这是标题

First Author1 & Ernst-August Doelle1,2

1 Wilhelm-Wundt-University

2 Konstanz Business School

Author note

author note here.

The authors made the following contributions. First Author: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Ernst-August Doelle: Writing - Review & Editing, Supervision.

Correspondence concerning this article should be addressed to First Author, Postal address. E-mail: [my@email.com](mailto:my@email.com)

Abstract

这是摘要。

*Keywords:* bayesian; GLMM

这是标题

# 1 引入

近年来，心理学研究所使用的统计方法日渐丰富。随着新的统计方法的推广与普及，心理学研究者逐渐意识到了一些传统统计方法的局限性。 重复测量方法分析，作为心理学实验中最常用的统计方法之一，也不例外。例如，重复测量方差分析通常只是对各个条件的均值进行分析，而忽略了这些均值的不确定性（uncertianty）。 其次，重复测量方法分析一般只能包含一个随机因素（例如，被试或刺激材料），而无法同时考虑多个随机因素；且它无法处理实验数据的复杂层级结构，比如对实验难度的操作嵌套于被试只内，而被试又嵌套于不同的实验组中。 此外，重复测量方差分析在处理条件均值的缺失时，通常只能删除该被试的所有数据，这会进一步造成数据的流失甚至可能会导致有偏差的估计。

相比之下，线性混合效应模型（Linear Mixed-effects Model, LMM），亦称多水平模型（Multilevel Model）或层级模型（Hierarchical Model）可以更好地处理这些问题。 例如，混合模型可以利用试次水平的数据考虑各个条件均值的不确定性，变量间可能存在的相互依存关系，以及数据中的层级结构，从而能够更好地处理数据缺失，提供更准确的估计和更高的统计检验力 (Tuerlinckx et al., 2006)。 随着统计工具（比如 R 语言）的成熟，在心理学研究中使用混合效应模型来拟合数据已经不再困难。 Bono 等 (2021) 最近的系统性综述发现，在心理学领域使用线性混合效应模型的研究已经越来越多。随着混合效应模型的应用场景逐渐丰富，以及模型的复杂程度日渐提升，贝叶斯（Bayesian）混合效应模型也因贝叶斯数据分析的独特优势受到了越来越多研究者的关注 (Sorensen & Vasishth, 2016) (Bürkner, 2017a) 。 首先，与频率论（Frequentist）相比，贝叶斯方法可以提供更直观的参数估计。贝叶斯分析的可信区间（credible interval）描述了各个参数值的合理性，即哪个参数值更可能是真值，且该区间的系数（例如95%）表示该区间包含真值的可能性。而这些通常被误认为是频率论置信区间（confidence interval）的属性 (Morey et al., 2016)。 其次，与频率论不同，贝叶斯分析可以为虚无假设（null hypothesis）提供直接的证据。例如，贝叶斯因子（Bayes factor）可以直观地描述在备择假设和虚无假设的基础上观测到当前数据可能性（likelihood）的比值 [@schmalz2021a]（详见顾昕，本期）。 再次，当混合效应模型试图包含复杂的随机效应结构时，基于频率论的模型通常难以收敛，或是对效应之间的相关性给出错误的估计 (Matuschek et al., 2017)。相比之下，基于贝叶斯的模型能够更好的处理这些复杂的随机效应结构 (Sorensen & Vasishth, 2016)。 此外，贝叶斯方法为混合效应模型中多重比较问题提供了一个更加自然的解决方案 (Gelman et al., 2013)。 最后，贝叶斯分析可以整合研究者的先验知识，从而提供更合理的参数估计。比如，研究者可以根据以往研究的效应量为当前混合效应模型设定先验分布。

鉴于贝叶斯混合模型的诸多优势，本文致力于向读者提供一个易于理解和上手的教程。本文主要包括如下几个部分： 我们首先介绍混合模型和贝叶斯统计的相关基本概念与原理。 接下来，我们将借助模拟数据讨论如何使用统计编程语言 R的 brms包来进行贝叶斯混合模型的数据分析，以及如何报告结果。 最后，我们将总结模型拟合过程中的常见问题，并对其应用进行展望。

## 2 具体介绍

线性混合效应模型（linear mixed-effects models，LMM；又称 hierarchical models 或 multilevel models），简称混合模型 (Baayen et al., 2008)，是一般线性回归模型的一种扩展模型，它通过随机效应参数捕捉观测数据之间的依赖性，比如每个被试间观测数据的依赖性。与忽略这些依赖关系的传统分析相比，混合模型可以提供更准确的效应估计，以及更高的统计检验力 (Tuerlinckx et al., 2006)。 在解释混合模型结果时，往往会提及两个特有的统计概念，即固定效应与随机效应（fixed and random effects）。其中，固定效应代表特定模型项（即主效应或交互作用）或参数在总体或群体层面对因变量的平均效应。固定效应通常是研究者感兴趣的实验操作变量，它的统计学意义的与方差分析或回归模型中对给定项的统计检验的结果基本相同，即固定效应可以用来检验因素水平之间的差异性假设。相比之下，随机效应解释的是数据中来自不同来源随机变量的变异，比如不同被试与不同实验刺激带来的变异，而这些效应往往是研究者不感兴趣以及想要控制的变异。 需要注意的是，这些随机变量的来源必须是分组变量，即称名变量，如实验条件、实验刺激或被试，而连续变量不能作为随机效应。在真实实验中，随机效应的意义在于，当研究者希望了解某个实验因子是否对所有被试具有普遍影响时，即实验因子的效应不会受到被试差异影响，通过在模型中指定随机效应，研究者可以排除被试差异带来的特异性，从而获得一个更普遍的对固定效应更精确的估计。

为了让读者更好的理解固定效应与随机效应的概念，以及混合效应模型与方差分析的关系，我们将从最基本的回归模型与方差分析的数学原理开始介绍，以便读者更好的理解为何混合效应模型可以解决传统分析方法无法解决的问题。

首先，一般线性回归模型的表达式如 [公式 1](#eq:eq1)。其中， 是因变量，代表观测变量， 为第 到 个观测点。 是自变量，代表了感兴趣的操作变量，比如，任务难度或者学生成智力。参数 是回归截距项，一般代表随机变量的总体平均值；参数 为回归斜率项，代表研究因子效应的大小。 为残差，代表模型不能解释的部分变异。 需要注意的是，使用一般线性模型需要满足残差服从独立同分布的前提假设，即每一个观测值的残差是相互独立的，并且在总体上所有观测值的残差应该遵循正态分布。对于现实中收集的数据，比如包含重复测量的实验数据，单个被试内的数据具有相互依存性，因此独立性的前提假设难以满足。

当任务难度为组间设计时，每个观测数据点为各被试的平均反应时，此时一般线性模型的前提假设更容易满足，因此可以通过一般线性模型来检验不同任务难度对于反应时的影响。然而，当任务难度为组内设计时，每个被试会在不同任务难度条件下产生两个不同的观测数据点，即存在重复测量，此时一般线性模型的前提假设难以满足，只能使用配对 t 检验或者重复测量方差分析对数据进行检验。 实际上，配对t检验或者重复测量方差分析只是额外的考虑了不同被试带来的变异，但并没有考虑每个被试内试次间相互影响变异，即对每个被试在不同任务难度下的反应时求平均值依然不能完全考虑各被试中不同试次的依赖程度。 为了解决这个问题，混合效应模型将任务难度作为固定效应，将被试随机效应，并且考虑了各被试中不同试次的变异，即使用混合效应模型时不需要将观测数据在被试或者实验条件层面进行平均或汇总。

