是原始数据, 蓝色代表组内误差不变, 加大处理效应; 绿色代表组间变异不变, 改变组内误差

1: AS 中 A 代表因素 A 的处理效应, S 代表被试个体误差, 具体是什么后面再说

(1) 调整原始数据, 保持 SSA 不变, 而 SSE 改变 (减小 SSE) (通常加大处理效应实现)

(2) 调整原始数据, 改变 SSA 不变 (增大 SSA), 而 SSE 不变

我直接给出我的思路, 图 ?? 中红色是原始数据, 它有两个水平, a_1 的各个值围绕其均值 3 波动, a_2 的各个值围绕其均值 7 波动.

绿色是我写的符合 (1) 的一组数据, 可以看到红色数据 a_1 处理的每个数据都围绕着 3 波动, 现在绿色的值在 a_1 内还是围绕 3 波动, 但是数据间的距离更短了, 它们显得更加紧密. 虽然从组内均值上来看 (3 和 7) 两组数据的平均值之差不变, 但是由于 a_1 的数据更集中于 3, a_2 的数据更集中于 7, 故实际上这样两组间的差异其实增大了.

蓝色是我写的符合 (2) 的一组数据, 可以看到数据间的紧密程度不变, a_1 整体向下移动, a_2 整体向上移动, 这样组间的差距被拉大.

原始数据方差分析结果是.

Table 1.1: 原始数据方差分析

变异来源	SS	df	MS	F	p
组间变异	40.000	1.000	40.000	16.000	0.004
组内变异	20.000	8.000	2.500		

Table 1.2: 保持 SSA 不变, 减小 SSE

a_1	1.5	2.5	3	3.5	4.5
a_2	5.5	6.5	7	7.5	8.5

绿色数据见表 ?? on page ??.

其方差分析结果是, 可以看到相比表 ?? 中的数据, 处理效应没变, 误差效应减小了, F 值变大, p 值减小.

变异来源	SS	df	MS	F	p
组间变异	40.000	1.000	40.000	32.000	< .001
组内变异	10.000	8.000	1.250		

Table 1.3: 绿色数据方差分析

Table 1.4: 保持 SSE 不变, 增大 SSA

a_1	1	1	2	3	4
a_2	6	7	8	8	10

蓝色数据见表 ??.

其方差分析结果是, 可以看到相比表 ?? 中的数据, 处理效应变大, 组内误差减没变, F 值变大, p 值减小.

变异来源	SS	df	MS	F	p
组间变异	90.000	1.000	90.000	36.000	< .001
组内变异	20.000	8.000	2.500		

Table 1.5: 蓝色数据方差分析

1.1 单因素完全随机实验设计

基本特点

完全随机设计 (completely randomized design) 的假定是, 由于被试是从总体中随机选取, 处理也是随机分配给被试, 那么被试间的差异也是随机、在统计上无差异的.

单因素完全随机设计中分配被试的图解例子如表 ??, 个因素 A 有 4 个水平, 每个处理下有 4 个被试. 完全随机设计是一种被试间设计,

Table 1.6: 单因素完全随机实验设计中被试的分配

a_1	a_2	a_3	a_4
S_1	S_2	S_3	S_4
S_5	S_6	S_7	S_8
S_9	S_{10}	S_{11}	S_{12}
S_{13}	S_{14}	S_{15}	S_{16}

这意味着每一个被试只接受一种处理或处理的组合², 故对于本例而言需要 16 个被试.

对接下来要描述的模型的字母进行一下简要的说明:

(1) 假设因素 A 有 p 个水平, 用 α_j 表示其第 j 个水平 ($1 \leq j \leq p$), 也就是说 A 的水平分别是 $\alpha_1, \dots, \alpha_j, \dots, \alpha_p$,

(2) p 个处理下的被试数都相等, 均为 n 个, 用 i 表示每个处理下第 i 个被试 ($1 \leq i \leq n$)

(3) 假设被试在没有施加处理的时候, 无限次实验结果的平均数是 μ . 该值是个理想值, 其实意思是被试测量值在理论上的基线, 不过这么说只是表达了概念. 虽然说明了更加明确的定义不过可以充分的感受到, 这个理论基线值是无法获得的, 毕竟无限次是个极限, 不是一个具体的数, 所以 μ 是个理论值.

(4) 假设实验中的误差是 $\varepsilon_{i(j)}$ ³, 其平均数为 0, 方差为 σ_e^2 , 我们进一步假设它服从正态分布

(4) 观察值 Y_{ij} 指第 j 个处理水平下, 第 i 个被试的观测值

后现的章节因素多了会加入一些新字母和下标, 和上述符号已经没有本质区别, 故以后我不打算再介绍新符号了.

单因素完全随机实验设计模型

定义 1.1.1 (单因素完全随机实验设计模型)

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{i(j)} \\ (j = 1, 2, \dots, p; i = 1, 2, \dots, n)$$

其中:

- Y_{ij} - 被试 i 在处理水平 j 上的分数
- μ - 总体平均数
- α_j - 水平 j 的处理效应
- $\varepsilon_{i(j)}$ - 误差效应

单因素随机区组实验设计检验的假说

单因素完全随机设计的原假说是: 多个处理水平上的总体平均数相等, 即

$$H_0: \mu_1 = \mu_2 = \dots = \mu_p$$

或处理效应等于为 0, 即

$$\alpha_j = 0$$

实验设计与计算举例

研究的问题

我想研究文章的生字密度对学生阅读理解的影响. 我的假设是: 阅读理解随着文章中生字密度的增加而下降. 因此, 该实验有一个自变量——生字密度, 通过研究前人文献我感兴趣的是四种生字密度 (1)5:1 (a_1)、(2)10:1 (a_2)、(3)15:1 (a_3)、(4)20:1 (a_4) 因变量是被试的阅读理解分数. 实施实验时, 我将 32 名被试随机分成四组, 每组被试阅读一种生字密度的文章, 并回答阅读理解测验中有关文章内容的问题. 这是一个典型的单因素完全随机设计. 虽然我不再检验实验中其他因素的影响, 但实际上存在着诸多对因变量影响的无关变量, 比如文章的长度、文章的主题熟悉性、文章类型等, 还有被试的年龄、受教育程度、阅读能力等. 这时, 控制无关变量可做的工作之一是在选取四篇文章时, 使它们在除生字密度以外的其他方面尽量匹配.

新的计算模型的引入以及数据的计算

我将引入一种新的计算模式, 这种模式主要是为了手算. 为了理解愿意, 我们必须手算一次每种实验设计的简单数据, 但是用统计教材上的算法根本无法实现一个较为精确的估算, 其原因是, 和方的计算公式选用:

$$SS = \sum_{i=1}^n (X_i - \bar{X})^2$$

式子中 \bar{X} 很大可能是小数, 且可能是循环小数, 所以每次相减都会四舍五入, 然后平方后进一步放大误差, 这样算到最后误差不断累积. 所以应当最后一步再做除法, 把上面的式子改写一下:

$$\begin{aligned} SS &= \sum_{i=1}^n (X_i - \bar{X})^2 \\ &= \sum_{i=1}^n \left(X_i^2 + \underbrace{\bar{X}^2}_{\text{常数}} - \underbrace{2\bar{X} X_i}_{\text{常数}} \right) \\ &= \sum_{i=1}^n X_i^2 + n\bar{X}^2 - 2\bar{X} \underbrace{\sum_{i=1}^n X_i}_{2n\bar{X}^2} \\ &= \sum_{i=1}^n X_i^2 - n \left(\frac{\sum_{i=1}^n X_i}{n} \right)^2 \\ &= \underbrace{\sum_{i=1}^n X_i^2}_{\text{离散} + \text{集中} + \text{可能个体差异}} - \underbrace{\frac{\left(\sum_{i=1}^n X_i \right)^2}{n}}_{\text{集中}} \end{aligned}$$

这样除法只进行了一次(上式“集中”那里算了一次除法,产生误差),其余的都是乘法和加法.那么在方差分析中,大量求 SS , 用这样的方法就可以利于手算,以便更好了解方差分析原理.

1. 计算表

表 ?? 中的数据被我标了 3 个颜色,不同颜色的区域反映了不同的变异,一个一个来看.灰色数据可以看到集中趋势、离散趋势,包含个体差异;绿色数据看不到个体差异,但是能看到组间差异;蓝色数据只能看到集中趋势.

某个区域的颜色为 L , 定义这个区域反应的变异 [因素] 的算法:

$$[A?B?C?\dots S?] = \frac{\sum (\text{某颜色区域的数据}^2)}{\text{该颜色一个单元格内被试数}}$$

公式中的 ? 表示的是前面的因素可有可无.如果你感觉有些抽象,不妨直接看一下下面是在单因素完全随机实验设计中是怎么计算的.以后所有的 SS 都是这样的算法.列表算 SS 的好处在多因素实验设计中会充分展现,现在可能还是显得麻烦.

2. 各种基本量的计算

$$\sum_{i=1}^n \sum_{j=1}^p Y_{ij} = 3 + 6 + 4 + \dots = 202.000$$

$$\frac{\left(\sum_{i=1}^n \sum_{j=1}^p Y_{ij} \right)^2}{np} = [Y] = \frac{(202)^2}{(8)^2 (4)^2} = 1275.125$$

$$\sum_{i=1}^n \sum_{j=1}^p Y_{ij}^2 = [AS] = (3)^2 + (6)^2 + \dots = 1544.000$$

$$\sum_{i=1}^n \frac{\left(\sum_{j=1}^p Y_{ij} \right)^2}{n} = [A] = \frac{(35)^2}{8} + \frac{(31)^2}{8} + \dots = 1465.250$$

其中, $[Y]$ 仅仅用了右下角的数据,这里的数据只有集中趋势,不体现个体差异,也不体现离散趋势,用字母 Y 表示集中趋势; $[A]$ 用到了下方绿色的数据,这部分数据由于合并了组内的数据,所以不体现个体差异,不过还能体现离散趋势和集中趋势.用字母 A 表示因素 A 的集中和离散趋势,要得到纯粹的 A 的离散趋势还要减去 $[Y]$ 这个集中趋势; $[AS]$ 用到了中间灰色区域的数据,体现了集中趋势,也体现了离散趋势,而且每个个体的数据都用到了,故也能体现个体差异, $[AS]$ 中各个符号前面都出现过,其中 A 代表了因素 A 的集中和离散趋势, S 表示是个体差异.

