

实验研究的设计及分析

杨志宏

2019-12-01

献给……

呃，爱谁谁吧

目录

第一章 实验设计概述	1
1.1 实验设计发展趋势	1
1.1.1 实验室研究	1
1.1.2 社会研究	1
1.1.3 多因素实验设计的特点	2
1.2 实验设计中的基本概念	2
1.2.1 因素与因素实验设计	2
1.2.2 处理与处理水平的结合	2
1.3 主效应与交互作用	2
1.3.1 简单效应	3
第二章 Bookdown 的基本用法	5
2.1 使用 bookdown 创作的流程	5
2.1.1 初始化	5
2.1.2 编辑内容	5
2.1.3 预览内容	5
2.1.4 发布	5
2.2 bookdown 对文档的组织	6
2.3 bookdown 语法	6
第三章 常用功能的实现	9
3.1 如何插入参考文献	9
3.2 如何将参考文献格式调整为国标格式	9

表格

插图

1.1 主效应、交互作用示意图	3
---------------------------	---

前言

在这本小书中，我将介绍控制实验的研究设计和数据分析。多因素实验设计是当前研究发展的趋势，它可在一定程度上克服早期实验室和现场研究的局限性，使实验研究更加深入，可探索更加复杂的现象，同时使研究结果更加精确、可靠。

同时，实验设计也是一本技术，它包括实验设计、统计分析和计算机数据处理三方面的知识，缺一不可。

以下是我的 R 进程信息：

```
sessionInfo()
```

```
## R version 3.6.1 (2019-07-05)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Catalina 10.15
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.6/Resources/lib/libRlapack.dylib
##
## locale:
##  [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
##  [1] stats      graphics  grDevices  utils      datasets
##  [6] methods    base
##
## loaded via a namespace (and not attached):
##  [1] compiler_3.6.1  magrittr_1.5    bookdown_0.16
##  [4] tools_3.6.1     htmltools_0.4.0 yaml_2.2.0
##  [7] Rcpp_1.0.3      stringi_1.4.3   rmarkdown_1.18
## [10] knitr_1.26      stringr_1.4.0   xfun_0.11
## [13] digest_0.6.23   rlang_0.4.2     evaluate_0.14
```

杨志宏
于世界之最温暖处

第一章 实验设计概述

实验设计，指的是实施实验处理的一个计划方案以及与计划方案有关的共计划分析。

实验设计包括：

1. 建立与研究假说有关的统计假说；
2. 确定实验中使用的实验处理（自变量）和必须控制的多余条件（无关变量）；
3. 确定实验中需要的实验单元（被试）的数量以及被试抽样的总体；
4. 确定将实验条件分配给被试的方法；
5. 确定实验中每个被试要记载的测量（因变量）和使用的统计分析。

做实验研究，需要具备两方面的知识：

1. 有关研究课题的理论背景、研究的基本假设与预期，要保证开展的研究在特定的领域中有继承、有发展，有一定的科学价值。
2. 有关实验研究的研究设计、统计学知识和软件操作知识。

1.1 实验设计发展趋势

自 17 世纪经验科学发展以来，实验研究由单变量研究，向多变量研究发展。多变量研究带来的问题是，多种因素之间本来是相互关联的，要同时研究多个因素的影响，难度加大。

自 20 世纪初以来，实验研究有两条发展路线：

1.1.1 实验室研究

实验室研究通过严密的实验控制，改变和操作极少数自变量，控制其它无关变量，以保证实验结果的可靠性，它适合研究比较单纯的问题，研究结果距离现实世界太远，研究的外部效度低。

1.1.2 社会研究

利用统计控制，通过合理取样大量被试，在较少或没有实验控制的情况下，探讨变量之间的关系，它的基本思想是，大样本测试可以通过统计规律保证结果的可靠性。

多因素实验设计与多元统计相结合，是近几十年的发展趋势，目的在于克服传统研究的不足，使研究情景接近现实，同时最大限度地保证结果的精度。

1.1.3 多因素实验设计的特点

1. 实验中通过实验设计方案的周密安排及相应的统计方法，尽可能地控制无关变量。
2. 实验中尽可能容纳较多的自变量

1.2 实验设计中的基本概念

1.2.1 因素与因素实验设计

因素 (factor) 指研究者在实验中感兴趣的一个变量，研究者通过操纵、改变它，来估计它对因变量的影响，这个变量也叫做自变量。

因素的水平 (level)，是实验中所操纵的变量的每个特定的值。如灯光强度、性别、年龄。在方差分析相结合的实验设计中，因素水平的数量不能过多。

因素设计，通常指多于一个因素的实验设计 (factorial experimental design)，如一个含有三个因素、每个因素有两个水平的实验设计，称为 $2 \times 2 \times 2$ 三因素设计。