当混合效应模型只考虑固定效应时，其表达式与一般线性模型类似，其表达式如下公式 [公式 2](#eq:eq2)：

其中， 为观测变量，比如反应时， 为被试编号， 为试次编号， 为实验条件，比如任务难度，简单条件编码为-1，困难条件编码为1。 与 为固定效应， 反应了所有试次的平均反应时， 反应了任务难度对于反应时的效应，即两个条件下所有被试反应时的差异。 为模型残差。 此时，由于模型没有加入随机效应，因此无法解释反应时的个体差异，以及固定效应在不同被试间的一致性。

为了解释反应时的个体差异，可以在混合模型中加入随机截距，其表达式如下 [公式 3](#eq:eq3)：

与 [公式 2](#eq:eq2) 相比，显而易见的是随机截距就是指影响截距项 的随机效应。其中， 不再是所有试次反应时的均值，而是所有被试平均反应时的均值。而 代表了每个被试的平均反应时（并不一定等于每个被试的平均反应时），即等于被试平均反应时的均值 加上个体偏移量 。 因此，加入随机截距的混合效应模型可以解释不同被试 在反应时上的个体差异 ，这已经足以解释各数据点之间的相关性，比如有些被试的反应更慢，而有些反应更快。 然而，该模型依然无法解释固定效应在不同被试间的一致性。很容易想象，任务难度对反应时的影响对不同被试是不同的，比如某些被试在两种任务难度上的反应都很快，这就导致任务难度的固定效应在这些被试上是小于另一部分被试的。

为了解释固定效应在被试间的一致性，进一步在混合模型中加入随机斜率，其表达式如下 [公式 4](#eq:eq4)：

与 [公式 3](#eq:eq3) 相比，随机斜率指影响斜率项 的随机效应。其中， 指固定效应在所有被试上的平均值。而 指每个被试独特的固定效应，即等于被固定效应在所有被试上的平均值 加上个体偏移量 ，使用 作为固定效应的个体偏移量是为了避免与因变量均值的个体偏移量 混淆。 此外，随机效应之间可能存在相关性，即随机截距 与随机斜率 服从均值为零的多元正态分布。因此，在建立混合模型前需要假设是否存在随机效应间的相关性，一般混合模型建立时会默认存在随机效应间的相关性。

至此，我们讨论了如何在混合模型中加入不同的随机效应以及他们的相关性来解释被试之间的异质性以及观测数据中的依赖性。需要注意的是，研究者常关心的效应（主效应与交互效应）往往可以与模型的固定效应相对应，并且固定效应的解释不会因为加入随机效应而变化，此外，加入随机效应还能增加对于固定效应估计的准确性。

在确定混合效应模型时，通常需要考虑包含最复杂的随机效应结构 (Barr et al., 2013)，即及考虑所有潜在存在的随机截距与随机斜率。然而，此时的模型可能会遭遇不收敛或给出异常估计的问题 (Bates et al., 2015)。 相比之下，这种包含复杂随机效应结构的模型通常可以在贝叶斯框架中得到更好的拟合 (Sorensen & Vasishth, 2016) (Eager & Roy, 2017)。除此之外，使用贝叶斯方法进行模型拟合还包含其他优势，比如对于参数估计、假设检验、以及结合先验经验的优势，这在前文都有讨论。因此，我们接下来讲介绍如何使用贝叶斯方法来拟合混合模型。

基于贝叶斯框架的模型拟合的特点在于，不同于频率主义框架中将模型参数与实验效应视作固定真值，即当抽样次数接近无穷大时两个样本之间的差距等同于两个总体真实存在的差异，贝叶斯框架将模型参数与实验效应视作存在不确定性的概率事件，即两个总体间的差异不是固定值，而是存在不确信噪音的概率分布。 用贝叶斯公式表示如下 [公式 5](#eq:eq5)：

公式右部 为后验概率分布，代表在得到实验数据 的条件下，实验效应为 的概率。 后验分布可以通过公式左边部分计算，其中为先验概率，代表了在未获得实验数据时对实验效应的假设，即可以通过以前的研究实验效应设定该实验的先验效应；为似然，代表了在实验效应已知的条件下实验数据出现的概率，可以理解为在混合效应模型已知的条件下，通过该模型预测或生成不同模拟数据的概率；为边际似然，代表了数据随机抽出的概率，不同于似然是在两个已知差异的总体中进行采样，边际似然则是在所有总体的组合中进行样本的采样，因此可以表示为在所有不同实验效应下似然的和，即，当为连续变量时，，更多详情请参考 (Kruschke & Liddell, 2018)。 贝叶斯公式的意义在于说明，参数（实验效应），数据和模型的关系，即如何在结合先验知识和数据的情况下推测模型的参数值。然而，将精力过多的投入到贝叶斯方法本身并不会增强我们对建立模型的理解，真正重要的关键在于如何结合贝叶斯方法去拟合线性模型。 为了便于理解，首先，我们假设实验数据来自于一个正态分布，，其中 为 个观测数据点， 为正态分布， 为数据分布的均值， 为变异。在贝叶斯框架下，该数据分布等同于 似然（参数 包含均值 和变异 ），因此，只需额外得到先验与边际似然就可以推测出后验分布。其中，实验数据 已知，模型参数 未知，贝叶斯推断的目的就在于推测模型参数值，即实验效应大小。当假设先验参数值 已知时（可以根据以往研究结果进行设定），此时似然也是已知的 。由于边际似然 只是一个标量，且只会对后验分布进行缩放，因此忽略边际似然的计算，可以根据贝叶斯公式推断得到参数后验分布 。需要注意的，参数后验分布 不同于先验参数分布 ，后者在前者的基础上额外考虑了数据似然的作用。换句话说，不同于频率学派框架直接通过数据得到模型参数，贝叶斯框架通过实验数据更新先验参数的值得到模型的后验参数。此时，贝叶斯框架可以看作是频率学派的一种拓展与延申。 通过贝叶斯模型拟合线性回归模型的方式与上述过程类似。 首先，将回归模型 转写为 ，此时模型参数 包括 、 和 三个部分。 代表了实验条件，比如任务难度，其中简单任务 ，困难任务 。可见，当任务为简单时，实验数据采样与均值为 ，变异为 的正态分布，当任务为困难时，数据采样与均值为 ，变异为 的正态分布。 因此，在已知两个实验条件下的实验数据，以及设定先验参数分布后，可以通过贝叶斯公式得到参数的后验分布。 根据同样的原理，将混合模型的公式带入贝叶斯框架中可以得到如下概率分布表达 [公式 6](#eq:eq6)：