由上面的分析可以看出,只有原始数据可以反映个体差异 S ,只要合并了数据就损失了个体差异;离散趋势只要有多组数据都可以反映;所有的数据都一定会反映集中趋势,故在所有式子都不再加入 Y ,否则过于麻烦.

3. 平方和分解与计算

Table 1.7: 单因素完全随机实验的 AS 表

	a_1	a_2	a_3	a_4	
	3	4	8	9	
	6	6	9	8	
	4	4	8	8	
	3	2	7	7	
	5	4	5	12	
	7	5	6	13	
	5	3	7	12	
	2	3	6	11	
总	35	31	56	80	202

定义 1.1.2 (单因素完全随机实验设计平方和分解)

$$\begin{aligned} SS_{\text{总}} &= SS_{\text{组间}} + SS_{\text{组内}} \\ &= SSA + SS_{\text{单元内}} \end{aligned}$$

$$SS_{\text{总}} = [AS] - [Y] = 268.875$$

$$SS_{\text{组间}} = SSA = [A] - [Y] = 190.125$$

$$SS_{\text{组内}} = SSE = SS_{\text{总}} - SS_{\text{组间}} = 78.750$$

4. 方差分析表及结果的解释

Table 1.8: 单因素完全随机设计的方差分析表

变异源	SS	df	MS	F	p
组间 (生字密度)	190.125	$p - 1 = 3$	63.375	22.533	$< .001$
组内 (个体误差)	78.750	$p(n - 1) = 28$	2.813		
总计	268.875	$np - 1 = 31$			

方差分析结果表明,生字密度的效应是显著的 ($F(3, 28) = 22.53(p < 0.01)$), 生字不同的文章对学生的学生阅读理解有显著的影响。从阅读理解的四种生字密度文章的平均数可以初步看出

5. 平方和与自由度分解图

$$SS_{\text{总}} \begin{cases} SS_{\text{组间}} \\ SS_{\text{组内}} \end{cases} \quad (1.1)$$

$$SS_{\text{总}} \quad df = np - 1 = 31 \quad SS_{\text{组间}} \quad df = p - 1 = 3 \quad SS_{\text{组内}} \quad df = p(n - 1) = 28$$

一些解释

1. 各种平方和的意义

$SS_{\text{总变异}}$ ——总平方和或总变异, 带有实验数据中所有的变异, 包括实验处理效应、无关变异和误差变异.

$SS_{\text{组间}}$ ——组间平方和, 或处理平方和, 指所有由于实验处理引起的变异, 在单因素设计中指 A 因素的处理效应.

$SS_{\text{组内}}$ ——组内平方和, 或误差平方和, 指所有不能用实验处理解释的变异, 它可能包括被试个体差异、其他无关变异和实验误差. 在单因素完全随机实验中, 不再对组内平方和做进一步分离, 因此在总变异中减去组间平方和就是组内平方和.

完全随机实验中 F 值的计算是:

$$F = \frac{MS_{\text{组间}}}{MS_{\text{组内}}}$$

F 值的实质是两个卡方商的分布, 而 $(n - 1)S^2/\sigma^2 \sim \chi_{n-1}^2$, 故 F 值可以比较两个方差是否存在显著差异, 在方差分析中就是检验实验处理带来的效应是否不同于实验误差, 如果差异达互一定统计显著水平, 表明处理的效应是存在的.

2. 方差显著的意义

方差分析显著表明，从整体来看组间的平均数是有差异的，即 $\alpha_1 \neq \alpha_2 \neq \alpha_3 \neq \alpha_4$ ，但是具体那些存在差异不知道。想检验到底是哪些平均数间存在差异要进行比较和对比，这不是本章的内容，因此不介绍了。

3. 同质性检查

F 检验要求被试在观测值上的变异是相同的，这靠随机化完成，检验是否真的做到了同质性就要进行同质性检验。同质性的要求是处理前被试同质，在心理学实验中通常不进行前测，因为我们假定实验处理不影响随机因素，故处理后处理内被试的差异也反映了处理前被试间的个体差异，比如在图 ?? 中红色点在不同处理水平的分散水平都一样， a_1 下它们离各个相近的点相差 1， a_2 下它们还是这样。

严格的因果关系要求让因变量的差异仅由自变量不同水平造成，研究者操纵自变量在不同水平上变化，除了自变量不同之外，其余的变量都是无关变量，需要全部保持恒定。不过这点在以人为研究对象的实验中无法达到，不同人在进入实验室前就已经经历了他/她自己独一无二的人生，因此实验即使控制地再严密，接受相同的被试之间在因变量的观测值上仍然存在差异，这种误差叫单元内误差。这些误差在今后我们将看到，从性质上只有两种，这里我们只能介绍其中的单元内误差。

同质性检验计算的方法是比较各组中，最大的误差和最小的误差间是否存在差异：

$$SS_{1\text{组}} = (3^2 + 6^2 + \dots) - \frac{(35)^2}{8} = 19.875$$

$$SS_{2\text{组}} = (4^2 + 6^2 + \dots) - \frac{(31)^2}{8} = 10.875$$

$$SS_{3\text{组}} = (8^2 + 9^2 + \dots) - \frac{(56)^2}{8} = 12.000$$

$$SS_{4\text{组}} = (9^2 + 8^2 + \dots) - \frac{(80)^2}{8} = 36.000$$

$$F = \frac{SS_{\text{最大}}}{SS_{\text{最小}}} = \frac{36.000}{10.875} = 3.31$$

$F(8, 8) = 3.31 (p > .05)$ 表明接受不同实验处理的四组被试是统计上无差异的。

4. 误差平方和的计算

完全随机实验中，误差平方和的计算有两种方法。一种是相减法，如上所述，在单因素完全随机实验中，用总平方和减去组间平方和，即可得误差变异。从这种意义上说，误差变异是指不能被实验处理所解释的变异。

另一种是直接计算法。在单因素完全随机实验中，误差平方和还可以通过先计算各处理组内的平方和，再将各组平方和相加而获得。利用同质性检查中所得到的，可直接计算误差平方和，组内平方和如下：

$$SS_{\text{组内}} = SS_{1\text{组}} + SS_{2\text{组}} + SS_{3\text{组}} + SS_{4\text{组}} = 78.750$$

Table 1.9: 计算单元内误差

a_1	a_2	a_3	a_4
3	4	8	9
6	6	9	8
4	4	8	8
3	2	7	7
5	4	5	12
7	5	6	13
5	3	7	12
2	3	6	11
35	31	56	80

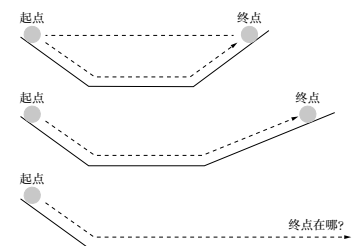


Figure 1.2: Galileo 斜面实验

这个误差就是上面提到的单元内误差. 这种误差在是心理学中的一种非常特殊的误差, 在物理实验中就不存在这样误差, 比如在 Galileo 的斜面实验 (slope experiment), 可以看到图 ?? 中每个圆球都会上升到同一个高度, 不同的球在各种属性上都是完全相同的, 所以重复实验对物理属性匹配了的圆球就会有相同效果. 但是对于人而言, 每个人都不一样, 都带着性别、性格、年龄、教育程度等进入实验室, 而这些因素会影响行为, 我们无法做到匹配完全相同的两个人. 这个变异在完全随机设计中给我们提供了一个随机误差的估计, 所有的处理效应都是和它相比, 尽管这部分变异不完全是随机的, 不过对于完全随机设计而言已经没有更好的办法了.

从这个角度上来说, 任何实验中每一组内都不可能只用一个被试, 如果只有一个被试就没有变异了, 起码两个人才有变异.

评价

完全随机实验设计的**优点**是, 实验设计和实施简单, 接受每个处理水平的被试数量可以不相等, 不需要匹配被试, 每个被试公接受一个处理水平. 完全随机实验的统计分析和对结果的解释简单, 并且与它的误差平方和相对应的自由度最大. 因此, 如果在不同的实验设计中得到的误差平方和相等, 那么完全随机实验设计比其他实验设计更敏感.

完全随机实验设计的**缺点**是, 它的组内变异并非全部由随机误差组成, 其中还包括了被试的个体差异. 虽然完全随机设计假设随机分配的各组被试在统计上是无差异的, 但实际上被试的个体差异带来的无关变异是存在的, 并且混杂在组内变异中, 导致 F 比率的分母项加大, 从而使实验较为不敏感. 另外, 当实验中含多个处理水平时, 需要的被试量也会较大.

1.2 单因素随机区组实验设计

基本特点

行为科学研究中, 被试的个体差异是误差变异的重要来源. 它常常会混淆实验处理效应, 因此是无关变异.

方差分析用随机误差代表一些随机发生的因素带来的影响, 如果某个效应带来的变异显著大于随因素带来的变异, 我们认为这个效应不同于随机因素, 是个在统计上有显著意义的效应. 然而, 在完全随机实验设计中, 这个误差的衡量方式是所有被试间的个体差异, 但是被试间的个体误差不能完全代表随机误差, 因变量对其中某些因素是敏感的. 在这些个体差异带来的无关变异中, 随机区组设计找出影响因变量最大的, 分离这个无关变量带来的变异, 使误差效应变小, 使这部分无关变异既不出现在处理效应中混淆处理作用的大小, 也不出现在误差中降低检验力. 这样在 F 检验中:

$$F = \frac{MSA}{MS_{\text{组内}}}$$

将 MSE 中由无关变量 B 带来的无关变异减去, 得到

$$MSE = MS - MSB$$

使误差效应更小, F 更容易显著.

单因素随机区组设计适用于这样的情境: 研究中有一个自变量, 自变量有两个或多个水平 ($P \geq 2$), 研究中还有一个无关变量, 也有两个或多个水平 ($n \geq 2$), 并且自变量的水平与无关变量的水平之间没有交互作用.

当无关变量是**被试变量**时, 一般首先将被试在这个无关变量上进行匹配, 然后将他们随机分配给不同的实验处理. 这样, 区组内的被试在此无关变量上更加同质, 他们接受不同的处理水平时, 可看作不受无关变量的影响, 主要受处理的影响而区组之间的变异反映了无关变量的影响, 我们可以利用方差分析技术区分出这一部分变异, 经减少误差变异, 获得对应效应的更精确估价.