1.2.2 处理与处理水平的结合

处理 (treatment) 与处理水平的结合 (treatment combinations)，是指实验中一个特定的、独特的实验条件。

被试接受 A2B2 实验处理，即在呈现速度为低速 (A2) 并且英语单词为高频单词 (B2) 的实验条件下，被试接受测试。

1.3 主效应与交互作用

主效应 (main effects)，是指一个因素的不同水平引起的变异。

在单因素实验中，由自变量的不同水平的数据计算的方差，就是这个变量的处理效应，即主效应。

在多因素实验中，计算一个因素的主效应时，应忽略实验中其他因素的不同水平的差异。

当一个因素的水平在另一个因素的不同水平上变化趋势不一致时，我们称这两个因素之间存在交互作用 (interaction)。如果变化趋势一致，表明这两个因素是相互独立的。

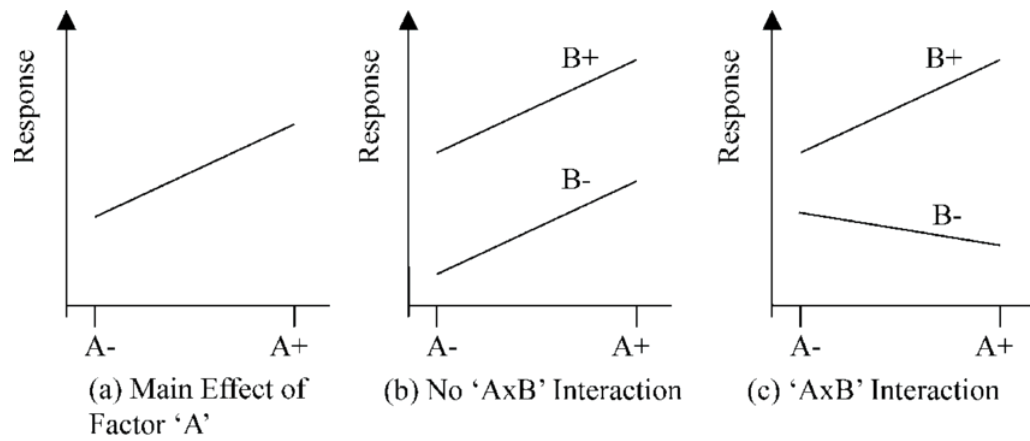


图 1.1: 主效应、交互作用示意图

1.3.1 简单效应

简单效应 (simple effects) , 指的是一个因素的水平在另一个因素的某个水平上的变异。

A 因素的两个水平在 B1 水平的方差, 叫 A 在 B1 水平的简单效应。

第二章 Bookdown 的基本用法

从本质上讲，bookdown 是基于 Pandoc、markdown、R、Rmarkdown、gitbook、tex、git 等一系列开源项目上的综合性工具，这套工具大大降低了用户自己整合上述多个工具的成本，以更加高效地方式完成内容创作、发布和维护。

2.1 使用 bookdown 创作的流程

2.1.1 初始化

使用下载的模板或者 bookdown 自带的模板，创建 bookdown 项目。建议使用前者，因为更加适合中文内容的排版。

当然，当用户熟悉了 bookdown 之后，还可以在这些模板的基础上进行必要的配置，使之更加符合用户的需求。

2.1.2 编辑内容

在合适的编辑器中（建议先使用 RStudio，后期熟悉后，可使用 VS Code 之类的功能更为复杂的通用性编辑器），按照 markdown 以及 Rmarkdown 的语法规则，进行内容的创作和编辑。

2.1.3 预览内容

选择 RStudio 中的“Build Book”功能，生成合适格式。还可以在控制台中键入如下命令，开启实时预览：

```
bookdown::serve_book()
```