此时的模型参数包括 、、、 和 五个个部分（如果考虑随机效应间的相关性还需设定其他参数，这里尽量简化模型方便理解）。

由于模型参数增加，贝叶斯在计算时会遇到困难，因此会采用mcmc采样，即brms中使用的算法。这种算法通过从后验分本中采集样本来模拟真实的后验分布。

# 3 假想的心理学实验

我们借助一个假想的实验来展示如何使用贝叶斯混合效应模型分析实验心理学的数据。该实验的目的是探究抑郁症患者加工不同类型图片的神经基础。在该实验中，我们招募了抑郁症患者和健康对照组被试各30人。所有的被试都观看了30张正性图片和30张负性图片，且每张图片呈现10次。在实验过程中，我们记录了被试的脑电。我们所关注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：抑郁症患者depression、对照组；被试间control) 2 (图片类型type：正性positive、负性negative；被试内) 的混合实验设计。

## 3.1 模拟数据

我们采用(DeBruine, 2021)的library(faux)包生成该假想实验的数据。具体假定参数如下：

subj\_n <- 60 # 总被试量：抑郁患者30人，健康对照组被试30人  
trial\_n <- 10 # 每张图片呈现的次数  
# 固定效应  
b0 <- 2.5 # 截距 (所有条件的均值)  
b1 <- 4.2 # 图片类型的固定效应 (主效应)  
b2 <- 4.5 # 组别的固定效应 (主效应)  
b3 <- 2.4 # 图片类型与组别的交互作用  
fixed\_true <- c(b0,b2,b1,b3)   
  
# 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)  
u\_sd\_subj\_true <- c(u0s,u1s)  
# 误差项  
sigma <- 4

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("negative","positive")) %>%  
 # 每个图片呈现10次  
 add\_random(trial = trial\_n) %>%  
 # 图片类型的编码：负性=-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  
# 设定-0.5和0.5的编码  
contrasts(df\_simu$group) <- MASS::contr.sdif(2)  
contrasts(df\_simu$type) <- MASS::contr.sdif(2)  
head(df\_simu, 5)#查看数据结构

# 3.3 数据分析

接下来对数据进行建构贝叶斯线性混合效应模型。在这部分将使用brms的R包 (Bürkner, 2017b) 进行分析。Brms目前常用贝叶斯模型分析之中(Bürkner, 2017b, 2017c; Ladislas et al., 2019; Vasishth et al., 2018)。关于brms包的更多详细内容可以见以下的网址 <https://paul-buerkner.github.io/brms/> 。

## 3.3.1 先验的设置

正如前文所述，在运用贝叶斯模型之前，我们需要先设置先验信息。而如何设置一个合适的先验信息对于初学者来说却是并不容易的事。先验分布是模型参数的预先设置的分布形态，它赋予了模型参数在数据分析之前相对合理性的信息，然后用数据的信息更新先验信息，以获得模型参数的后验信息。在我们的数据模型中需要建立**LPP = *+*\*group +  *\** type +++** 这样的线性混合效应模型，那么需要指定固定截距**，固定斜率**,固定斜率 以及误差项，随机斜率，随机截距的先验分布。在brm函数使用中如果没有设置先验，那么会默认采用负无穷到正无穷的平均分布，但是这样极容易影响后验估计的准确性，致使得出错误的结论(Veenman et al., 2022)。在设置先验分布时，我们需要确定参数的先验分布的大致形状，常见的分布有二项分布、正态分布、泊松分布等。而先验分布的大致分布可以利用先验信息进行获取。具体方法可以参考(Sarma & Kay, 2020)。在这个例子中，我们预设参数是符合正态分布的，接着需要确定正态分布的平均数以及方差等超参数。这时便需要进行先验预测（Prior predictive distribution），以检验所设置的超参数是否合适。我们可以采用brms包里面的sample\_prior ="only"进行检验。 在model\_predict中，LPP是因变量，group和type是自变量。family代表模型中使用的反应分布和相关函数。prior定义了模型的先验信息，如果不设置这个参数，那么模型建立中*先验*会默认使用均匀分布。cores当并行执行链时使用的核数，默认为1。chains 代表马尔可夫链的数量(默认为4)。iter参数用于马尔可夫链蒙特卡罗（Markov Chain Monte Carlo , MCMC）算法的总迭代次数。warmup参数指定在过程开始时运行的迭代次数来校准MCMC，以便最后只保留iter - warmup ’迭代来近似后端分布的形状。如果想要了解更多，可以见(McElreath, 2016)。 之后我们使用pp\_check进行检验，看先数据分布（即因变量LPP的概率分布）的平均数（mean）、最大值（max），最小值（min）是否在先验预测分布中，如果平均数、最大值、最小值都包括在其内，那么这个先验分布的设置是合适的。关于prior的设置可具体参考set\_prior函数对先验的设置。

## 定义先验分布  
prior = c(  
 prior(normal(0, 3), class = Intercept),  
 prior(normal(0, 3), class = b),  
 prior(normal(0, 3), class = sigma),  
 prior(normal(0, 3), class = sd),  
 prior(lkj(2), class = cor)  
)  
model\_predict <- brm(LPP ~ group \* type + (1 + type | subj),   
 data = df\_simu,  
 family="gaussian",  
 cores = 4,  
 chains = 4,  
 warmup =1000,  
 iter = 5000,  
 sample\_prior ="only",  
 prior=prior)

pp\_check(model\_predict, type = "stat", stat = "min")

pp\_check(model\_predict, type = "stat", stat = "max")

pp\_check(model\_predict, type = "stat", stat = "mean")

