

应用示例

温秀娟 潘晚珂 金海洋

2022-10-17

应用示例

我们借助一个假想的心理学实验展示如何使用贝叶斯混合效应模型分析数据。该假想实验的研究目的是探究抑郁症人群加工不同情绪刺激的神经基础。在该实验中，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。我们所感兴趣的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，具体代码如下：

```
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

# 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: 抑郁症组 == -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

head(df_simu, 10) # 查看生成的数据矩阵

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

```

数据预处理

接下来我们使用 R 的 `brms` 工具包建构和拟合贝叶斯线性混合效应模型（Bürkner, 2017; Ladislav et al., 2019; Vasishth et al., 2018）。`brms` 工具包凭借其具有与 `lme4` 相似的语法（Bates, Mächler, et al., 2015）和强大的扩展功能等优势受到许多研究者的青睐，是目前比较流行的贝叶斯数据分析工具之一。在使用 `brms` 工具包分析数据之前，我们需要对模拟的数据集进行一些预处理。首先，大部分流行的统计软件（例如 SPSS）通常分析的是“宽”（wide）数据，即每一行是一个被试的数据。相比之下，`brms` 以及其他一些 R 语言的工具包通常处理的是“长”（long）数据，即每一行是一个试次的的数据。因此，我们需要将数据集 `df_simu` 整理为“长数据”格式，如下所示：

```
head(df_simu, 10) # 查看数据结构

## # 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
```

此外，我们还需要对数据集 `df_simu` 中的称名（自）变量（nominal factors）进行编码（contrast coding）。这里我们使用逐次差分对比编码（即 R 中的 `MASS::contr.sdif()`）对组别（`group`）和图片类型（`type`）这两个自变量进行编码。通过此设定，贝叶斯模型的输出结果刚好对应大部分研究者熟悉的主效应和交互作用。

```
# 设定逐次差分对比编码 (即 -0.5 和 0.5)
contrasts(df_simu$group) <- MASS::contr.sdif(2)
contrasts(df_simu$type) <- MASS::contr.sdif(2)
```

只有固定效应的模型

在建立贝叶斯线性混合效应模型时，研究者需要考虑如何设定固定效应和随机效应。我们首先建立最简单的模型，即一个仅包含固定效应的模型 (`model_1`)。具体来说，我们将组别 (`group`) 和图片类型 (`type`) 两个自变量以及他们的交互作用作为固定效应。该模型假定所有被试的平均波幅都是相同的，且图片类型对于每个被试的影响也是相同的。

```
# 建立只有固定效应的模型
model_1 <- brm(LPP ~ group * type,
  data = df_simu,
  chains = 4,
  cores = 4,
  warmup = 1000,
  iter = 5000,
  seed = 2022,
  file = "brm_fit/model_1.Rds"
)
```

在上述代码中，LPP 是因变量，即被试观看不同图片时 LPP 的波幅，`group * type` 表示组别 `group` 和图片类型 `type` 两个自变量的主效应以及他们的交互作用。`df_simu` 为数据集。其他参数是模型拟合时所基于的 MCMC 采样算法的相关参数。为了保证后验分布的有效性，MCMC 通常会同时运行多个独立的采样过程，每个独立的采样过程称为一条链 (chain)。只有当各个独立链的抽样“融合”在一起时 (详见下文)，MCMC 获得的参数后验分布才更可能是有效的 (McElreath, 2020)。在 `model_1` 中，我们设置了 4 条链 (默认为 4)。与之相对应的是 `cores` 参数，它可以设定运行 MCMC 时使用的计算机中央处理器核心 (Central Processing Unit cores) 的个数。`cores=4` (默认为 1) 表示使用 4 个中央处理器核心同时分别各运

