

数据模拟代码

温秀娟 潘晚珂 金海洋

2022-10-17

```
# 首先加载需要使用的 R 包
if (!require(papaja)) {library(papaja)}

## 载入需要的程辑包: papaja

## 载入需要的程辑包: tinylabels

if (!require(tidyverse)) {library(tidyverse)}

## 载入需要的程辑包: tidyverse

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.1
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

if (!require(faux)) {library(faux)}

## 载入需要的程辑包: faux
##
## *****
## Welcome to faux. For support and examples visit:
```

```
## https://debruine.github.io/faux/
## - Get and set global package options with: faux_options()
## *****
##
## 载入程辑包: 'faux'
##
## The following object is masked from 'package:purrr':
##
##      %||%
set.seed(2022)
```

在模拟实验中，40 名抑郁症患者和 40 名健康对照组被试观看 30 张积极和 30 张中性图片，期间我们采集了他们的脑电数据。因变量是晚期正电位（late positive potentials, LPP）的波幅。简单来说，这是一个 2 (组别 **group**: 抑郁症组 **depression**、对照组 **control**) \times 2 (图片类型 **type**: 积极 **positive**、中性 **neutral**) 的混合实验设计，其中组别为被试间因素，图片类型为被试内因素。该假想实验的数据是使用 DeBruine (2021) 的 **faux** 工具包生成，下面是模拟这个实验所预设的参数。

```
subj_n <- 80    # 总被试量: 抑郁患者 30 人, 健康对照组被试 30 人
trial_n <- 30   # 每张图片呈现的次数

# 固定效应
b0 <- 0.5       # 截距 (所有条件的均值)
b1 <- 6.5       # 图片类型的固定效应 (主效应)
b2 <- 0.1       # 组别的固定效应 (主效应)
b3 <- 0.1       # 图片类型与组别的交互作用

# 随机效应
u0s <- 2        # 被试的随机截距
u1s <- 2        # 被试的随机斜率 (图片类型)

# 误差项
sigma <- 2
```

之后根据假定的实验设计和参数来生成模拟数据：

```
# 生成假定实验的条件的数据矩阵
df_simu <- add_random(subj = subj_n) %>%
  # 添加被试的组别信息（被试间）
  add_between("subj", group = c("depression", "control")) %>%
  # 添加图片类型的信息（被试内）
  add_within("subj", type = c("netural", "positive")) %>%
  # 每种图片呈现 30 次
  add_random(trial = trial_n) %>%
  # 图片类型的编码：中性 = -0.5；正性 = 0.5
  add_contrast("type", "anova", colnames = "type_code") %>%
  # 被试组别的编码：抑郁组 = -0.5；控制组 = 0.5
  add_contrast("group", "anova", colnames = "group_code") %>%
  # 添加基于被试的随机截距和斜率（图片类型）
  add_ranef("subj", u0s = u0s, u1s = u1s, .cors=0.5) %>%
  # 添加观察值的误差项
  add_ranef(sigma = sigma) %>%
  # 最后根据设置的固定效应和随机效应参数值，生成因变量。
  mutate(LPP = (b0+u0s) + # 截距
             (b1+u1s) * type_code + # 图片材料的斜率
             b2 * group_code + # 组别的斜率
             b3 * type_code * group_code + # 交互作用
             sigma) # 误差项

df_simu <- df_simu %>%
  select(subj, group, type, LPP) # 去除冗余的信息

# 查看生成的数据
head(df_simu, 10)

## # A tibble: 10 x 4
##   subj group    type    LPP
##   <chr> <fct>    <fct>  <dbl>
## 1 s01  depression netural -1.64
```

```
## 2 s01 depression netural -4.00
## 3 s01 depression netural -1.53
## 4 s01 depression netural -3.22
## 5 s01 depression netural -4.34
## 6 s01 depression netural 1.49
## 7 s01 depression netural -3.46
## 8 s01 depression netural 2.93
## 9 s01 depression netural -3.27
## 10 s01 depression netural -1.44
```