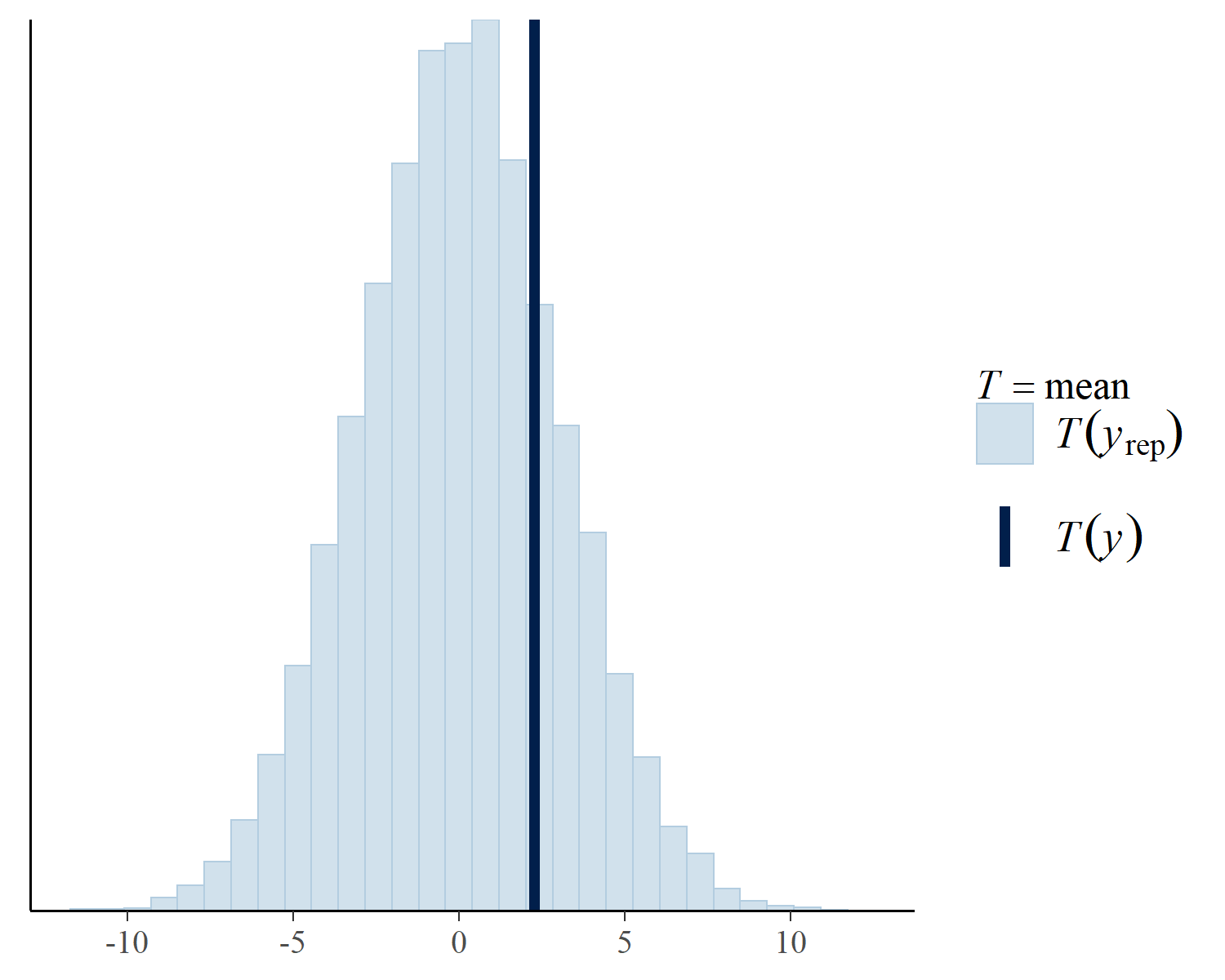
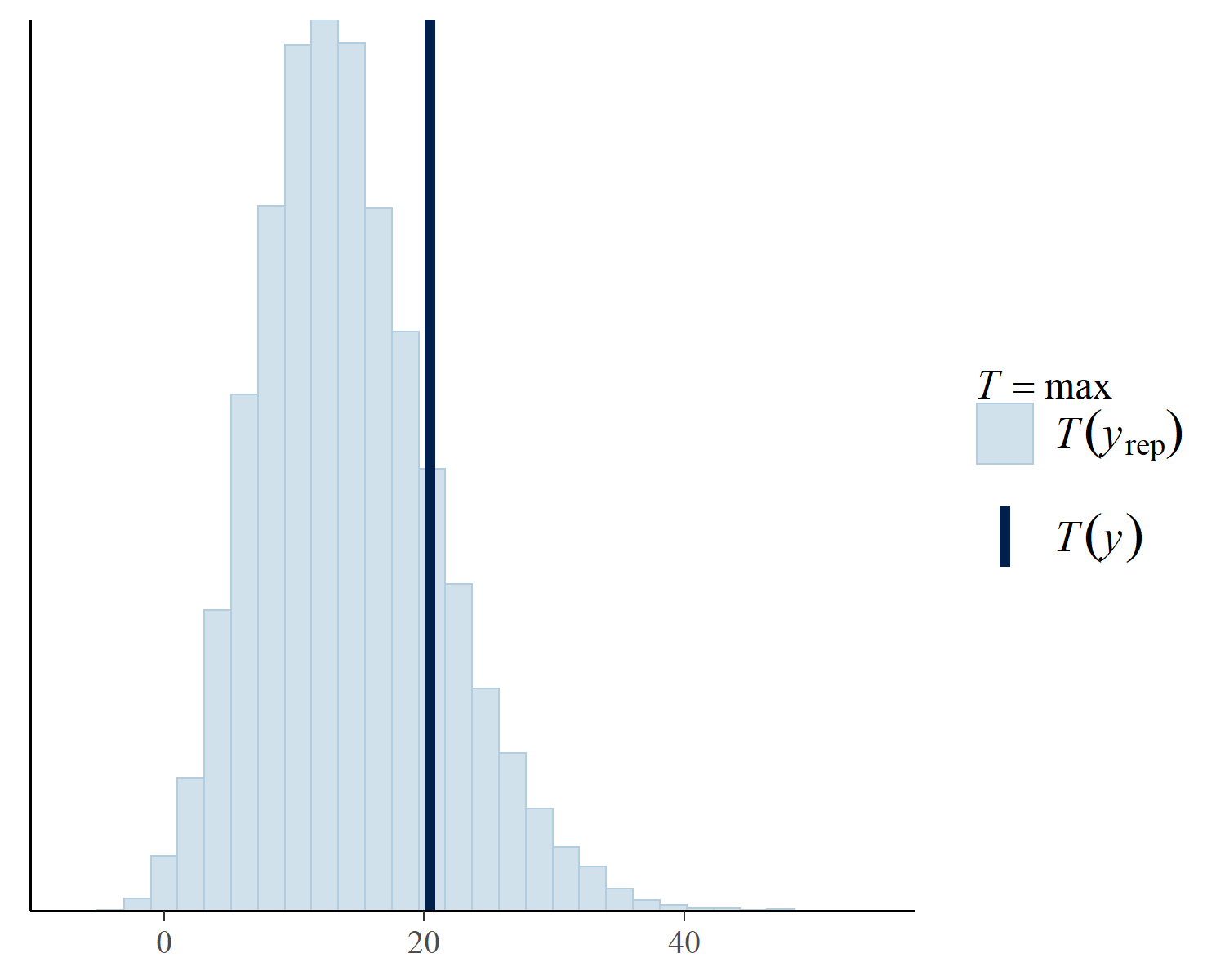
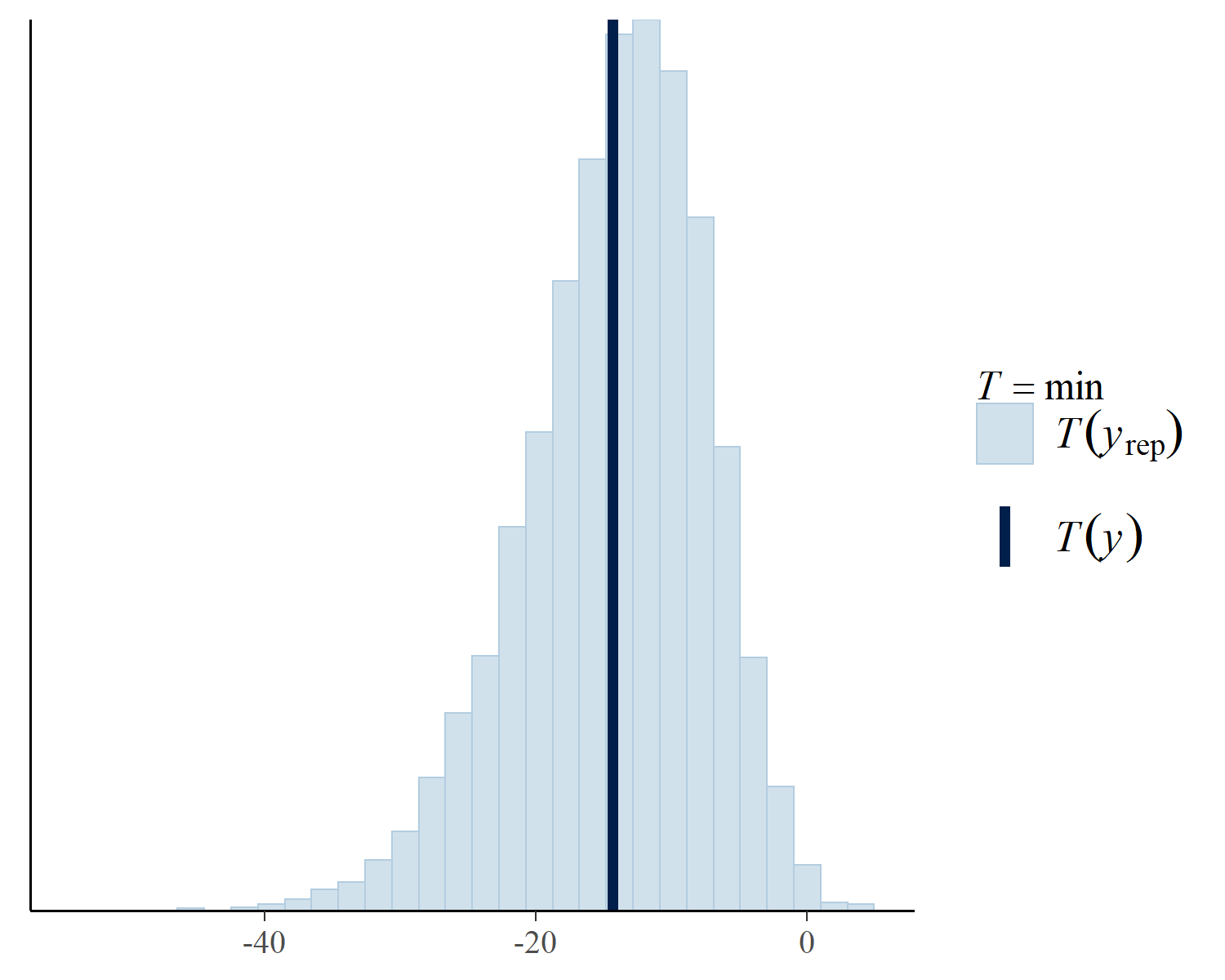