行 1 条马尔可夫链（共 4 条链），借此提高数据分析的速度。`iter` 参数（iteration 的简写）指的是每条链的迭代次数，即对每个参数进行多少次的抽样（sample）。`warmup` 参数（或称为 burn in）设定了在每条链采样的开始阶段用于“校准”的迭代次数。因此，每条链最后保留的采样数量为 `iter` 和 `warmup` 的差值。例如，在 `model_1` 中每一条链会对每个参数抽取 4000 个样本（即 5000-1000），4 条链一共会对每个参数抽取 16000 个样本。通过 `summary()` 函数可以展示模型拟合的相关参数和结果：

```
summary(model_1) # 可以查看参数情况
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LPP ~ group * type
## Data: df_simu (Number of observations: 4800)
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;
## total post-warmup draws = 16000
##
## Population-Level Effects:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept          0.68      0.04    0.59    0.76 1.00   19010   12439
## group2M1            0.28      0.09    0.11    0.45 1.00   21511   12581
## type2M1             6.50      0.09    6.33    6.67 1.00   22435   12931
## group2M1:type2M1    0.02      0.17   -0.32    0.35 1.00   20521   11983
##
## Family Specific Parameters:
##           Estimate Est.Error 1-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma          3.03      0.03    2.97    3.09 1.00   20821   12632
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Population-Level Effects 部分为固定效应的结果。其中，Intercept 是模型的固定截距，表示所有被试的平均 LPP 波幅。group2M1 表示在平均图片类型的效应之后，健康对照组的 LPP 波幅减去抑郁症组 LPP 波幅的差

值,即组别的主效应。与此类似,`type2M1` 是在平均组别的效应之后,积极图片诱发的 LPP 波幅减去中性图片诱发的 LPP 波幅的差值,即图片类型的主效应。`group2M1:type2M1` 是组别和图片类型交互作用,即 $(LPP_{(control,positive)} - LPP_{(control,neutral)}) - (LPP_{(depression,positive)} - LPP_{(depression,neutral)})$ 。

在输出的结果中, **Estimate** 表示参数的后验分布的均值, **Est.Error** 表示后验分布的标准差 (standard deviation)。l-95% CI 和 u-95% CI 表示参数的 95% 可信区间的两个边界,例如 `group2M1` 的 95% 的可信区间是 [0.11, 0.46]。Rhat 表示不同链条采样的融合程度。Rhat 越接近于 1,表示不同链条采样融合的更好,即模型的收敛情况更好。如果 Rhat 大于 1.01 则说明模型的收敛情况并不理想。此时,研究者不应该使用获得的后验分布进行进一步的分析,而应该检查模型设定是否合理等。ESS 为有效样本量 (effective sample size),是对 MCMC 算法采样到的有效且独立的样本数量的估计。ESS 越高表明能用来有效估计后验分布的样本数量越大。通常建议 Bulk-ESS 是 MCMC 链数量的 100 倍及其以上。例如,当运行 4 条 MCMC 链时,所需的 Bulk-ESS 至少为 400。Tail-ESS 表示后验分布两端的有效样本量。通常情况下,过小的 Bulk-ESS 和 Tail-ESS 会伴随着过大的 Rhat。合理的设定模型和先验分布(见 3.5 部分)一般可以解决大部分 Rhat 过大的情况。此外,研究者还可以通过可视化的方式检查各条链的融合情况:

图 1 描述了 `model_1` 中 MCMC 不同链抽样的轨迹图,反映了不同链的融合情况。在图 1 中,不同链参数的抽样很好地“融合”在一起,说明其对应的参数后验分布是相对可靠的。由于融合的链看起来像一条毛毛虫,因此图 1 的轨迹图又称“毛毛虫”图。如果一条链的抽样与其他链的抽样存在很大的差异(通常也会伴随着较大的 Rhat),那这些后验抽样也很可能无法有效地表征参数的后验分布。此时,研究者需要重新考虑模型的合理性。

添加随机效应

如前所述,仅包含固定效应的模型(`model_1`)假定所有被试的平均脑电波幅都是相同的,即所有被试共享同一个截距值。但这种情况在现实中不太可能发生。更可能的情况是,每个被试的平均脑电波幅都各不相同。因此,我们需要在模型中增加随机截距来考虑被试在平均脑电波幅上的个体差异。具体来说,我们需要在 `model_1` 的基础上增加基于被试 (by-subject) 的随

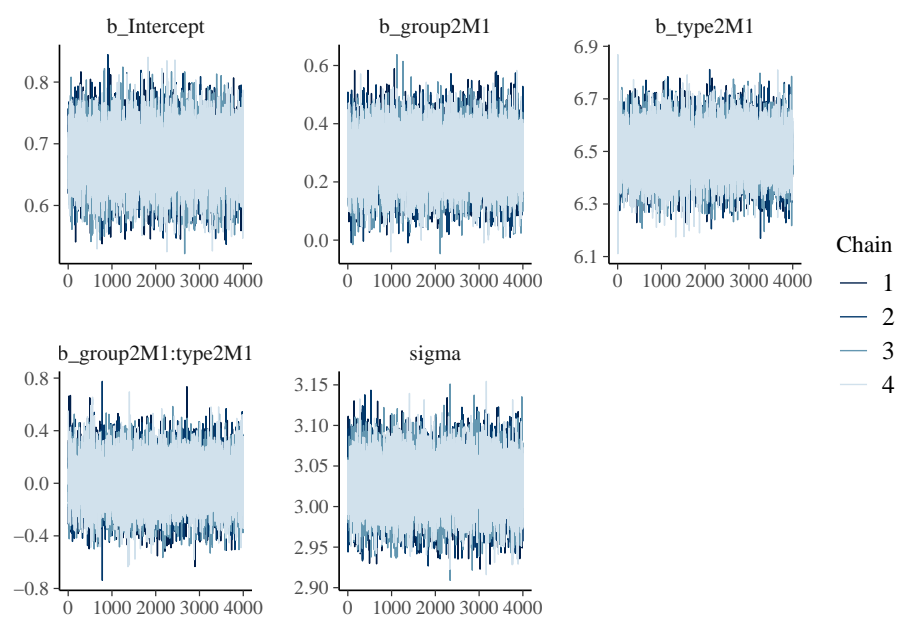


图 1: model_1 的 MCMC 各条链抽样的轨迹图（又叫“毛毛虫”图）。x 轴表示迭代数，y 轴表示参数值。不同颜色的线表示不同的链。

机截距（对应公式 2 中的 $S_{(0,i)}$ ），即 $(1 \mid \text{subj})$ 。其中 `subj` 为被试编号，1 表示随机截距：