另外环境因素也是潜在可考虑的区组变量, 例职, 每天的时间、每年的季节、地点、仪器等方面的因素也可以进行区组, 经减少误差变异, **时间**是一个特别有效的区组变量, 因为它常常还会带来一些附加自变量, 如身体的生理周期、疲劳等.

单因素随机区组实验设计中被试与处理的分配如表 ??, 图中可以看出实验中有一个自变量, 该自变量有 4 个水平. 实验中还有一个无关变量, 将 16 个被试在无关变量上进行匹配, 分为 4 个区组, 每个区组内 4 个同质被试, 随机分配每个被试接受一个处理水平.

单因素随机区组的下标字母与单因素完全随机差不多, 不同的是, 完全随机设计中, 每个处理内被试有多个, 故 i 表示的是处理内第 i 个被试, 每个处理都有 n 个被试; 而在随机区组设计中, 不存在这种处理内的被试, 我们看到表 ?? 中 a_1 下的 4 个被试又属于不同的区组. 因此, 现在的处理内的被试换成了区组, 故一共有 n 个区组, i 表示第 i 个区组.

Table 1.10: 单因素随机区组实验设计中被试的分配

	a_1	a_2	a_3	a_4
区组 1	S_1	S_2	S_3	S_4
区组 2	S_5	S_6	S_7	S_8
区组 3	S_9	S_{10}	S_{11}	S_{12}
区组 4	S_{13}	S_{14}	S_{15}	S_{16}

单因素随机区组实验设计模型

定义 1.2.1 (单因素随机区组设计模型)

$$Y_{ij} = \mu + \alpha_j + \pi_i + \varepsilon_{i(j)} \\ (j = 1, 2, \dots, p; i = 1, 2, \dots, n)$$

其中

- Y_{ij} - 被试在区组 i 和处理水平 j 上的分数
- μ - 总体平均数
- α_j - 水平 j 的处理效应
- π_i^2 - 区组效应, 且 $\pi_i^2 \sim N(0, \sigma_\pi^2)$
- $\varepsilon_{i(j)}$ - 误差效应

单因素随机区组实验设计检验的假说

1. 处理水平的总体平均数相等

$$\mu_{.1} = \mu_{.2} = \cdots = \mu_{.p}$$

或处理效应等于 0，即

$$\alpha_j = 0$$

2. 区组的总体平均数相等

$$\mu_{1.} = \mu_{2.} = \cdots = \mu_{n.}$$

或区组效应等于 0，即

$$H_0 : \pi_i^2 = 0$$

因随机区组设计中假设区组变量和自变量间没有交互作用，所以在假设中就没有关于交互作用假设.

实验设计和计算举例

研究问题

我们仍利用第一节中文章的生字密度对阅读理解影响的研究做例子，由于考虑到学生的智力可能对阅读理解测验分数产生影响，但它又不是该实验中感兴趣的因素，我决定把学生的智力作为一个无关变量，通过实验设计将它的效应分离出去，以更好地探讨生字密度对阅读理解的影响. 我选用了单因素随机区组实验设计. 这时，我的研究假说、实验的自变量、因变量都是不变的，只是增加了一个无关变量. 在实验实施前，我首选要给 32 个学生做智力测验，并按智力将学生分为 8 个区组，然后随机分配每个区组内的 4 个同质被试分别阅读一种生字密度的文章.

数据计算

1. 计算表

Table 1.11: 单因素随机区组实验的 AS 表

	a_1	a_2	a_3	a_4	Σ
区组 1	3	4	8	9	24
区组 2	6	6	9	8	29
区组 3	4	4	8	8	24
区组 4	3	2	7	7	19
区组 5	5	4	5	12	26
区组 6	7	5	6	13	31
区组 7	5	3	7	12	27
区组 8	2	3	6	11	22
Σ	35	31	56	80	202

2. 各种基本量的计算

$$\begin{aligned}\frac{\left(\sum_{i=1}^n \sum_{j=1}^p Y_{ij}\right)^2}{np} &= [Y] = \frac{(202)^2}{(8)^2 (4)^2} = 1275.125 \\ \sum_{i=1}^n \sum_{j=1}^p Y_{ij}^2 &= [AS] = (3)^2 + (6)^2 + \cdots = 1544.000 \\ \sum_{i=1}^n \frac{\left(\sum_{j=1}^p Y_{ij}\right)^2}{n} &= [A] = \frac{(35)^2}{8} + \frac{(31)^2}{8} + \cdots = 1465.250 \\ \sum_{i=1}^n \frac{\sum_{j=1}^p (Y_{ij})^2}{p} &= [S] = \frac{(24)^2}{4} + \frac{(29)^2}{4} + \cdots = 1301.000\end{aligned}$$

3. 平方和的分解

定义 1.2.2 (单因素随机区组设计平方和的分解)

$$\begin{aligned}SS_{\text{总变异}} &= SS_{\text{处理间}} + SS_{\text{处理内}} \\ &= SSA + (SS_{\text{区组}} + SS_{\text{残差}})\end{aligned}$$

$$\begin{aligned}SS_{\text{总变异}} &= [AS] - [Y] = 268.875 \\ SSA &= [A] - [Y] = 190.125 \\ SS_{\text{区组}} &= [S] - [Y] = 25.750 \\ SS_{\text{残差}} &= SS_{\text{总变异}} - SSA - SS_{\text{区组}} = 52.875\end{aligned}$$

4. 方差分析表及结果解释

变异源	SS	df	MS	F	p
A (生字密度)	190.125	$p - 1 = 3$	63.375	25.170	< .001
区组 (智力)	25.850	$n - 1 = 7$	3.696	1.47	0.23
残差	52.875	$(n - 1)(p - 1) = 21$	2.518		
总计	268.875	$np - 1 = 31$			

Table 1.12: 单因素随机区组实验的方差分析表

方差分析可以看出, 实验中的自变量——生字密度的效应是统计显著的 ($F(3, 21) = 25.170, p < .001$), 说明学生对生字密度不同的文章的阅读理解有显著差别. 实验中无关变量——智力的效应是统计上不显著的 ($F(7, 21) = 1.47, p > .05$), 表明本实验中智力不同的学生的阅读理解没有明显差异. 方差分析表中还可以看出, 生字密度和智力的 F 检验都使用了同一个误差项 $MSE = 2.518$.

5. 平方和与自由度分解图

一些解释

1. 各种平方和的意义

$SS_{\text{总变异}}$ ——在随机区组实验中, 总平方和应首先分解为处理间平方和处理内平方和应首先分解为处理间平方和处理内平方和.

$SS_{\text{处理间}}$ ——处理间平方和指所有由实验处理引起的变异, 在单因素设计中指 A 因素的处理效应 SSA

$SS_{\text{处理内}}$ ——在随机区组实验中, 处理内平方和可进一步分解为两部分: 区组平方和和误差平方和.

$SS_{\text{区组}}$ ——区组效应, 在该实验中指总变异中由被试的智力引起的变异.

$SS_{\text{残差}}$ ——残差指总变异中不能被实验处理和区组效应解释的变异. 在随机区组实验设计中, 接受同实验条件的同质被试只有一个, 因此, 不能计算单元内误差, 而残差作为误差变异的估计. 残差的计算是从总变异中减去处理效应和区组变异.

2. 做区组的方法

除了感兴趣的自变量外, 还有很多无关变量都影响因变量. 如果发现一个变量特别影响因变量, 但对其效应不感兴趣, 可以考虑将其作为一个区组. 本例中区组变量是智力, 我们知道一个孩子的智力情况会影响他的阅读理解, 如果一个组里都是智力低的孩子, 一个组里都是智力高的孩子. 若智力高的组分配的是低生字密度, 生字密度和智力都会促进阅读理解, 处理效应被加大, 当本身处理效应不显著时, 会面临错误判定处理效应显著的失误; 若智力高的组分配的是高生字密度, 自变量效应方向和无关变量方向相反, 它们抵消后会使得处理效应被掩盖, 造成处理效应不显著.

3. 随机区组提高实验敏感性的体现

前面我提到从 F 值的大小思考实验的敏感性, 单因素完全随机的 F 值:

$$F = \frac{MSA}{MS_{\text{组内}}}$$

单因素随机区组设计将 SS 进一步分解为 $SS_{\text{区组}}$ 和 $SS_{\text{残差}}$, 既然区组变量带来的变异不是随机误差, 因此将其从组内误差中分离出来, 这样误差项就减小了, 从而单因素随机区组的 F 值就是:

$$F = \frac{MSA}{MS_{\text{残差}}}$$

而 $MS_{\text{残差}} < MS_{\text{组内}}$, 因而提升了实验的敏感性. 从数据上, 我们可以看到, 在表 ?? 中, 组内误差是 78.750, 在同样的数据用了随机区组的算法后, 表 ?? 中可以看出, 将组内误差分解成了区组带来的无关变异 25.850, 和剩余的残差 52.875. 完全随机的 F 是 22.533, 在随机区组中 F 是 25.170, 这看上去没有提升多少, 原因是这组数据中区组变量的效应是不显著的, 可以看到残差项还是很大, 说明这个随机区组变量并不是特别有效.

那么什么是有效的随机区组设计呢? 在表 ?? 中右边红色的数据是对把自变量处理水平 a_1, a_2, a_3, a_4 合并起来的数据, 这样这个数据只和集中趋势和区组无关变量带来的变异相关, 我们把它按小到大排个序:

19, 22, 24, 24, 26, 27, 29, 21

这和绿色的数据:

31, 35, 56, 80

相比, 变异就小了很多. 一个好的区组设计应该让红色的那部分数据呈现一个明显的梯度, 让它们之间相差得大一点, 这样就可以带走更多变异. 这点不是实验操作上可以保证的, 主要是要看这个区组变量是不是一个有效的区组变量, 是否对因变量有较大影响.

4. 残差的实质

从方差分析中的自由度 (见表 ??) 知, 残差的自由度是 $(p-1)(n-1)$, 这个形式表明这个残差本质是个交互作用. 现在还不是时候从根本上揭示残差的实质.

随机区组中, 用残差来估价随机误差, 今后我们将看到, 所有的误差项只有两种形式, 一种是单元内误差 (当有同质被试接受相同的处理时产生), 一种是残差 (实验是交互作用). 在随机区组设计中, 由于没有接受相同处理的同质被试, 故不存在单元内误差.