图1： 混合效应模型的均值、最大值、最小值的先验预测分布。 先验预测分布的均值、最小值和最大值的分布被标记为yrep

从图1可知，先验分布是合适的。如果觉得与实际数据的分布相差还是很大，可以再次调整预设的先验分布的超参数，再查看目前先验预测的分布是否符合设想的参数分布，如果不合适可调整至合适的参数为止。

## 3.3.2 添加固定效应

为了更好地理解贝叶斯混合线性模型，首先我们建立只有固定效应的模型。model\_1中brm函数后的参数设置参见先验的设置部分的说明。与先验的设置不同，在模型建立中采用sample\_prior ="yes",以期纳入现有数据的分布影响，从而得到数据模型的后验分布。summary可以帮助我们展示模型拟合后的参数情况。plot可以显示模型拟合后的参数的分布情况以及MCMC情况的图示化结果。

# 定义参数的先验信息  
prior <- c(  
 prior(normal(0, 3), class = Intercept),  
 prior(normal(0, 3), class = b),  
 prior(normal(0, 3), class = sigma)  
)  
#建立模型  
model\_1 <- brm(LPP ~ group \* type,   
 data = df\_simu,  
 family="gaussian",  
 cores = 4,  
 chains = 4,  
 warmup = 1000,  
 iter = 5000,  
 sample\_prior ="yes",  
 prior = prior)

## Family: gaussian   
## Links: mu = identity; sigma = identity   
## Formula: LPP ~ group \* type   
## Data: df\_simu (Number of observations: 1200)   
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;  
## total post-warmup draws = 16000  
##   
## Population-Level Effects:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## Intercept 2.26 0.13 2.00 2.52 1.00 19351 12166  
## group2M1 4.70 0.27 4.17 5.23 1.00 18270 12388  
## type2M1 4.30 0.26 3.78 4.82 1.00 17345 12367  
## group2M1:type2M1 2.44 0.52 1.43 3.44 1.00 19975 11594  
##   
## Family Specific Parameters:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## sigma 4.60 0.09 4.42 4.79 1.00 19928 13219  
##   
## 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代表固定效应部分参数的后验情况。Intercept 代表模型中的固定截距。group2M1参数估计代表，指的是组别变量的固定斜率。type2M1参数估计代表，指的是图片类型变量的固定斜率。而group2M1以及type2M1中的2M1代表组别之间以及类型之间的效应的相减，即或者。 group2M1:type2M1代表组别和图片类型交互作用的固定斜率。Estimate表示参数的估计值，Est.Error代表估计值的标准误。l-95% CI和u-95% CI代表参数的可信区间。sigma代表残差项的后验分布参数。 Rhat代表参数的收敛情况。在模型拟合的数据中如果Rhat>1.01说明模型的收敛情况并不理想，Rhat应该尽量接近于1，这说明模型的收敛情况较好，但是尽量不要超过1.01。另外也可以使用bayestestR 包的函数进行判断模型的收敛情况。

if (!require(bayestestR)) {library(bayestestR)}  
post\_diag<- diagnostic\_posterior(model\_1)  
effectsize::interpret\_rhat(post\_diag$Rhat)

表1：model\_1的后验均值, 标准误, 95%的可信区间以及Rhat

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 后验均值 | 标准误 | 下限 | 上限 | Rhat | Bulk\_ESS | Tail\_ESS |
| 截距 | 2.258 | 0.133 | 2.001 | 2.521 | 1.000 | 19,350.535 | 12,165.533 |
|  | 4.703 | 0.267 | 4.174 | 5.225 | 1.001 | 18,269.737 | 12,388.054 |
|  | 4.300 | 0.264 | 3.778 | 4.824 | 1.000 | 17,345.354 | 12,367.401 |
|  | 2.437 | 0.516 | 1.426 | 3.441 | 1.001 | 19,974.942 | 11,593.742 |

表 1显示了model\_1拟合后的参数的后验情况，包括后验分布均值，标准误，95%的可信区间以及Rhat。

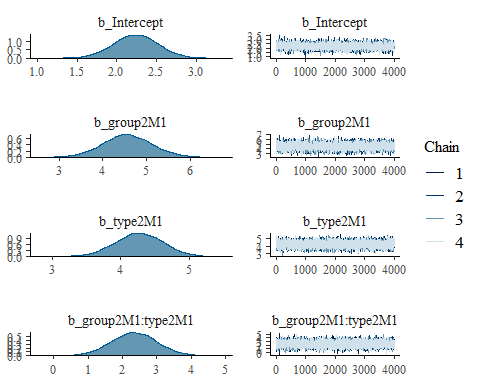


图2 左列是参数的后验分布图，右列是参数抽样的轨迹图。

图2描述了model\_1中固定截距、固定斜率、残差的后验分布以及MCMC抽样情况。图的左边是参数的后验分布，x轴代表参数值；图的右边是逼近后验分布的两个行为的模拟(即4条链的轨迹图)，x轴代表迭代数，y轴代表参数值。这轨迹图显示了平均值周围的随机波动。