```
model_2 <- brm(LPP ~ group * type + (1 | subj),
  data = df_simu,
  cores = 4,
  chains = 4,
  warmup = 1000,
  iter = 5000,
  seed = 2022,
  file = "brm_fit/model_2.Rds"
)
```

```
summary(model_2) # 查看参数情况
```

```
## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LPP ~ group * type + (1 | subj)
## Data: df_simu (Number of observations: 4800)
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;
## total post-warmup draws = 16000
##
## Group-Level Effects:
## ~subj (Number of levels: 80)
##
```

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	2.15	0.17	1.84	2.52	1.01	934	1946

```
##
## Population-Level Effects:
##
```

	Estimate	Est.Error	1-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.67	0.24	0.19	1.15	1.00	502	1018
group2M1	0.26	0.49	-0.72	1.20	1.01	450	1095
type2M1	6.50	0.06	6.38	6.62	1.00	13087	11075
group2M1:type2M1	0.02	0.13	-0.23	0.26	1.00	13188	10601

```
##
## Family Specific Parameters:
```



```
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma      2.18      0.02      2.14      2.23 1.00      15492      11004
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

Population-Level Effects 为 model_2 固定效应的结果，其相关参数的含义与 model_1 中的相同（详见 3.2 部分）。Group-Level Effects 为随机效应结构的结果，其中 `sd(Intercept)` 为随机截距的标准差，反映了 LPP 平均波幅在不同被试间的离散程度。

添加随机斜率

虽然 model_2 考虑了不同被试平均 LPP 波幅的个体差异，但它仍然假定所有被试受到实验处理（即图片类型）的影响是相同的。然而在实际情况中，每个被试受到图片类型的影响更可能是不同的。因此，我们需要设置随机斜率考虑图片类型对不同被试的不同影响。具体来说，我们可以在 model_2 模型的基础之上加入随机斜率（对应公式 3 中的 $S_{(\delta,i)}$ ），即 `(type | subj)`，其中 `subj` 为被试编号，`type` 为图片类型的随机斜率。与此同时，我们也在模型中加入了随机截距与随机斜率间的相关：

```
model_3 <- brm(LPP ~ group * type + (1 + type | subj),
  data = df_simu,
  cores = 4,
  chains = 4,
  warmup = 1000,
  iter = 5000,
  seed = 2022,
  file = "brm_fit/model_3.Rds"
)
```

```
summary(model_3) # 查看参数情况
```

```
## Family: gaussian
```

```
## Links: mu = identity; sigma = identity
## Formula: LPP ~ group * type + (1 + type | subj)
## Data: df_simu (Number of observations: 4800)
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;
## total post-warmup draws = 16000
##
## Group-Level Effects:
## ~subj (Number of levels: 80)
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sd(Intercept)	2.17	0.18	1.85	2.56	1.00	1178	
sd(type2M1)	1.81	0.16	1.53	2.16	1.00	2998	
cor(Intercept,type2M1)	0.32	0.11	0.10	0.52	1.00	2103	
sd(Intercept)	2349						
sd(type2M1)	6078						
cor(Intercept,type2M1)	3937						