评价

随机区组的**优点**是, 在许多研究情境中, 它比完全随机实验设计更加有效. 这是由于它使研究者从总变异中分离出了一个无关变量的效应, 从而减小了实验误差, 可获得对处理效应的更加精确的估价. 随机区组实验设计可用于含任何处理水平数的实验中, 并且区组的数量也不受限制, 因而有较好的灵活性.

随机区组实验设计的**缺点**是, 如实验中含有多处理水平, 可能给形成同质区组、寻找同质被试带来困难. 另外, 使用随机区组设计比使用完全随机设计有更多的限定, 例如使用随机区组实验设计的前提假设是, 实验中的自变量与无关变量之间没有交互作用. 如果交互作用是存在的, 使用随机区组实验设计是不合适的. 这在一定程度上限制了随机区组实验设计的应用. 但由于无法对该残差的显著性进行检验, 故对于这个误差的估计是否合理就没有检验的方法.

经典随机区组设计的改进

基本思想

经典随机区组设计中, 误差用残差衡量, 它的实质是区组变量与自变量间的交互作用, 问题是我们无法保证它们间一定不存在交互作用, 我们甚至没有检验这个交互作用显著性的手段.

单元内误差是一定是一个可靠的误差估计, 不过在经典随机区组设计中没有接受相同处理的同质被试, 那对其进行改进即可, 让区组内人数扩大, 使接受同一个处理和在同一个区组内的被试有多个, 这样可以算出单元内误差, 检验残差是否显著.

评价

我们力求区组内被试同质, 但是实际上很难做到. 本例来说, 区组变量是智力, 这是一个连续变量. 增加区组内被试时, 有很大可有使区组内的变异变大, 很有可能让区组间的变小.

所以随机区组设计不会用这种方式.

1.3 单因素拉丁方实验设计

基本特点

拉丁方设计最大特色是同时计算了残差和单元内误差, 扩展了区组实验设计. 对于完全随机方差分析中的 F 检验而言:

$$F = \frac{MSA}{MS_{\text{组内}}} \quad MS_{\text{组内}} = \text{无关变异} + \text{随机误差}$$

由于完全随机设计中误差项 MSE 中混入了个体差异和一些其它的无关变异, 所以随机区组设计找到一个影响较大的无关变异, 计算其大小并从 MSE 中分离, 这样新的 MSE 就等于 $MSE_{\text{组内}} - MS_{\text{区组}}$, 误差项减小, 可以提高实验的精度. 不过这样还不够, 首先, 对于影响因变量的无关变量不止有一个, 也许除了生字密度学、生的智力外, 还有别的无关变量对阅读理解有较大影响. 故我们希望能再分离出一个无关变量的效应; 其次, 随机区组的误差是否是一个合理的误差也不清楚所以我还希望在每一个区组内加入多个被试, 以检验残差的是否显著.

总的来说, 拉丁方实验设计同时分离两个无关变量的效应, 同时通过分配多个被试给同一自变量水平和无关变量水平的结合, 使得可以计算出单元内误差来检验残差是否显著. 不过该设计对两个无关变量有较大要求. 拉丁方实验设计适合下列条件:

- 1. 自变量和无关变量的水平限制** 自变量的水平和两个无关变量的水平相等, 即自变量水平是 p ($p \geq 2$), 两个无关变量水平都是 p ($p \geq 2$);
- 2. 自变量与无关变量无交互作用** 如果这个假设不能满足峰实验中的一个或多个效应的检验可能有偏差
- 3. 拉丁方限制** 随机分配处理水平给 p^2 个方格单元, 每个处理水平仅在每行、每列中出现一次

Table 1.13: 2×2 拉丁方快

A	B
B	A

Table 1.14: 3×3 拉丁方快

A	B	C
B	C	A
C	A	B

Table 1.15: 4×4 拉丁方快

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Table 1.16: 5×5 拉丁方快

A	B	C	D	E
B	C	D	E	A
C	D	E	A	B
D	E	A	B	C
E	A	B	C	D

我列出了五个拉丁方格: 表 ??, 表 ??, 表 ??, 表 ??, 从列子中可以看到拉丁方格每一行每一列每一个字母只出现一次, 下面以 4×4 拉丁方表格介绍一下拉丁方格随机化的方法:

标准块					随机化行				
	1	2	3	4		1	2	3	4
1	A	B	C	D	3	C	D	A	B
2	B	C	D	A	1	A	B	C	D
3	C	D	A	B	2	B	C	D	A
4	D	A	B	C	4	D	A	B	C

随机化行					随机化列				
	1	2	3	4		4	3	1	2
3	C	D	A	B	3	B	A	C	D
1	A	B	C	D	1	D	C	A	B
2	B	C	D	A	2	A	D	B	C
4	D	A	B	C	4	C	B	D	A

Table 1.17: 拉丁方格随机化

单因素拉丁方实验设计被试分配到处处理上的例子如表 ??, 从表中可以看出, 实验中的自变量 A 有 4 个水平, 无关变量 B 和无关变量 C

也各有 4 个水平，形成 4×4 的拉丁方格，32 个被试参加了实验，每个方格内有 2 个被试，每个被试只接受一种独特的实验条件处理。

单因素拉丁方设计模型

定义 1.3.1 (单因素拉丁方设计模型)

$$Y_{ijkl} = \mu + \alpha_j + \beta_k + \gamma_l + \varepsilon_{pooled}$$

$$(j = 1, 2, \dots, p; i = 1, 2, \dots, n)$$

$$(k = 1, 2, \dots, p; l = 1, 2, \dots, p)$$

其中

Y_{ijkl}	-	两无关变量在第 jk 个水平上，被试 i 的分数
μ	-	总体平均值或真值
α_j	-	自变量水平 j 的处理效应
β_k	-	无关变量 B 水平 k 的变异
γ_l	-	无关变量 C 水平 l 的变异
ε_{pooled}	-	合并误差

模型中的 ε_{pooled} ，pooled 指的是“集合的，集中的”，这是两项误差的合并，这点我在后面会介绍。

单因素拉丁方实验设计适合检验的假说

(1) 处理水平的总体平均数相等，即：

$$H_0 : \mu_{1..} = \mu_{2..} = \dots = \mu_{p..}$$

或因素 A 效应等于 0，即：

$$H_0 : \alpha_j = 0$$

(2) 无关变量（横行）的总体平均数相等，即

$$H_0 : \mu_{.1.} = \mu_{.2.} = \dots = \mu_{.p.}$$

或无关变量 B 的效应等于 0，即

$$\beta_k = 0$$

(3) 无关变量（纵行）的总体平均数相等，即

$$H_0 : \mu_{..1} = \mu_{..2} = \dots = \mu_{..p}$$

或无关变量 C 的效应等于 0，即

$$\gamma_l = 0$$

Table 1.18: 单因素拉丁方实验设计中被试的分配

	c_1	c_2	c_3	c_4
	a_1	a_2	a_3	a_4
b_1	S_1	S_9	S_{17}	S_{25}
	S_2	S_{10}	S_{18}	S_{26}
b_2	a_2	a_3	a_4	a_1
	S_3	S_{11}	S_{19}	S_{27}
b_3	S_4	S_{12}	S_{20}	S_{28}
	a_3	a_4	a_1	a_2
b_4	S_5	S_{13}	S_{21}	S_{29}
	S_6	S_{14}	S_{22}	S_{30}
b_4	a_4	a_1	a_2	a_3
	S_7	S_{15}	S_{23}	S_{31}
	S_8	S_{16}	S_{24}	S_{32}

实验设计与计算举例

研究的问题与实验设计

我在做 4 种文章的生字密度对学生阅读理解影响时, 从 4 个班级随机选取 32 名学生, 每个班 8 人, 实验在周三、四、五、六下午分 4 次进行. 在这个研究中, 自变量——生字密度有 a_1, a_2, a_3, a_4 4 个水平, 考虑到来自不同班级的学生可能在阅读理解方面存在差异, 从而影响实验结果, 但班级的差异又不是研究者感兴趣的, 可以把班级做为一个带有 b_1, b_2, b_3, b_4 4 个水平的无关变量; 另外, 实验时间的不同也可能影响学生的情绪, 从而影响实验结果, 因此可将实验时间做为第 2 个无关变量, 它也有 4 个水平: c_1, c_2, c_3, c_4 . 实验实施前, 我需要首先建构一个 4×4 的拉丁方标准块, 将每个班级的 8 名学生随机分配在 c_1, c_2, c_3, c_4 的拉丁方格中, 每个方格中的两个学生接受完全相同的实验条件. 然后, 将拉丁方格标准块随机化, 并按随机块的方案实施实验. 举例如下:

Table 1.19: 拉丁方设计 ABCS 表

	c_1	c_2	c_3	c_4
	a_1	a_2	a_3	a_4
b_1	3	2	6	9
	4	3	5	8
	a_2	a_3	a_4	a_1
b_2	8	3	4	7
	7	2	3	6
	a_3	a_4	a_1	a_2
b_3	8	12	5	6
	9	13	6	4
	a_4	a_1	a_2	a_3
b_4	5	8	12	7
	4	7	11	5

Table 1.20: 拉丁方设计 ABC 表

	c_1	c_2	c_3	c_4	Σ
$n = 2$					
	a_1	a_2	a_3	a_4	
b_1	7	5	11	17	40
		a_2	a_3	a_4	
b_2	15	5	7	13	40
		a_3	a_4	a_1	
b_3	17	25	11	10	63
		a_4	a_1	a_2	
b_4	9	15	23	12	59
		a_1	a_2	a_3	
Σ	48	50	52	52	202

Table 1.21: 拉丁方设计 A 表

a_1	a_2	a_3	a_4
$np = 8$			
35	31	56	80

实验数据及其计算

1. 计算表

2. 各种基本量的计算

$$\sum_{i=1}^n \sum_{k=1}^p \sum_{l=1}^p Y_{ijkl} = 3 + 4 + \cdots = 202.000$$

$$\frac{\sum_{l=1}^p \sum_{k=1}^p \sum_{i=1}^n Y_{ijkl}}{np^2} = [Y] = \frac{(202)^2}{(2)(4)^2}$$

$$\sum_{l=1}^p \sum_{k=1}^p \sum_{i=1}^n (Y_{ijkl}^2) = [ABCS] = (3)^2 + (4)^2 + \cdots = 1544.000$$

$$\sum_{l=1}^p \sum_{k=1}^p \frac{\left(\sum_{i=1}^n Y_{ijkl} \right)^2}{n} = [ABC] = \frac{(7)^2}{2} + \frac{(5)^2}{2} + \cdots = 1533.000$$

$$= [A] = \frac{(35)^2}{8} + \frac{(31)^2}{8} + \cdots = 1465.250$$

$$\sum_{k=1}^p \frac{\left(\sum_{l=1}^p \sum_{i=1}^n Y_{ijkl} \right)^2}{np} = [B] = \frac{(40)^2}{8} + \frac{(40)^2}{8} + \cdots = 1331.250$$

$$\sum_{l=1}^p \frac{\left(\sum_{k=1}^p \sum_{i=1}^n Y_{ijkl} \right)^2}{np} = [C] = \frac{(48)^2}{8} + \frac{(50)^2}{8} + \cdots = 1276.500$$