## 3.3.3 添加随机效应

接下来在后面的模型中加入一些随机效应。而随机效应项一般是由随机截距和随机斜率组成。正如前文提到随机效应的设置是为了更好地排除由个体差异的因素所带来的影响，这样可以使固定效应的参数更加精准的估计。在brms包中提供了对随机效应设置的语法，我们可以在brm函数采用(1 | xxx )类似的结构进行设置随机截距，xxx可以设置随机变量，如被试subj或者实验因素type，而在我们需要建立的模型中，我们只对subj指定了随机效应。然后在1后加上需要设置随机斜率的因子， 如（1 + type | subj）。下面我们在上述模型中添加随机截距。

prior = c(  
 prior(normal(0, 3), class = Intercept),  
 prior(normal(0, 3), class = b),  
 prior(normal(0, 3), class = sigma),  
 prior(normal(0, 3), class = sd)  
)  
model\_2 <- brm(LPP ~ group \* type + (1 |subj),   
 data = df\_simu,  
 family="gaussian",  
 cores = 4,  
 chains = 4,  
 warmup =1000,  
 iter = 5000,  
 sample\_prior ="yes",  
 prior = prior)

summary(model\_2)#查看参数情况  
plot(model\_2) #查看每个参数的分布和抽样情况

## Family: gaussian   
## Links: mu = identity; sigma = identity   
## Formula: LPP ~ group \* type + (1 | subj)   
## Data: df\_simu (Number of observations: 1200)   
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;  
## total post-warmup draws = 16000  
##   
## Group-Level Effects:   
## ~subj (Number of levels: 60)   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## sd(Intercept) 2.01 0.23 1.59 2.51 1.00 5502 7316  
##   
## Population-Level Effects:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## Intercept 2.24 0.29 1.68 2.80 1.00 6542 9347  
## group2M1 4.57 0.57 3.46 5.68 1.00 6995 9235  
## type2M1 4.30 0.24 3.83 4.77 1.00 38560 11656  
## group2M1:type2M1 2.45 0.48 1.50 3.38 1.00 38370 10959  
##   
## Family Specific Parameters:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## sigma 4.17 0.09 4.01 4.35 1.00 30497 10815  
##   
## 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是model\_2的固定效应部分，解释与model\_1类似。与model\_1不同，model\_2加入了随机截距部分。 Group-Level Effects则是随机截距的参数的后验分布情况。

然后在model\_2模型的代码基础之上加上随机斜率,探究在不同被试下图片类型对LPP波幅的影响。

prior = c(  
 prior(normal(0, 3), class = Intercept),  
 prior(normal(0, 3), class = b),  
 prior(normal(0, 3), class = sigma),  
 prior(normal(0, 3), class = sd),  
 prior(lkj(2), class = cor)  
)  
model\_3 <- brm(LPP ~ group \* type + (1 + type | subj),   
 data = df\_simu,  
 family="gaussian",  
 cores = 4,  
 chains = 4,  
 warmup =1000,  
 iter = 5000,  
 sample\_prior ="yes",  
 prior=prior)

summary(model\_3)#查看参数情况  
plot(model\_3) #查看每个参数的后验分布和抽样情况

plot(model\_3, variable = "^b", regex = TRUE) #查看固定效应部分的参数分布情况

## Family: gaussian   
## Links: mu = identity; sigma = identity   
## Formula: LPP ~ group \* type + (1 | subj)   
## Data: df\_simu (Number of observations: 1200)   
## Draws: 4 chains, each with iter = 5000; warmup = 1000; thin = 1;  
## total post-warmup draws = 16000  
##   
## Group-Level Effects:   
## ~subj (Number of levels: 60)   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## sd(Intercept) 2.01 0.23 1.59 2.51 1.00 5502 7316  
##   
## Population-Level Effects:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## Intercept 2.24 0.29 1.68 2.80 1.00 6542 9347  
## group2M1 4.57 0.57 3.46 5.68 1.00 6995 9235  
## type2M1 4.30 0.24 3.83 4.77 1.00 38560 11656  
## group2M1:type2M1 2.45 0.48 1.50 3.38 1.00 38370 10959  
##   
## Family Specific Parameters:   
## Estimate Est.Error l-95% CI u-95% CI Rhat Bulk\_ESS Tail\_ESS  
## sigma 4.17 0.09 4.01 4.35 1.00 30497 10815  
##   
## 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是model\_3的固定效应部分，解释与model\_1类似。与model\_2不同，model\_2加入了随机斜率部分。 Group-Level Effects则是随机截距以及随机斜率参数的后验分布情况。

至此我们已经完成了线性混合效应模型的建构。

# 3.3.4 模型拟合效果与真实值比较

在这部分，我们验证一下模型拟合效果，将模型拟合参数与预设的参数真实值之间进行对比。

#比较固定效应中各参数与之前设置的总体参数（即真实值）之间的比较。  
as.data.frame(model\_3) %>%  
 select(starts\_with("b\_")) %>%  
 mcmc\_recover\_hist(true = fixed\_true)