```
## Population-Level Effects:
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
Intercept	0.67	0.25	0.18	1.17	1.00	596	1057
group2M1	0.30	0.49	-0.67	1.25	1.00	630	1315
type2M1	6.49	0.20	6.09	6.89	1.00	1482	2921
group2M1:type2M1	0.03	0.43	-0.81	0.86	1.00	1601	3638

```
## Family Specific Parameters:
##
```

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.99	0.02	1.95	2.03	1.00	23572	12071

```
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

与 model_2 结果相比,model_3 中的 Group-Level Effects 部分额外展示了基于被试的随机斜率和他与随机截距相关性的后验分布情况。其中,

`sd(type2M1)` 为基于被试的图片类型随机斜率的标准差，反映了不同被试 LPP 波幅受到图片类型不同影响的离散程度。`cor(Intercept, type2M1)` 为随机截距和随机斜率的相关，该相关结果表明两者之间存在一定的正相关，即平均 LPP 波幅更大的被试，其对积极图片的 LPP 波幅也比中性图片更大。或许有研究者会考虑是否也应该在模型中增加组别的随机斜率。但由于在该实验中，每名被试只参加了实验组或控制组，所以我们无法估计组别的随机斜率。一般来讲，研究者应该为感兴趣的所有被试内因素设定随机效应，而不需要为被试间因素设定随机效应。至此，我们完成了对该模拟数据的模型构建，即根据实验设计设定了所有合理的固定效应和随机效应结构。在该示例中，我们从仅包含固定效应的线性模型出发，在不同的模型中依次展示了如何添加随机截距和随机斜率。但在实际的贝叶斯数据分析中，研究者只需要根据实验设计直接设定最合理的模型，而无需在不同的模型中依次添加不同的随机效应。

先验的设置

贝叶斯分析的一大优势是可以通过设定先验分布在数据分析中整合研究者的先验知识。由于我们在先前建立的模型（例如 `model_3`）中没有自定义先验分布，因此它们使用的是 brms 的默认先验分布。这些先验分布通常是无信息先验（noninformative prior），即相对实验数据来讲，基本不能为参数估计提供有效信息的先验 (Seaman et al., 2012)。使用这些默认的无信息先验不仅不能发挥贝叶斯数据分析的最大作用，甚至可能会对数据分析的结果产生负面的影响 (Seaman et al., 2012)。例如，贝叶斯因子很容易受到先验分布的影响，不合理的先验会导致有偏差的贝叶斯因子。因此，为贝叶斯分析设定合理的先验至关重要。

先验分布是研究者以概率分布的形式描述自己关于特定研究问题的先验知识。它赋予了模型参数在观测到数据之前的相对合理性。这些先验信息与观测数据的信息相互整合成为模型参数的后验分布。作为设定先验分布的第一步，研究者需要知道自己可以对一个模型的哪些参数设置先验。这里可以借助 `get_prior()` 函数得到相应的信息：

```
get_prior(LPP ~ group * type + (1 + type | subj),  
          data = df_simu
```

```

)