3. 平方和分解与计算

定义 1.3.2 (单因素拉丁方设计平方和分解)

$$\begin{aligned} SS_{\text{总变异}} &= SS_{\text{处理间}} + SS_{\text{处理内}} \\ &= SSA + (SSB + SSC + SS_{\text{单元内}} + SS_{\text{残差}}) \end{aligned}$$

$$\begin{aligned} SS_{\text{总变异}} &= [ABCS] - [Y] &&= 268.875 \\ SSA &= [A] - [Y] &&= 190.125 \\ SSB &= [B] - [Y] &&= 56.125 \\ SSC &= [C] - [Y] &&= 1.375 \\ SS_{\text{残差}} &= [ABC] - [Y] - SSA - SSB - SSC &&= 10.250 \\ SS_{\text{单元内}} &= SS_{\text{总变异}} - SSA - SSB - SSC - SS_{\text{残差}} &&= 11.000 \end{aligned}$$

4. 方差分析表及对结果的解释

变异源	SS	df	MS	F	p
A (生字密度)	190.125	$p - 1 = 3$	63.375	25.170	<.001
B (班级)	56.125	$p - 1 = 3$	18.708	27.19	<.001
C (实验时间)	1.375	$p - 1 = 3$	0.458	0.67	0.58
残差	52.875	$(p - 1)(p - 2) = 6$	1.708	2.48	0.18
单元内误差	11.000	$p^2(n - 1) = 16$	0.688		
总计	268.875	$np^2 - 1 = 31$			

Table 1.22: 单因素拉丁方实验的方差分析表

5. 平方和与自由度的分解

一些解释

评价

交互作用难以保证，毕竟三个因素无关变量水平数要和自变量水平数相等，也难经保证 == 给我们很好的思想，但是不是特别常用

1.4 单因素被试内实验设计

基本特点

在单因素完全随机实验中，组内变异实际上由两部分组成的：实验中测量误差引起的变异和未控制的无关变量带来的变异，其中主要是被试个体差异带来的变异. 减少误差变异的一个方法是控制个体差异引起的无关变异，达互这个目标的途径之一是使用随机区组设计，带走一个无关变量引起的无关变异；另一个方法是使用拉丁方设计，带走两个无关变量引起的无关变量；因个体差异是无关变量引起的无关变异的主要来源，故一个更有效的方法是重要测量实验设计，在单因素实验设计中，重复测量设计也是被试内设计，这种设计的基本思路是分离出所有被试间个体差异.

前面介绍的完全随机设计、随机区组设计、拉丁方设计中，每一个被试都只接受一个处理，这种设计被称为被试间实验设计或非重复测量设计。完全随机设计中，误差用所有被试间的变异估价，随机区组设计和拉丁方设计进一步将其中的一个或两个无关变量带来的无关变异带走，可不论怎么样，总还是有被试个体差异混杂在误差变异中。

重复测量实验设计的基本方法是：实验中每个被试接受所有的处理水平。这种实验设计的目的是利用被试自己做控制，使被度的各方面特点在所有的处理中保持恒定，经最大限度地控制由被试的个体差异带来的变异。

使用重复测量设计的前提是研究者必须事先假设，当若干处理水一连续实施给同一被试时，被试接受前面的处理对接受后面的处理没有长期影响。重复测量设计在有些情况下是不合适的，当处理的实施对被试有长期影响时，如学习、记忆效应，不能使用重复测量设计。例如，在一个教学研究中，要比较两种教学方法对学生成绩的影响。我们不可能使用同一班学生先后接受两种教学方法，然后比较它们对学生学习成绩的影响，因为前一种教法的教学不可避免地对学生接受后一种教法的教学产生影响。在心理与教育研究中，话多实验处理会对被试产生学习、记忆效应，因此用重复测量设计要特别谨慎。

另外顺序效应也是重复测量设计中应特别注意的问题。被试连续接受处理时，练习、疲劳等效应是难免的，因此重复测量中需要考虑平衡顺序效应的问题。

单因素被试设计模型

定义 1.4.1 (单因素被试设计模型)

$$Y_{ij} = \mu + \alpha_j + \pi_i + (\alpha\pi)_{ij} + \varepsilon_{ij} \\ (j = 1, 2, \dots, p; i = 1, 2, \dots, n)$$

其中

- Y_{ij} - 两无关变量在第 jk 个水平上，被试 i 的分数
- μ - 总体平均值或真值
- α_j - 自变量水平 j 的处理效应
- π_i - 无关变量 C 水平 i 的变异
- $(\alpha\pi)_{ij}$ - 水平 α_j 和被试 i 的交互作用
- ε_{ij} - 四舍五入剩余的点误差，不体现在平方和分解中

单因素重复测量实验设计检验的假说

处理水平的总体平均数相等，即：

$$H_0: \mu_1 = \mu_2 = \dots = \mu_p$$

或处理效应等于这，即

$$\alpha_j = 0$$

实验设计与计算举例

研究的问题与实验设计

我们继续以 4 种文章的生字密度对学生阅读理解的影响的研究为例. 为了更好地控制被试变量, 我仅用 8 名被试, 每个被试阅读 4 篇生字密度不同的文章, 并测他们对各篇文章的阅读理解分数. 选择使用重复测量实验设计是由于我假设, 当实验安排合适时, 被试阅读一篇文章不会对阅读另一篇文章产生影响. 但是, 在这种实验设计中, 疲劳效应和顺序效应是必须考虑的. 为了减少疲劳效应, 研究者决定将 4 篇文章在四个下午分 4 次施测. 平衡顺序效应的方式有两种: 以随机顺序实施 4 和种生字密度的文章, 或以拉丁方排序实施 4 种生字密度的文章. 后一平衡效应的方法举例如表 ??.

Table 1.23: 单因素被试实验设计平衡顺序效应

		呈现方式			
		1	2	3	4
被试	S_1, S_2	a_1	a_2	a_3	a_4
	S_3, S_4	a_2	a_3	a_4	a_1
	S_5, S_6	a_3	a_4	a_1	a_2
	S_7, S_8	a_4	a_1	a_2	a_3

实验数据及其计算

1. 计算表

单因素被试内设计中的计算表有一张 (表 ??), 这张表和随机区组设计的表很像, 不过为了突显被试内设计可以带走更多的无关变异, 我对数据进行了一些调整. 表 ?? 中红色的数据相比表 ?? 中的红色数据呈现一个更加明显的梯度, 也就是区组之间的变异小于被试间的变异, 而下方绿色的数据仍然不变. 之后在结果中可以看到被试的 F 是更大的.

2. 各种基本量的计算

$$\frac{\left(\sum_{i=1}^n \sum_{j=1}^p Y_{ij}\right)^2}{np} = [Y] = \frac{(202)^2}{(8)^2 (4)^2} = 1275.125$$
$$\sum_{i=1}^n \sum_{j=1}^p Y_{ij}^2 = [AS] = (3)^2 + (6)^2 + \cdots = 1544.000$$
$$\sum_{i=1}^n \frac{\left(\sum_{j=1}^p Y_{ij}\right)^2}{n} = [A] = \frac{(35)^2}{8} + \frac{(31)^2}{8} + \cdots = 1465.250$$
$$\sum_{i=1}^n \frac{\sum_{j=1}^p (Y_{ij})^2}{p} = [S] = \frac{(19)^2}{4} + \frac{(19)^2}{4} + \cdots = 1330.500$$

3. 平方和的分解与计算

定义 1.4.2 (单因素被试内和方分解)

$$SS_{\text{总变异}} = SS_{\text{被试间}} + SS_{\text{被试内}}$$
$$= SS_{\text{被试间}} + (SS_A + SS_{\text{残差}})$$

Table 1.24: 单因素被试内设计的 AS 表

	a_1	a_2	a_3	a_4	Σ
被试 1	4	3	5	7	19
被试 2	3	2	6	8	19
被试 3	3	5	7	11	26
被试 4	2	4	6	9	21
被试 5	5	4	7	8	24
被试 6	5	4	8	12	29
被试 7	6	3	8	12	29
被试 8	7	6	9	13	35
Σ	35	31	56	80	202

$$SS_{\text{总变异}} = [AS] - [Y] = 268.875$$
$$SS_{\text{被试间}} = [S] - [Y] = 55.375$$
$$SS_{\text{被试内}} = SS_{\text{总变异}} - SS_{\text{被试间}} = 213.500$$
$$SS_A = [A] - [Y] = 190.125$$
$$SS_{\text{残差}} = SS_{\text{被试内}} - SS_A = 23.375$$

计算时，被试内设计第一步就拿走了被试间个体差异，完全不参与计算. 下面的方差分析表 (表 ??) 中也可以看到，第一排就是被试间变异，我们没有检验其显著性，因为这就是一个无关变异，并且它显著或是不显著都不是我们所感兴趣的.

4. 方差分析表及对结果的解释

变异源	SS	df	MS	F	p
被试	55.375	$n - 1 = 7$	7.911		
A (生字密度)	190.125	$p - 1 = 3$	63.375	56.936	< .001
残差	23.375	$(n - 1)(p - 1) = 21$	1.113		
总计	268.875	$np - 1 = 31$			

Table 1.25: 单析表

方差分析表中可以看出，实验中的自变量——生字密度的效应是统计显著的 ($F(3, 21) = 25.17, p < .01$)，自变量的 F 检验的误差项是 $MSE = 1.113$. 方差分析表中还有一项“被试间变异”($SS_{\text{被试间}} = 55.375$)，这部分变异事业有所有被试之间个体差异引起的变异. 由于从总变异中分享出了被试间变异，因此与完全随机实验相比，重复测量实验中提高了实验处理的 F 检验的敏感性.