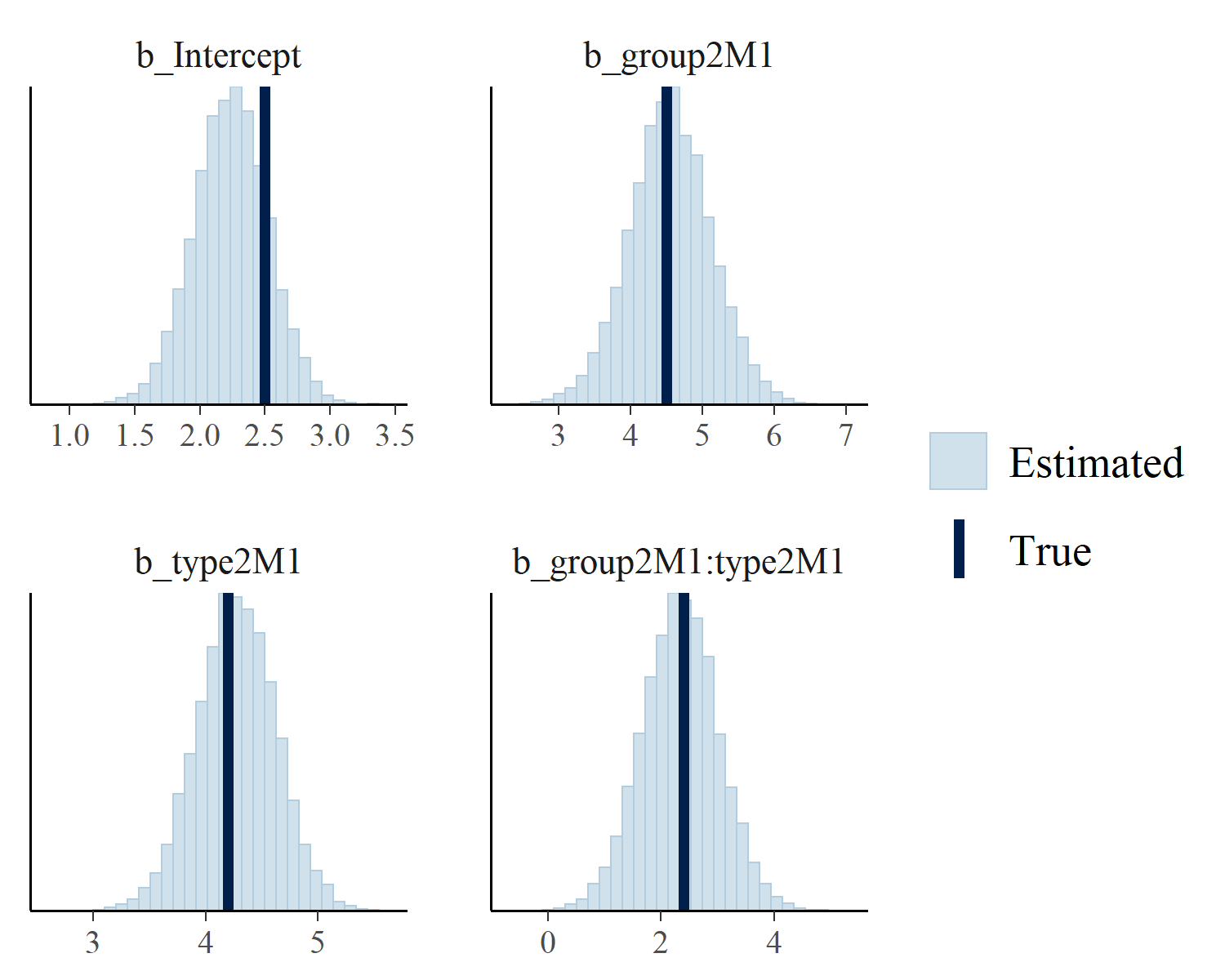


图3 固定效应中各参数与之前设置的总体参数（即真实值）之间的比较

as.data.frame(model\_3) %>%  
 select(starts\_with("sd\_subj")) %>%  
 mcmc\_recover\_hist(true = u\_sd\_subj\_true)

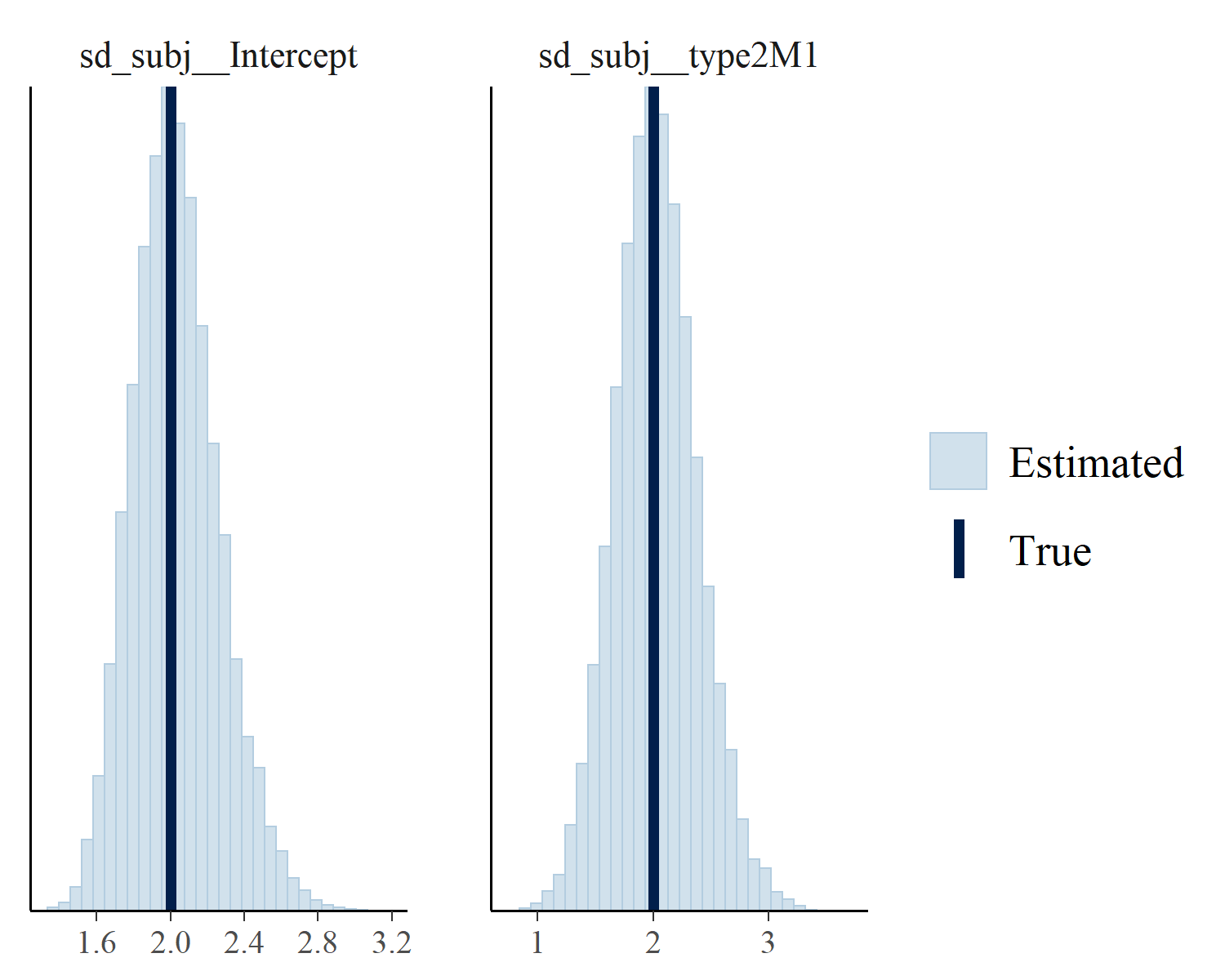


图4 预设的被试随机效应总体参数(即真实值)和拟合的被试的随机效应的估计参数

从图3和图4中可知,模型拟合的参数的分布还是包含预设值的，这表明模型拟合度还是比较好的。不过对于研究者来说，更重要的还是想比较组间差异：两组之间的LPP波幅是否存在显著性的差异。我们可以通过贝叶斯因子（Bayesian factor, BF）对模型拟合的一些参数进行比较 (Heck et al., 2022;Schad et al., 2022;Wagenmakers et al., 2010)，从而更好地观察两个组相关的自变量因素对因变量是否存在差异，这样更能够解释两组之间是否存在差异。

# 3.3.5组间差异以及组内差异的比较

研究中通常需要比较组间差异或者组内各条件的差异。使用贝叶斯模型进行组间或者组内条件比较时需要注意，在贝叶斯线性混合效应模型中，一般可以将组间条件和组内条件的参数纳入，从而比较组间因素与组内因素相关的参数。我们可以通过library(BayestestR)(Makowski et al., 2019)计算模型中的可信区间、最大后验概率估计(Maximum A Posteriori, MAP)、显著性方向概率(Probability of Direction, pd)、全后验分布百分比（percentage of the full posterior distribution,ROPE）、贝叶斯因子（Bayes factor, BF）等 (Makowski et al., 2019)。在贝叶斯模型参数比较中，比较推荐报告以上的指标 (Makowski et al., 2019)。ROPE中的百分比是一个重要的指数，它一定程度上可以作为显著性指数的标准 (Makowski et al., 2019)。关于除ROPE指标外其他贝叶斯模型模型参数报告的更多详细内容可见 <https://easystats.github.io/bayestestR/articles/guidelines.html>。

describe\_posterior(model\_3,centrality = "all", dispersion = TRUE,ci =0.95,test = "all")