2.1.4 发布

利用 GitHub 或 Gitee 的 page 服务，可将生成的静态页面发布到网络。

2.2 bookdown 对文档的组织

和一篇文章相比，一本书包含多个章节。在 bookdown 中，一个章节对应一个后缀名为 .Rmd 的 R Markdown 文件，每个 R Markdown 文件（除了首页）都必须以一级标题开头。例如：

```
# bookdown 对书本文档的组织
```

默认情况下，bookdown 按照文件名的顺序，从前到后合并在一起，并渲染成 gitbook 格式或者 pdf、epub 或者 word 文件。

如果存在 index.Rmd 文件，则该文件会被渲染成 index.html。以下划线开头的文件，会被 bookdown 忽略。

bookdown 还提供了自定义章节顺序的机制。在配置文件 _bookdown.yml 中，通过 rmd_files 字段实现，例如：

```
rmd_files: ["index.Rmd", "abstract.Rmd", "intro.Rmd"]
```

还可以分别为不同格式指定包含的内容：

```
rmd_files:
  html: ["index.Rmd", "abstract.Rmd", "intro.Rmd"]
  latex: ["abstract.Rmd", "intro.Rmd"]
```

必须要指出的是，每个章节的一级标题后，需要在后面加上标签，如：

```
# Bookdown 的基本用法 {#basic}
```

2.3 bookdown 语法

首先，bookdown 是基于 markdown 和 pandoc 的工具，因此，Markdown 的原生语法¹肯定是可以使用的。在此不再赘述。

其次，Pandoc 在 markdown 语法的基础上，提供了一些增强²，bookdown 也是支持的。

最后，bookdown 是在 R bookdown 包³的基础上进行扩展，故而也支持 R Markdown 语法。如：

```
dim(iris)
```

```
## [1] 150 5
```

¹<https://daringfireball.net/projects/markdown/syntax>

²<https://pandoc.org/MANUAL.html#pandocs-markdown>

³<https://rmarkdown.rstudio.com/index.html>

再如，下面的代码将列出 R Markdown 支持的所有语言：

```
names(knitr::knit_engines$get())
```

```
## [1] "awk"          "bash"          "coffee"        "gawk"          "groovy"
## [6] "haskell"      "lein"          "mysql"          "node"          "octave"
## [11] "perl"         "psql"          "Rscript"        "ruby"          "sas"
## [16] "scala"        "sed"           "sh"             "stata"         "zsh"
## [21] "highlight"    "Rcpp"          "tikz"           "dot"           "c"
## [26] "fortran"      "fortran95"     "asy"            "cat"           "asis"
## [31] "stan"         "block"         "block2"         "js"            "css"
## [36] "sql"          "go"            "python"         "julia"         "sass"
## [41] "scss"         "theorem"       "lemma"          "corollary"     "proposition"
## [46] "conjecture"   "definition"    "example"        "exercise"      "proof"
## [51] "remark"       "solution"
```


第三章 常用功能的实现

3.1 如何插入参考文献

1. 插入参考文献时，需要创建.bib 格式的参考文献数据库。
2. 在合适的地方插入引用 (Xie, 2015)，语法 (毛向樱, 2018) 如下：

引用 [Xie2015]

3. 单独创建一个显示全部参考文献的.Rmd 文件，内容如下：

按照上述步骤，将在章节末尾和全书末尾生成参考文献。

3.2 如何将参考文献格式调整为国标格式

在首页文件 index.Rmd 中，用户可以通过 `biblio-style` 指定参考文献的格式，例如：

```
----  
bibliography: ["one.bib", "another.bib", "yet-another.bib"]  
biblio-style: "GBT7714-2005"  
link-citations: true  
----
```

与此同时，还需要将上述文件复制到项目目录中，不过上述设定，仅对 PDF 和 word 输出时起作用，对于 gitbook 格式，则没有效果。

参考文献

Xie Y. Dynamic documents with R and knitr[M/OL]. 2nd ed. Boca Raton, Florida: Chapman and Hall/CRC, 2015. <http://yihui.name/knitr/>

毛向樱. 网络语言暴力及国外治理经验借鉴[J]. 武汉冶金管理干部学院学报, 2018, 28(1): 25–27.

索引

level, 2

主效应, 2

交互作用, 2

因素, 2

处理, 2

处理水平的结合, 2

实验室研究, 1

实验设计, 1

简单效应, 3