5. 平方和与自由度分解图解

一些解释

- $SS_{\text{总变异}}$ ——在重复测量实验中，总变异应首先分解为被试间平方和被试内平方和
- $SS_{\text{被试间}}$ ——被试间平方和，即总变异中所有由被试的个体差异引起的变异
- $SS_{\text{被试内}}$ ——被试内平方和包括同一被试在接受不同实验处理时产生的变异（铺好床吖啊内因素的处理效应），以有偶然因素引起的实验误差. 在单因素重复测量实验设计中，被试内平方和被分解为两个部分：A 因素的处理效应和误差变异
- $SS_{\text{残差}}$ ——重复测量实验中的残差变异与随机区组实验中的残差性质相同，重复测量实验方差分析中残差计算是先从总变异中减去被试间平方和，然后再减去处理效应. 由于事先从总变异中分离出了所有的由被试个体差异带来的变异，重复测量实验中的 $SS_{\text{残差}}$ 一般很小

评价

1.5 四种实验设计比较

1. 被试数量

实验设计	被试数	字母说明
单因素完全随机	np	自变量 p 个水平, 每个组内均 n 个被试
单因素随机区组	np	自变量 p 个水平, n 个区组
单因素拉丁方	np^2	无关变量及自变量均 p 个水平, 拉丁方格内 n 个被试
单因素被试内	n	n 个被试

拉丁方在本章中所用的被试数和其它非重复测量的被试数一样, 其原因是它的单元内被试数只有 2, 而自变量与无关变量水平数相等, 都是 4, $2 \cdot 4^2 = 8 \cdot 4$. 可以看到被试内实验设计非常节省被试数.

2. 平方和分解模式

平方和的分解是根据实验设计模型而定, 这个模型指导着观察分数的来源, 不同的分数组成部分由不同的变异源引起, 故我们假定它们都相互独立, 由此可以推算出下面的平方和分解模式.

可以清楚地看到被试间实验设计在不断寻找处理内变异中的无关变异, 而被试内干脆直接把被试个体差异全部带朝走, 虽然被试间个体差异确实是个误差, 但却不进入计算.

实验设计的核心就是要分离出误差效应, 我们期望这部分误差尽可能小, 让处理效应更容易显著, 从单因素实验设计中我们已经看到不同的实验设计是如何一步一步减小误差的.

实验设计	平方和分解
单因素完全随机	$SS_{\text{总变异}} = SS_{\text{处理间}} + SS_{\text{处理内}}$ $= SSA + SS_{\text{单元内}}$
单因素随机区组	$SS_{\text{总变异}} = SS_{\text{处理间}} + SS_{\text{处理内}}$ $= SSA + (SS_{\text{区组}} + SS_{\text{区组}})$
单因素拉丁方	$SS_{\text{总变异}} = SS_{\text{处理间}} + SS_{\text{处理内}}$ $= SSA + (SSB + SSC + SS_{\text{残差}} + SS_{\text{单元内}})$
单因素被试内	$SS_{\text{总变异}} = SS_{\text{被试间}} + SS_{\text{被试内}}$ $= SS_{\text{被试间}} + (SSA + SS_{\text{残差}})$

3. 误差项自由度

实验设计	误差项自由度
单因素完全随机	$(n-1)p$
单因素随机区组	$(n-1)(p-1)$
单因素拉丁方	$(n-1)(p-2)$
单因素被试内	$(n-1)(p-1)$

这么看来完全随机实验设计也不是一定就不好，至少它的误差自由度很大， F 临界值会更小一些，而其它的设计虽然使 F 值变大，不过其临界值也增大了。所以 F 值变大并不一定意味着处理效应更容易显著，毕竟误差自由度也变化了，要综合考虑。

本章开头我给出的一个练习中我们直观在数据中计算了 MSA 和 MSE 变化的情况，加上单因素实验设计的四种实验设计我们得到了更多对于数据的认识。现在我们简要总结一下减小 SSE 的方式

1. **选取更加同质的被试** 还是以测量学生在新旧教学方法条件下语文成绩为例，如果我真的从全部人类的部体中随机抽样，结果我抽样出一个 7 岁的小学生和一个 20 岁的大学生，这样结果就没法比了。因此在取样中我们通常是带有目的选取尽可能同质的被试，尽管对我们把结论推论到总体有一定的影响。
2. **带走无关变异** 随机区组、拉丁方、重复测量正用此方法

我们介绍完了单因素实验设计。从两因素开始因为有了多个因素，交互作用出现了，我们将获得更多信息；混合实验设计也出现了。上述两者是多因素实验设计的核心。

基本特点	
实验设计与计算举例	
一些解释	
基本特点	
实验设计与计算举例	
一些解释	

2 两因素被试间实验设计

2.1 两因素交互作用的进一步探讨

交互作用含义

前面我用教学方法和智力的例子大体说明了这个问题，教学方法对阅读成绩的成绩不仅与其本身是新教学方法还是旧教学方法有关，也受到学生智力的影响。换句话说，高智力学生可以接受到新教学方法的好处，相反低智力学生不仅不能体会其好处，而且可能因为他们不适应新的教学方法导致成绩不增反降。因此我们得到交互作用的定义：

定义 2.1.1 (交互作用) 指的是一个因素的的效应在一个或多个因素不同的水平上不同。

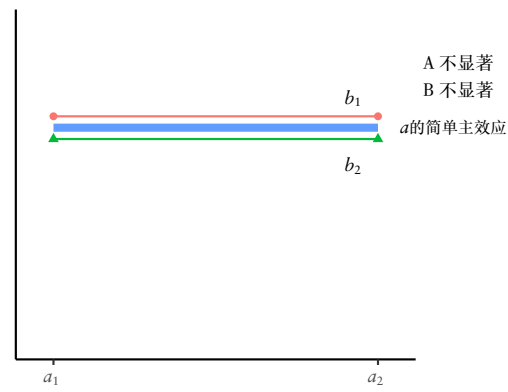
图 1.1 on page 4展示了两因素交互作用可能存在 8 种情况，我将单因素时得到的结果用蓝色线段加粗标出来，可以看到，A 因素主效应不显著会对应 (1) (4) 四种情况，而只有 (1) 和 (2) 是真的可以说明因素 A 对因变量没有影响；(3) 和 (4) 这种情况下，A 因素的效应在 B 因素的两个水平下是相反的，而且总体上恰好抵消了，显得好像 A 因素对因变量没有效应似的，在多因素解释效应时要小心。

在交互作用显著的 (3)(4)(7)(8) 中，(3)(4)(7) 表示的交互作用中两条线段交叉，表明 A 的效应在不同 B 的水平上相反，称为**非同序交互作用 (disordinal interaction)**。(8) 表示的交互作用中，B 不同组中 A 的效应还是一个方向的，称之为**同序交互作用 (ordinal interaction)**。

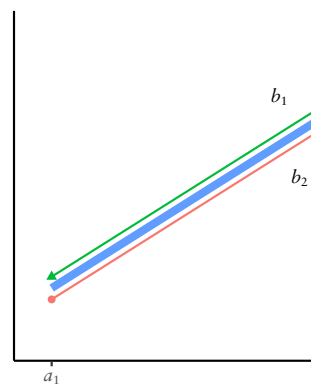
交互作用图解

当方差分析表明两个因素的交互作用是显著的时候，研究者常常需要进一步了解交互作用的含义是什么。一种简单的方法是画出交互作用的图解，以观察一个因素的各水平在另一个因素的每个水平上的变化。作图解时，应首先计算出每个处理水平结合中所得到的平均数，然后以平均数作图。我们仍举上面两因素完全随机设计的例子，列出它的 AB 表 (表 1.1 和表 1.2)

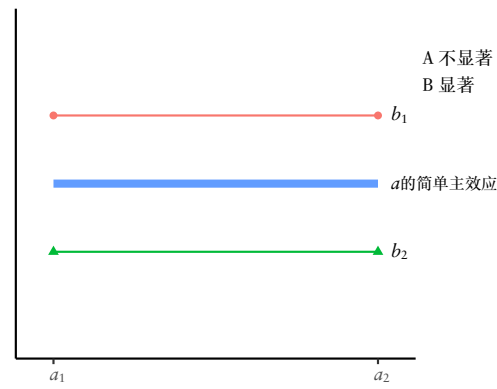
交互作用不显著



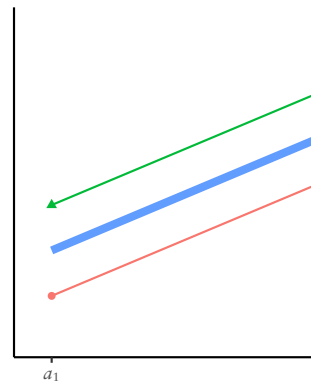
(1)



(5)

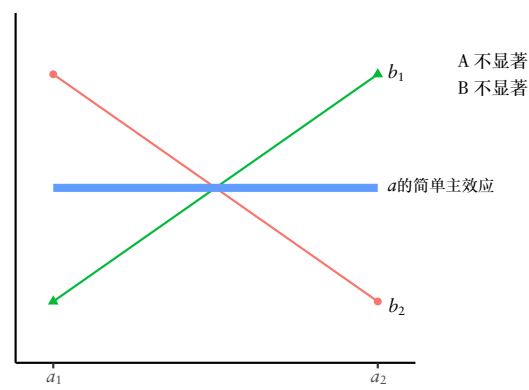


(2)

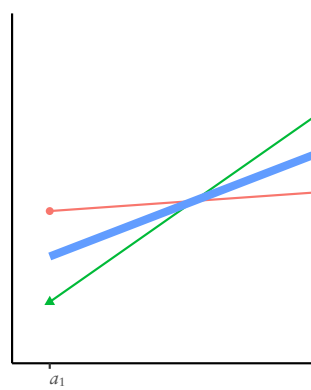


(6)

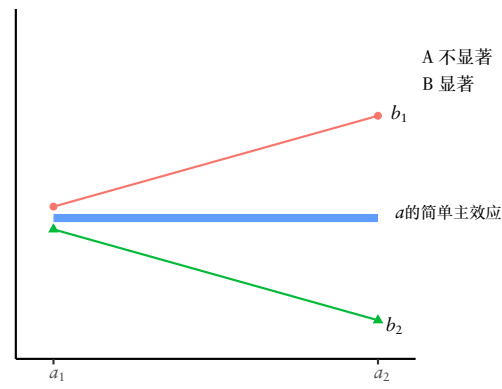
交互作用显著



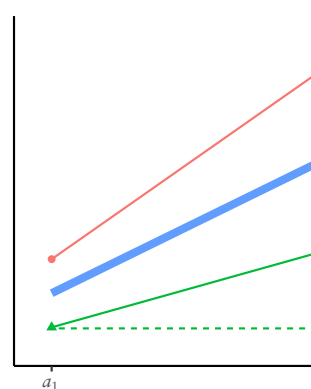
(3)



(7)



(4)



(8)

Figure 2.1: 两因素交互作用可能的 8 种情况，中间蓝色的加粗线为 A 因素的简单主效应，可以看到只考虑 (1) 和 (5) 中三条线本是重合的，但是为了可以看得清楚故将其稍微拉开了一点，(8) 中绿色的虚线和实线表示两种情况

根据 AB 平均数表, 可双两个方向作用 (图 1.2)

对两个图的解释是不一样的, 图 1.2(a) 表明, B 因素在 A 因素的两个水平上影响趋势是不一致的. B 因素的三个水平在 a_1 水平似乎没有明显差别, 而在 a_2 水平有较大差异. 图 1.2(b) 表明, A 因素在 B 因素的三个水平上的影响趋势也不一致. A 因素的两个水平在 b_1 水平没有明显差异, 而在 b_2, b_3 水平存在较大差异.