表2 模型参数比较的均值，标准差，95%CI，ROPE，Rhat，BF

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 参数 | 均值 | 标准差 | 95%CI | ROPE | % in ROPE | Rhat | BF |
| 截距 | 2.24 | 0.29 | [1.66,2.80] | [-0.56, 0.56] | 0% | 1.00 | 1.09e+06 |
| group2M1 | 4.53 | 0.57 | [3.39, 5.62] | [-0.56, 0.56] | 0% | 1.00 | 2.03e+06 |
| type2M1 | 4.27 | 0.35 | [3.57, 4.97] | [-0.56, 0.56] | 0% | 1.00 | 5.17e+10 |
| group2M1:type2M1 | 2.35 | 0.68 | [0.99, 3.69] | [-0.56, 0.56] | 0% | 1.00 | 55.35 |

使用describe\_posterior可以输出贝叶斯模型比较的参数情况。从以上的输出可知，在两组比较中，我们会发现，参数的ROPE的百分比（即 % in ROPE）小于1%，这说明我们可以拒绝零假设，两组之间存在显著的差异性。而的ROPE的百分比也小于1%，这说明我们可以拒绝零假设，抑郁症患者和健康对照组中正性图片和负性图片存在显著性的差异，即组内各条件存在显著性差异。

当然在心理学研究中，有不少的贝叶斯统计的支持者主张使用贝叶斯因子进行假设检验 (Rouder et al., 2018)。在以下的代码中我们将采用直接计算贝叶斯因子对组间差异以及组内差异进行检验。使用brms的hypothesis函数便可直接进行假设检验，并从中提取出贝叶斯因子。

#组内比较  
n=hypothesis(model\_3,"type2M1=0")  
withinGroup\_comparion\_BF=n$hypothesis$Evid.Ratio  
withinGroup\_comparion\_BF

## [1] -2.510157e-20

plot(n)#查看参数比较的先验分布和后验分布

#组间比较  
m=hypothesis(model\_3,"group2M1=0")  
group\_comparion\_BF=m$hypothesis$Evid.Ratio  
group\_comparion\_BF #查看贝叶斯因子

## [1] 2.147205e-17

plot(m)#查看参数比较的先验分布和后验分布

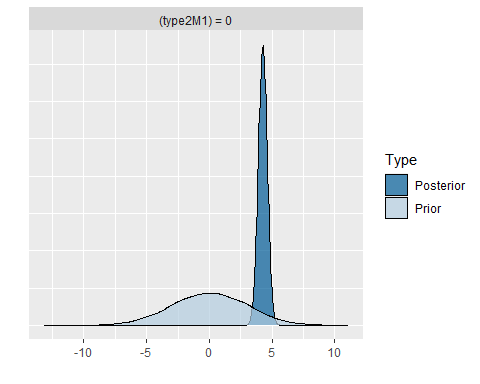
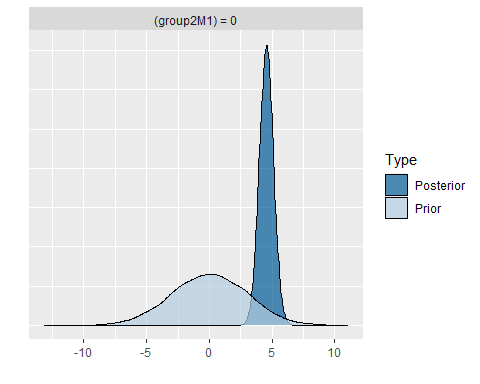
 

图5 参数比较的先验分布和后验分布

通过hypothesis$Evid.Ratio，我们可以得到，而 = 1/ 。在 以及中，>0且<1，则则趋于无穷大，这表明越支持备择假设，说明在组间和组内都存在显著的差异。

# 4 模型拟合常见问题

## 4.1 模型不收敛

模型不能收敛时，需要考虑是否模型建立得过于复杂，可以适当减少模型纳入的参数，达到简化模型。

## 4.2 模型比较

在数据处理中，我们纳入不同的参数，建立不同的贝叶斯混合线性模型。如何判断构建的哪种模型比较好，这便需要进行模型比较了。在贝叶斯分析中，常用LOOIC指标进行判断 (Vehtari, 2017)。在brms的R包中，我们可以采用loo函数进行计算LOOIC。数值越小，代表模型拟合度越佳。

loo\_compare(loo(model\_1),loo(model\_2),loo(model\_3))

## elpd\_diff se\_diff  
## model\_3 0.0 0.0   
## model\_2 -18.5 6.6   
## model\_1 -111.9 14.4

## 4.3 缺失值处理

由于实验操作过程中难免会出现失误，可能最后数据中难免出现缺失值。最简单的解决方法无非是将缺失的地方进行删除，但是如果该数值并不是完全随机缺失的，这就可能导致数据分析中出现偏差。一般可以采用两种方法处理缺失值。一是在模型拟合之前将缺失值进行多重插值；二是在模型拟合过程中进行动态地进行缺失值的多重插值。这个可以采用mi函数进行多重插值。更多详细内容可以参考https://cran.r-project.org/web/packages/brms/vignettes/brms\_missings.html。

# 参考文献

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. *ArXiv*, *1506*.

Bono, R., Alarcón, R., & Blanca, M. J. (2021). Report Quality of Generalized Linear Mixed Models in Psychology: A Systematic Review. *Frontiers in Psychology*, *12*, 666182. https://doi.org/10.3389/fpsyg.2021.666182

Bürkner, P.-C. (2017a). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2017b). Brms Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2017c). *Advanced Bayesian Multilevel Modeling with the R Package brms* (11 citation(s); arXiv:1705.11123). arXiv. http://arxiv.org/abs/1705.11123

DeBruine, L. (2021). *Faux: Simulation for Factorial Designs R package version 1.1.0. Zenodo*. https://doi.org/10.5281/zenodo.2669586

Eager, C., & Roy, J. (2017). *Mixed Effects Models are Sometimes Terrible*.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (0 ed.). Chapman and Hall/CRC. https://doi.org/10.1201/b16018

Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, *25*(1), 155–177. https://doi.org/10.3758/s13423-017-1272-1

Ladislas, N., Cédric, B., Hélène, L., Anne, V., & Paul-Christian, B. (2019). *An Introduction to Bayesian Multilevel Models Using brms: A Case Study of Gender Effects on Vowel Variability in Standard Indonesian*. 50.

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

McElreath, R. (2016). *Statistical rethinking: A bayesian course with examples in R and stan*. CRC Press.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*(1), 103–123. https://doi.org/10.3758/s13423-015-0947-8

Sarma, A., & Kay, M. (2020). Prior Setting in Practice: Strategies and Rationales Used in Choosing Prior Distributions for Bayesian Analysis. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376377

Sorensen, T., & Vasishth, S. (2016). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists. *The Quantitative Methods for Psychology*, *12*(3), 175–200. https://doi.org/10.20982/tqmp.12.3.p175

Tuerlinckx, F., Rijmen, F., Verbeke, G., & De Boeck, P. (2006). Statistical inference in generalized linear mixed models: A review. *British Journal of Mathematical and Statistical Psychology*, *59*(2), 225–255. https://doi.org/10.1348/000711005x79857

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Veenman, M., Stefan, A., & Haaf, J. M. (2022). *Bayesian Hierarchical Modeling: An Introduction and Reassessment* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/pskvx