##          prior      class      coef group resp dpar nlpar lb ub
##          (flat)         b
##          (flat)         b      group2M1
##          (flat)         b group2M1:type2M1
##          (flat)         b      type2M1
##          lkj(1)      cor
##          lkj(1)      cor          subj
## student_t(3, 0.5, 5) Intercept
##   student_t(3, 0, 5)      sd                      0
##   student_t(3, 0, 5)      sd          subj          0
##   student_t(3, 0, 5)      sd      Intercept  subj          0
##   student_t(3, 0, 5)      sd      type2M1  subj          0
##   student_t(3, 0, 5)      sigma          0
##      source
##      default
## (vectorized)
## (vectorized)
## (vectorized)
##      default
## (vectorized)
##      default
##      default
## (vectorized)
## (vectorized)
## (vectorized)
##      default

```

如结果所示，我们可以为 `model_3` 模型的以下参数设定先验：固定截距（`class` 为 `Intercept`），固定斜率（`class` 为 `b` 的参数，包含 `coef` 为 `group2M1`, `group2M1:type2M1` 和 `type2M1`），基于被试（`group` 为 `subj`）的随机效应的标准差（`class` 为 `sd`，包含随机截距 `coef` 为 `Intercept` 和随机斜率 `coef` 为 `type2M1`），基于被试的随机效应间的相关（`class` 为

cor), 以及残差 (class 为 sigma)。由于 sd 和 sigma 不能为负数, 因此这些参数的最小值为 0 (即列 lb, 为 low bound 的缩写)。在实际数据分析中, 研究者一般可以根据以往研究的结果为当前研究的模型设定先验分布。如果可以从先前类似的研究获得参数的后验概率分布, 研究者可以将这些概率分布作为当前研究模型的先验分布。即使无法获得参数准确的概率分布, 研究者仍然可以根据先前研究结果以及自己已有的知识为模型设定先验分布。此时, 研究者可以通过先验预测检验 (prior predictive check) 来测试这些先验的合理性, 并进行相应的调整, 进而为模型确定合理的先验。

假设我们无法从先前研究中获取参数的概率分布作为模型的先验。此时, 我们也可以凭借已有的经验为模型设定先验。比如, 我们首先尝试设置无信息或弱信息的先验分布:

```
## 定义先验分布
prior_01 <- c(
  prior(normal(0, 10), class = Intercept),
  prior(normal(0, 10), class = b),
  prior(normal(0, 5), class = sigma),
  prior(normal(0, 5), class = sd),
  prior(lkj(2), class = cor)
)
```

在上述代码中, 我们把固定截距和固定斜率分别设定为一个均值为 0, 标准差为 10 的正态分布 (在代码中表示为 `normal(0, 10)`)。对于 `prior(normal(0, 10), class = Intercept)`, 我们可以简单地把它理解为, 我们设定所有被试的平均值脑电波幅大致有 95% 的概率落在 -20 和 20 之间 (即两个标准差之内)。对于平均脑电波幅来讲, 这是一个相对比较大的区间。因此, 它对应的先验分布是一个无信息或弱信息的先验分布。类似地, 固定斜率, 包含组别和图片类型的主效应以及他们的交互作用, 也大致落在相同的区间。尽管固定斜率的先验是无信息或弱信息先验, 但这些先验 (即 `normal(0, 10)`) 仍然比 `brms` 中的默认先验 (它对应的区间大约是负无穷到正无穷) 提供了更多的信息。此外需要注意的是, 尽管我们使用正态分布作为残差 `sigma` 和随机效应的标准差 `sd` 设置先验, 但由于这些参数不能小于 0, 因此他们使用的实际先验分布是截断正态分布 (truncated normal distribution)。最后, 为随机效应间的相关性设置的是

弱信息先验 `prior(lkj(2), class = cor)`，其中 `lkj` 为 LKJ 相关性矩阵 (LKJ-Correlation prior matrix) (Lewandowski et al., 2009)。为了检验先验分布的合理性，我们可以使用先验预测检验审查只使用先验信息时因变量（即 LPP 的波幅）的情况。进行先验预测检验需要对先验分布进行采样（sample），即在模型中设定先验分布（`prior=prior_01`）的基础上加入 `sample_prior = "only"`。其余参数的设定和我们在分析数据时设定相同（例如 `model_3`），具体代码如下：

```
model_ppc_01 <- brm(  
  LPP ~ group * type + (1 + type | subj),  
  data = df_simu,  
  cores = 4,  
  chains = 4,  
  warmup = 1000,  
  iter = 5000,  
  sample_prior = "only",  
  prior = prior_01,  
  seed = 2022  
)
```

```
## Compiling Stan program...
```

```
## Start sampling
```

通过上述代码，我们可以获得 16,000 组 $((5000-1000) * 4)$ 来自先验分布的抽样。其中，每一组都包含模型的所有参数，即固定截距，固定斜率，随机效应等。每一组先验参数都可以生成一套新的与数据集（`df_simu`）结构相同的模拟观测数据（即因变量）。我们可以通过考查这些模拟观测数据的合理性（比如检查每一套模拟数据的极值）去判断先验的设定是否合适。例如在一个以反应时为因变量的 stroop 实验中，如果那些由先验分布生成的模拟数据集中大部分数据集的最大值都超过预期（比如 10 秒），这很可能说明模型某些参数先验的设定不太合理（比如截距或标准差过大）。在当前模拟实验中，尽管我们无法根据先前研究获取精确的先验，但是可以利用现有的知识，即试次水平的脑电事件相关电位的波幅通常在一定的范围之内，来确定先验。因此，我们计算 16,000 套由先验分布生成的模拟观测数据各

自的最小值 (min)、均值 (mean) 和最大值 (max)，并通过它们的范围是否符合我们的先验知识来判断先验分布的合理性。上述分析可以通过 brms 提供的 `pp_check()` 函数实现。以下代码展示了 `prior=prior_01` 的先验预测的结果：