图解法的优点是简单、直观. 直接利用各处理水平结合所得的平均观测值作图, 可以使读者对结果模式有一个非常直观的了解. 它的弱点是, 解释是主观的, 有时不同的研究者可能会对同一结果作出不同的解释, 尤其是当出现复杂的交互作用时, 研究者单靠经验、直观进行解释是很困难的. 因此, 图解一般只能作为检查交互作用的第一步, 它需要同统计检验结合起来, 以便进一步用数据对交互作用的意义作出更精确可靠的解释.

简单效应检验

基本特点

检查交互作用含义的另一个方法是简单效应检验, 简单效应检验与主效应检验不大相同. 主效应检验是在忽略其它因素的情况下检验一个因素的处理效应, 我们比较一下主效应和简单效应的不同之处.

设观测值为 Y_{ijk} , i 是处理组合内的第 i 个被试, j 是因素 A 的水平 j , k 是因素 B 的水平 k

1. A 因素和 B 因素的主效应用组内平均数与真值的差估计, 组内平均数忽略了被试和其它因素, 可表示为

$$\begin{aligned} \alpha_1 &= \mu_{.1} - \mu & \beta_1 &= \mu_{.1} - \mu \\ \alpha_2 &= \mu_{.2} - \mu & \beta_2 &= \mu_{.2} - \mu \\ &\vdots & &\vdots \\ \alpha_j &= \mu_{.j} - \mu & \beta_k &= \mu_{.k} - \mu \\ \alpha_p &= \mu_{.p} - \mu & \beta_q &= \mu_{.q} - \mu \end{aligned}$$

可检验的假说是:

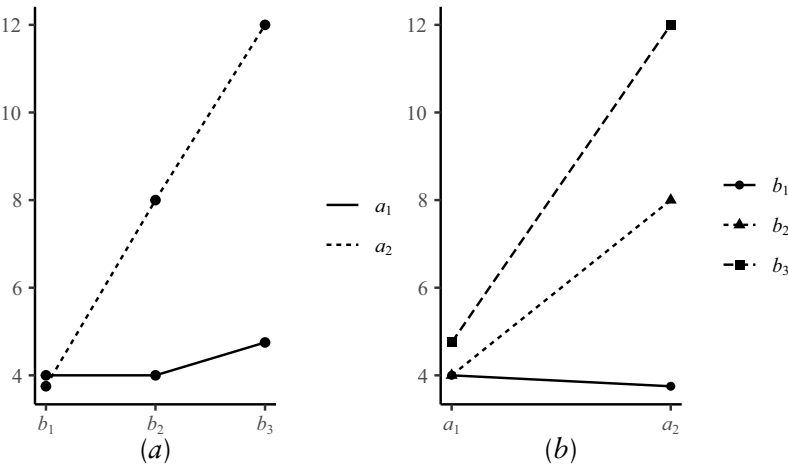


Table 2.1: 两因素完全随机设计 ABS 表

	a_1	a_2
$n = 4$		
b_1	16	15
b_2	16	32
b_3	19	48

Table 2.2: 两因素完全随机设计 AB 平均数表

	a_1	a_2
b_1	4	3.75
b_2	4	8
b_3	4.75	12

Figure 2.2: AB 交互作用两个方向的图

(1) A 因素处理水平上的总体平均数相等, 即:

$$H_0: \mu_{.1} = \mu_{.2} = \cdots = \mu_{.p}$$

或 A 因素处理效应为 0, 即:

$$\alpha_j = 0$$

(2) B 因素处理水平上的总体平均数相等, 即:

$$H_0: \mu_{.1} = \mu_{.2} = \cdots = \mu_{.q}$$

或 B 因素处理效应为 0, 即:

$$\beta_k = 0$$

2. 简单效应

本例中, 对于因素 A, 我们分别计算它的效应在 b_1, b_2, b_3 上的变异, 图 1.2(b) 上三条线段因为斜率不一样所以我们怀疑可能存在交互作用, 而上述拆分的三个变异的显著性正好反应这条线在统计上是显著的平还是显著的斜如果 A 在 b_1 上的变异不显著, 而在 b_2, b_3 上的变异显著, 则说明确实存在交互作用.

所以说简单效应的计算也就是将一个因素对因变量的效应拆分成多组, 有另一个因素有多少水平就拆分成多少个组, 我们检验它在另一个因素不同水平上的显著性.

$$\begin{array}{ll} \alpha_1(\text{在 } b_1 \text{ 水平上}) = \mu_{.11} - \mu_{..1} & \beta_1(\text{在 } a_1 \text{ 水平上}) = \mu_{.11} - \mu_{.1.} \\ \alpha_2(\text{在 } b_1 \text{ 水平上}) = \mu_{.21} - \mu_{..1} & \beta_2(\text{在 } a_1 \text{ 水平上}) = \mu_{.12} - \mu_{.2.} \\ \vdots & \vdots \\ \alpha_j(\text{在 } b_1 \text{ 水平上}) = \mu_{.j1} - \mu_{..1} & \beta_k(\text{在 } a_1 \text{ 水平上}) = \mu_{.1k} - \mu_{.1.} \\ \vdots & \vdots \\ \alpha_j(\text{在 } b_k \text{ 水平上}) = \mu_{.jk} - \mu_{..k} & \beta_k(\text{在 } a_j \text{ 水平上}) = \mu_{.jk} - \mu_{.j.} \\ \vdots & \vdots \\ \alpha_p(\text{在 } b_q \text{ 水平上}) = \mu_{.pq} - \mu_{..q} & \beta_q(\text{在 } a_p \text{ 水平上}) = \mu_{.pq} - \mu_{.p.} \end{array}$$

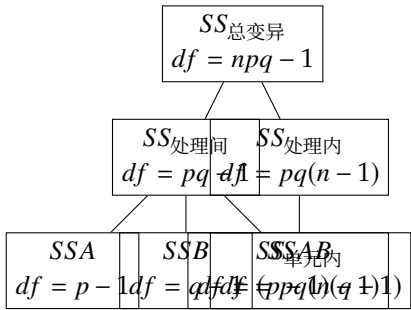
计算举例

使用

变异源	SS	df	MS	F	p
A	80.667	$p - 1 = 1$	80.667	43.343	< .001
B	81.083	$q - 1 = 2$	40.542	21.784	< .001
$A \times B$	56.583	$(p - 1)(q - 1) = 2$	28.292	15.201	< .001
$A_{(在b_1水平)}$	0.125	$p - 1 = 1$	0.125	0.067	0.798
$A_{(在b_2水平)}$	32.000	$p - 1 = 1$	32.000	17.194	< .001
$A_{(在b_3水平)}$	105.125	$p - 1 = 1$	105.125	56.485	< .001
$B_{(在a_1水平)}$	1.500	$q - 1 = 2$	0.750	0.403	0.674
$B_{(在a_2水平)}$	136.167	$q - 1 = 2$	68.083	36.582	< .001
单元内误差	33.500	$pq(n - 1) = 18$	1.861		

Table 2.3: 两因素完全随机方差分析与两个方向简单效应检验

2.2 两因素随机机区组实验设计



3 两因素重复测量实验设计

基本特点

两因素混合实验设计模型

定义 3.0.1 (两因素混合实验设计模型)

$$Y_{ijk} = \mu + \alpha_j + \pi_{i(j)} + \beta_k + (\alpha\beta)_{jk} + (\beta\pi)_{ki(j)} + \varepsilon_{ijk}$$
$$(j = 1, 2, \dots, p; k = 1, 2, \dots, q; i = 1, 2, \dots, n)$$

其中

- μ - 总体平均数, 或真值
- α_j - A 因素的水平 j 的处理效应
- $\pi_{i(j)}$ - 嵌套在 α_j 水平内的被试 i 的效应
- β_k - B 因素的水平 k 的处理效应
- $(\alpha\beta)_{jk}$ - 水平 α_j 和水平 β_k 的交互作用
- $(\beta\pi)_{ki(j)}$ - 嵌套在 β_k 水平和被试 i 的交互作用的残差
- ε_{ijk} - 误差

Table 3.1: Add caption

	b_1	b_2	b_3
a_1	S_1	S_1	S_1
	S_2	S_2	S_2
	S_3	S_3	S_3
	S_4	S_4	S_4
a_2	S_5	S_5	S_5
	S_6	S_6	S_6
	S_7	S_7	S_7
	S_8	S_8	S_8

两因素混合实验设计检验的假说

实验设计与计算举例

研究的问题

- 计算表
- 各种基本量的计算
- 平方和分解与计算

定义 3.0.2 (两因素混合实验设计平方和分解)

$$SS_{\text{总变异}} = SS_{\text{被试间}} + SS_{\text{被试内}}$$
$$= (SSA + SS_{\text{被试}(A)}) + (SSB + SSAB + SS_{B \times \text{被试}(A)})$$