```
if (!require(cowplot)) {library(cowplot)}

## 载入需要的程辑包：cowplot

P1=pp_check(model_ppc_01, type = "stat", stat = "min", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.

P2=pp_check(model_ppc_01, type = "stat", stat = "max", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.

P3=pp_check(model_ppc_01, type = "stat", stat = "mean", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.

plot_grid(P1,P2,P3,nrow = 1,align = "h")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

结果显示，基于 `prior_01` 生成的每一套模拟观测数据的最小值大部分落在 $[-75, 0]$ ，均值大部分落在 $[-25, 25]$ ，而最大值大部分落在 $[0, 80]$ 。根据先验知识，试次水平的脑电波幅通常很少会低于 -50 (V) 或高于 50 (V) (伪迹除外)。因此，我们可以适当地调整参数的先验分布以符合我们的先验知识，即减小参数先验分布的不确定性。在接下来的尝试中，我们将固定截距和固定斜率的先验设置成均值为 0，标准差为 5 的正态分布，并进行与之前相同的先验预测检验：

```
## 定义先验分布
prior_02 <- c(
  prior(normal(0, 5), class = Intercept),
```

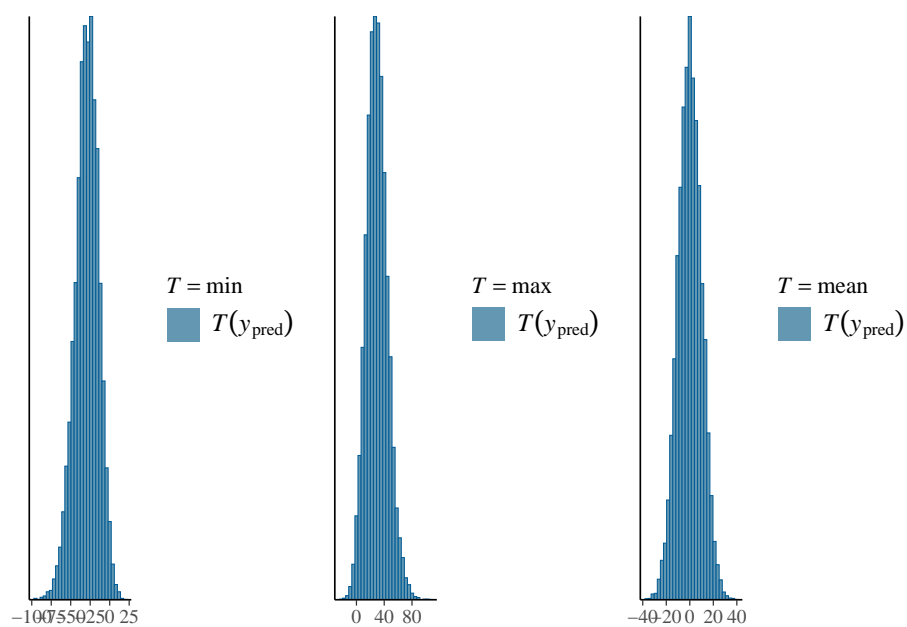


图 2: 基于先验 prior_01 生成的 16, 000 套模拟观测数据 ($y(\text{rep})$) 的最小值、均值和最大值的分布。


```

    prior(normal(0, 5), class = b),
    prior(normal(0, 5), class = sigma),
    prior(normal(0, 5), class = sd),
    prior(lkj(2), class = cor)
  )
model_ppc_02 <- brm(LPP ~ group * type + (1 + type | subj),
  data = df_simu,
  cores = 4,
  chains = 4,
  warmup = 1000,
  iter = 5000,
  sample_prior = "only",
  prior = prior_02,
  seed = 2022
)

## Compiling Stan program...

## Start sampling
if (!require(cowplot)) {library(cowplot)}
P4=pp_check(model_ppc_02, type = "stat", stat = "min", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.
P5=pp_check(model_ppc_02, type = "stat", stat = "max", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.
P6=pp_check(model_ppc_02, type = "stat", stat = "mean", prefix = "ppd")

## Using all posterior draws for ppc type 'stat' by default.
plot_grid(P4,P5,P6,nrow = 1,align = "h")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

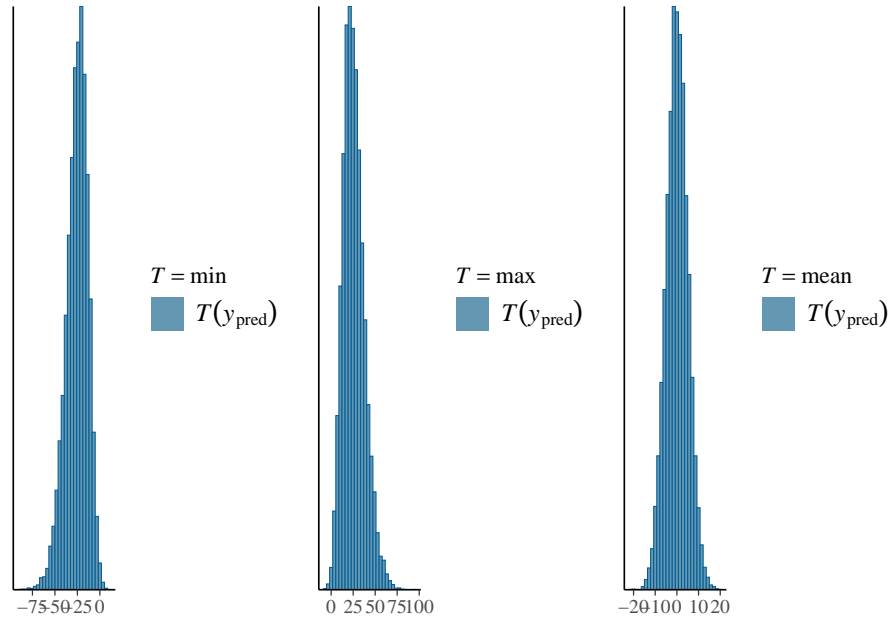


图 3: 基于先验 prior_02 生成的每一套模拟观测数据 ($y(\text{rep})$) 的最小值、均值和最大值的分布

经过同样的步骤, 我们可以获得基于先验 prior_02 生成的 16,000 套模拟观测数据 ($y(\text{pred})$) 的最小值、均值和最大值的分布 (图 3)。结果显示模拟观测数据的最小值大部分落在 $[-60, 0]$, 均值大部分落在了 $[-10, 10]$, 最大值大部分落在 $[0, 60]$ 。与 prior_01 的先验预测结果 (图 2) 相比, prior_02 预测结果更多落在了 $[-50, 50]$ 内, 能够相对更合理的描述我们的先验知识。

确定合理的先验分布通常需要重复多次类似的过程, 且并不是先验分布的信息越强就越好, 它需要合理反映研究者的先验知识。在确定合理的先验分布后, 我们可以利用该先验进行模型拟合, 即在 model_ppc_02 的基础上删除 sample_prior = "only" 或将其更新为 sample_prior = "yes"。我们将新的模型命名为 model_4:

```
model_4 <- brm(
  LPP ~ group * type + (1 + type | subj),
  data = df_simu,
```

```

cores = 4,
chains = 4,
warmup = 1000,
iter = 5000,
sample_prior = "yes",
prior = prior_02,
seed = 2022,
file = "brm_fit/model_4.Rds"
)

```

```
summary(model_4)
```

```

## Family: gaussian
## Links: mu = identity; sigma = identity
## Formula: LPP ~ group * type + (1 + type | subj)
## Data: df_simu (Number of observations: 4800)
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;
## total post-warmup draws = 16000
##
## Group-Level Effects:
## ~subj (Number of levels: 80)
##
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
## sd(Intercept)      2.17      0.18      1.85      2.54 1.01      1166
## sd(type2M1)         1.82      0.16      1.54      2.16 1.00      2255
## cor(Intercept,type2M1) 0.32      0.10      0.10      0.51 1.00      1913
##
##           Tail_ESS
## sd(Intercept)      2101
## sd(type2M1)         4405
## cor(Intercept,type2M1) 4172
##
## Population-Level Effects:
##
##           Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## Intercept      0.68      0.24      0.20      1.15 1.01      457      1094
## group2M1        0.27      0.49     -0.71      1.24 1.00      563      1275

```