Table 3.2: 两因素混合实验设计方差分析表

变异源	SS	df	MS	F	
A(主题熟悉性)	80.667	$p - 1 = 1$	80.667	15.869	0
被试间 (A)	30.500	$p(n - 1) = 6$	5.083		
B(生字密度)	81.083	$q - 1 = 2$	40.542	162.167	<
AB	56.583	$(p - 1)(q - 1) = 2$	28.292	113.167	<
B× 被试 (A)	3.000	$p(n - 1)(q - 1) = 12$	0.250		

$$\begin{aligned}
SS_{\text{总变异}} &= [ABS] - [Y] = 251.833 \\
SS_{\text{被试间}} &= [AS] - [Y] = 111.166 \\
SS_A &= [A] - [Y] = 80.667 \\
SS_{\text{被试(A)}} &= SS_{\text{被试间}} - SS_A = 30.500 \\
SS_{\text{被试内}} &= SS_{\text{总变异}} - SS_{\text{被试间}} = 140.667 \\
SS_B &= [B] - [Y] = 81.083 \\
SS_{AB} &= [AB] - [Y] - SS_A - SS_B = 56.584 \\
SS_{B \times \text{被试(A)}} &= SS_{\text{被试内}} - SS_B - SS_{AB} = 3.000
\end{aligned}$$

4. 方差分析表及结果的解释

5. 平方和与自由度分解图

一些解释

1. 各种平方和的意义

2. $SS_{\text{被试(A)}}$ 的实质

忽略 B 因素, 将其各个水平的值求和, 得到表 2.3 的计算表, 观察这个表的被试分配到处理的模式可以看到, 因素 A 的两个水平下, a_1 水平下有 4 个被试, 每个被度仅接受一种处理; a_1 水平下有 4 个被试, 每个被度仅接受一种处理. 这样就好像是一个单因素完全随机设计. 接下来我们计算这个完全随机设计的单元内误差, 注意算的时候要忘记 B 当成完全随机, 但是被试数要带上 B 因素的被试数, 否则平方和会拉大.

Table 3.3: 忽略 B 因素 (被试内变量) 的 AS 表

	a_1	a_2
$n = 3$		
$S_1 = 12$		$S_5 = 24$
$S_2 = 19$		$S_6 = 27$
$S_3 = 13$		$S_7 = 23$
$S_4 = 7$		$S_8 = 21$
Σ	51	95

(1) 基本量的计算

$$\begin{aligned}
[Y] &= \frac{(146)^2}{24} = 888.167 \\
[AS] &= \frac{(12)^2}{3} + \frac{(19)^2}{3} + \cdots = 999.333 \\
[A] &= \frac{(51)^2}{12} + \frac{(95)^2}{12} = 968.833
\end{aligned}$$

(2) 平方和的计算

$$SS_{\text{总变异}} = [AS] - [Y] = 111.166$$

$$SS_{\text{组间}} = [A] - [Y] = 80.667$$

$$SS_{\text{组内}} = SS_{\text{总变异}} - SS_{\text{组间}} = 30.500$$

所以在混合因素设计中的

$$F = \frac{MSA}{MS_{\text{被试}(A)}} = \frac{SSA/(p-1)}{SS_{\text{被试}(A)}/(n-1)p}$$

类似于完全随机设计中的

$$F = \frac{MS_{\text{组间}}}{MS_{\text{组内}}} = \frac{SS_{\text{组间}}/(p-1)}{SS_{\text{组内}}/p(n-1)}$$

上面 $SS_{\text{组内}}$ 是用相减法算得，因其实质是单元内误差，也可以直接计算，其算法我们也已知晓，即每个处理内的被试间的变异。

$$SS_{1\text{组}} = \frac{(12)^2}{3} + \frac{(19)^2}{3} + \frac{(13)^2}{3} + \frac{(7)^2}{3} - \frac{(51)^2}{12} = 24.25$$

$$SS_{2\text{组}} = \frac{(24)^2}{3} + \frac{(27)^2}{3} + \frac{(23)^2}{3} + \frac{(21)^2}{3} - \frac{(95)^2}{12} = 6.25$$

$$SS_{\text{组内}} = 24.25 + 6.25 = 30.5$$

3. $SS_{B \times \text{被试}(A)}$ 的实质

$$[Y_1] = \frac{(51)^2}{12} = 216.75$$

$$[B_1 S_1] = (3)^2 + (4)^2 + \dots = 245.000$$

$$[S_1] = \frac{(12)^2}{3} + \frac{(19)^2}{3} + \frac{(13)^2}{3} + \frac{(7)^2}{3} = 241$$

$$[B_1] = \frac{(16)^2}{4} + \frac{(16)^2}{4} + \frac{(19)^2}{4} = 218.25$$

$$[Y_2] = \frac{(95)^2}{12} = 752.803$$

$$[B_2 S_2] = (4)^2 + (8)^2 + \dots = 888.250$$

$$[S_2] = \frac{(24)^2}{3} + \frac{(27)^2}{3} + \frac{(23)^2}{3} + \frac{(21)^2}{3} = 758.333$$

$$[B_2] = \frac{(15)^2}{4} + \frac{(32)^2}{4} + \frac{(48)^2}{4} = 888.25$$

Table 3.4: 在 a_1 水平的 BS 表

	b_1	b_2	b_3	Σ
S_1	3	4	5	12
S_2	6	6	7	19
S_3	4	4	5	13
S_4	3	2	2	7
Σ	16	16	19	51

$$\begin{aligned}
SS_1 &= [B_1 S_1] - [Y_1] = 28.25 \\
SSB_1 &= [B_1] - [Y_1] = 1.5 \\
SS_{\text{被试}_1} &= [S_1] - [Y_1] = 24.25 \\
SS(B_1 S_1) &= [B_1 S_1] - [Y_1] - SSB_1 - SS_{\text{被试}_1} = 2.5
\end{aligned}$$

$$\begin{aligned}
SS_2 &= [B_2 S_2] - [Y_2] = 142.917 \\
SSB_2 &= [B_2] - [Y_2] = 136.167 \\
SS_{\text{被试}_2} &= [S_2] - [Y_2] = 6.25 \\
SS(B_2 S_2) &= [B_2 S_2] - [Y_2] - SSB_2 - SS_{\text{被试}_2} = 0.5
\end{aligned}$$

然后合并两个残差

$$\begin{aligned}
MS_{\text{残差}(pooled)} &= \frac{SSB_1 S_1 + SSB_2 S_2}{(n-1)(q-1) + (n-1)(q-1)} \\
&= \frac{2.5 + 0.5}{12} \\
&= 0.25
\end{aligned}$$

Table 3.5: 在 a_2 水平的 BS 表

	b_1	b_2	b_3	Σ
S_5	4	8	12	24
S_6	5	9	13	27
S_7	3	8	12	23
S_8	3	7	11	21
Σ	15	32	48	95

Table 3.6: Add caption

a_1	a_1	a_1	a_2	a_2	a_2
b_1	b_2	b_3	b_1	b_2	b_3
S_1	S_1	S_1	S_1	S_1	S_1
S_2	S_2	S_2	S_2	S_2	S_2
S_3	S_3	S_3	S_3	S_3	S_3
S_4	S_4	S_4	S_4	S_4	S_4

基本特点

两因素被试内实验设计模型

定义 3.0.3 (两因素被试内实验设计模型)

$$Y_{ijk} = \mu + \alpha_j + (\alpha\pi)_{ji} + \beta_k + (\beta\pi)_{ki} + (\alpha\beta)_{jk} + (\alpha\beta\pi)_{jki} + \varepsilon_{ijk}$$

$(j = 1, 2, \dots, p; k = 1, 2, \dots, q; i = 1, 2, \dots, n)$

其中

- μ - 总体平均数或真值
- π_i - 由被试 i 引起的变异, 被试间变异
- α_j - A 因素的水平 j 引起的处理效应
- $(\alpha\pi)_{ji}$ - 水平 α_j 和被试 i 的交互作用
- β_k - B 因素的水平 k 引起的处理效应
- $(\beta\pi)_{ki}$ - 水平 β_k 和被试 i 的交互作用
- $(\alpha\beta)_{jk}$ - 水平 α_j 和水平 β_k 的交互作用
- $(\alpha\beta\pi)_{jki}$ - 水平 α_j, β_k 和被试 i 的交互作用

Table 3.7: 两因素被试内实验设计方差分析表

变异源	SS	df	MS	F	p
被试间	27.167	$n - 1 = 3$			
A (主题熟悉性)	80.667	$p - 1 = 1$	80.667	72.600	0.003
$A \times$ 被试	3.333	$(p - 1)(n - 1) = 3$	1.111		
B	81.083	$q - 1 = 2$	40.542	153.632	< .001
$B \times$ 被试	1.583	$(n - 1)(q - 1) = 6$	0.264		
$A \times B$	56.583	$(p - 1)(q - 1) = 2$	28.292	119.824	< .001
$A \times B \times$ 被试	1.417	$(n - 1)(p - 1)(q - 1) = 6$	0.236		

两因素被试内实验设计检验的假说

实验设计与计算举例

研究的问题

1. 计算表
2. 各种基本量的计算
3. 平方和分解与计算

定义 3.0.4 (两因素被试内实验设计平方和分解)

$$\begin{aligned}
 SS_{\text{总变异}} &= SS_{\text{被试间}} + SS_{\text{被试内}} \\
 &= SS_{\text{被试间}} + (SSA + SS_{A \times \text{被试}} + SSB + SS_{B \times \text{被试}} + SSAB + SS_{A \times B \times \text{被试}})
 \end{aligned}$$

$$\begin{aligned}
 SS_{\text{总变异}} &= [ABS] - [Y] &= 251.833 \\
 SS_{\text{被试间}} &= [S] - [Y] &= 27.167 \\
 SS_{\text{被试内}} &= SS_{\text{总变异}} - SS_{\text{被试间}} &= 224.667 \\
 SSA &= [A] - [Y] &= 80.667 \\
 SS_{A \times (\text{被试})} &= [AS] - [Y] - SSA - SS_{\text{被试间}} &= 3.333 \\
 SSB &= [B] - [Y] &= 81.083 \\
 SS_{B \times (\text{被试})} &= [BS] - [Y] - SSB - SS_{\text{被试间}} &= 1.583 \\
 SSAB &= [AB] - [Y] - SSA - SSB &= 56.583 \\
 SS_{A \times B \times \text{被试}} &= SS_{\text{被试内}} - SSA - SS_{A \times \text{被试}} - SSB - SS_{B \times \text{被试}} - SSAB = 1.417
 \end{aligned}$$

4. 方差分析表及结果的解释

5. 平方和与自由度分解图

一些解释

1. 各种平方和的意义