```
## type2M1          6.49      0.21      6.07      6.91 1.00      1492      2834
## group2M1:type2M1  0.02      0.43     -0.85      0.85 1.00      1583      2804
##
## Family Specific Parameters:
##      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
## sigma      1.99      0.02      1.95      2.03 1.00      20329      11728
##
## Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
## and Tail_ESS are effective sample size measures, and Rhat is the potential
## scale reduction factor on split chains (at convergence, Rhat = 1).
```

与 `model_3` 相比, `mode_4` 设定了使用先验 `prior_02`, 且在抽取后验样本的同时, 也抽取了先验样本 (即 `sample_prior = "yes"`)。这样的设定可以为我们下一步借助贝叶斯因子进行假设检验奠定基础。

贝叶斯因子

在实际应用中, 研究者通常感兴趣的是实验条件之间是否存在差异, 比如本研究中的组别 (组间因素) 和图片类型 (被试内因素) 是否会对 LPP 产生影响。探索这类问题的一种常用方法是使用贝叶斯因子。贝叶斯因子可以量化基于不同的假设或模型 (例如零假设和备择假设) 观测到当前数据的相对似然 (Rouder et al., 2018), 研究者可以根据其数值大小判断观测到的数据更可能支持哪种假设或模型。一种流行的计算贝叶斯因子的方法是 Dickey Savage density ratio (Heck, 2019; Wagenmakers et al., 2010)。它可以通过 `brms` 提供的 `hypothesis()` 函数实现。我们首先利用贝叶斯混合效应模型的拟合结果 (`model_4`) 检验图片类型是否会对 LPP 产生影响:

```
# 组内比较
result <- hypothesis(model_4, "type2M1=0")
result$hypothesis$Evid.Ratio # 查看贝叶斯因子 BF_01
```

```
## [1] 4.609467e-17
```

其中, `type2M1=0` 表示图片类型的零假设, 即 LPP 在积极图片与中性图片上的差异为 0。`hypothesis$Evid.Ratio` 描述了贝叶斯因子 BF_{01} , 表

示基于零假设观测到当前数据的似然与基于备择假设观测到当前数据的似然的比值。为了依据 BF_{01} 进行统计推断，研究者通常需要预先设定贝叶斯因子的判断标准（胡传鹏等，2018），例如， $BF_{01} < 1/10$ 表示存在较强的证据支持备择假设， $BF_{01} < 1/3$ 表示存在中等强度的证据支持备择假设， $BF_{01} > 3$ 表示存在中等强度的证据支持零假设， $BF_{01} > 10$ 表示存在较强的证据支持零假设（Wagenmakers et al., 2017；胡传鹏等，2018）。通过比较 BF_{01} 和判断标准，研究者可以进行相应的统计推断。上述结果显示， $BF_{01} = 4.61e-17$ ，远远小于 $1/10$ ，说明存在较强的证据支持图片类型的备择假设，因此我们推断 LPP 的波幅在图片类型之间存在差异。接下来，我们检验 LPP 在不同组别间的差异是否为 0：

```
# 组间比较
result <- hypothesis(model_4, "group2M1=0")
result$hypothesis$Evid.Ratio # 查看贝叶斯因子 BF_01

## [1] 8.89509
```

$group2M1=0$ 表示组别（group）的零假设，即抑郁组与对照组被试在 LPP 上的表现差异为 0。结果显示， $BF_{01} = 8.90 > 3$ ，这表明存在中等强度的证据支持组别的零假设，从而我们推断组别之间的 LPP 波幅不存在差异。

此外，我们还可以检验交互作用是否为 0：

```
result <- hypothesis(model_4, "group2M1:type2M1=0")
result$hypothesis$Evid.Ratio # 查看贝叶斯因子 BF_01

## [1] 11.74594
```

$group2M1:type2M1=0$ 表示组别（group）和图片类型（type）交互作用的零假设，即抑郁组中积极图片和中性图片 LPP 波幅差异与对照组中积极图片和中性图片 LPP 波幅差异的差异（即差的差）为 0。结果显示， $BF_{01} = 11.75 > 10$ ，这表明存在较强的证据支持交互作用的零假设。