

贝叶斯统计方法的应用与现状*

李贵玉¹, 顾昕²

(1. 华东师范大学课程与教学研究所, 上海 200062; 2. 华东师范大学教育心理学系, 上海 200062)

摘要: 贝叶斯统计方法是心理学数据分析的热门方法。研究全面论述贝叶斯方法在心理学领域的应用与方向。现阶段贝叶斯方法以模拟研究为主, 应用方向为心理学研究常用的项目反应理论、认知诊断、计算机自适应、结构方程模型。同时, 评述发现贝叶斯方法正逐步被国内心理学研究者所接受。最后, 文章讨论了当下贝叶斯统计在心理学研究中应用的局限性及可能的原因, 建议统计学者开发界面友好的贝叶斯软件, 并在心理学课程中加入贝叶斯知识。

关键词: 贝叶斯统计; 先验; 后验; 回归模型; 系统性评述

中图分类号: B841.2

文献标识码: A

文章编号: 1003-5184(2021)05-0466-08

1 引言

统计推断是心理学数据分析的基本方法。心理学实证研究多采用极大似然估计、置信区间、零假设显著性检验等频率统计理论进行统计推断。近年来, 贝叶斯统计方法被广泛应用于心理学研究的各个领域(Wagenmakers et al., 2018)。贝叶斯统计假设未知参数服从特定概率分布, 根据主观经验或历史数据给出未知参数或模型的先验分布, 并基于实验数据更新先验分布, 得到未知参数或模型的后验分布, 用于统计推断。

贝叶斯统计方法较频率统计有着诸多优势。首先, 贝叶斯统计方法允许研究者加入实质性的先验信息, 能够有效提高参数估计的精确度(McNeish, 2016)。其次, 贝叶斯统计方法分析是动态的。先验分布将随着数据的收集而不断更新, 加深研究者对未知参数与模型的理解。这一动态过程使得贝叶斯统计方法在网络模型(Clewley, Chen, & Liu, 2011), 长时追踪数据(Sims, Neth, Jacobs, & Gray, 2013)等领域应用广泛。再次, 研究者可根据理论与模型设定先验分布中的超参数, 天然的层次结构使得贝叶斯统计方法可直接用于分析各类多层模型(Hierarchical models, Scheibehenne & Pachur, 2014)。最后, 贝叶斯检验通过衡量数据支持或拒绝零假设的程度, 给出研究理论可信度高、低或数据不敏感三个可能的推断结论, 获得的信息远比显著性检验要丰富(钟建军, Dienese, 陈中永, 2017; Gu, Mulder,

Dekovic, & Hoijtink, 2014)。

鉴于贝叶斯统计方法的优点, 国内外众多心理统计学家致力于推广贝叶斯方法(胡传鹏, 孔祥祯, Wagenmakers, Ly, 彭凯平, 2018; 王孟成, 邓俏文, 毕向阳, 2017; 吴凡, 顾全, 施壮华, 高在峰, 沈模卫, 2018)。但是, 这类综述性研究仅阐述贝叶斯统计方法的使用、优点, 以及在特定领域的应用, 对于贝叶斯统计应用现状与趋势的讨论不多, 更没有详实的评述贝叶斯统计方法的应用领域与方向。与此相对, 系统性评述(systematic review)是一种新近流行的文献综述方式, 通过搜集、归类、分析文献的方式, 得到量化的综述结论。贝叶斯统计方法的系统性评述已见于教育学、运筹学、流行病学及心理学(konig & van de Schoot, 2017; Van de Schoot, Winter, Ryan, Zondervan – Zijnenburg, & Depaoli, 2017)等多个领域。其中, Van de Schoot 等(2017)利用系统性评述的方法分析过去 25 年间研究或使用贝叶斯统计方法的论文发表情况, 表明贝叶斯统计在心理学领域的应用率持续增长。但是, 已有评述均基于英文索引论文, 无法反映贝叶斯统计方法在国内心理学研究的应用现状与趋势。

文章采用系统性评述的方式, 量化分析 20 年间(1999 ~ 2019)心理学领域贝叶斯统计方法相关中文文献, 综述贝叶斯方法的研究进展与应用现状。文章的主要目标为: 帮助心理学实证研究者了解贝叶斯统计方法的应用领域与现状; 帮助心理学方法

* 基金项目: 国家自然科学基金青年项目(32100894), 上海市浦江人才项目(2020PJC034)。

通讯作者: 顾昕, E-mail: guxin57@hotmail.com。

研究者掌握贝叶斯统计方法的发展趋势与现实需求;探究目前贝叶斯统计方法应用的不足与原因,为今后贝叶斯统计方法与应用研究方向提供支撑性的信息。文章的研究框架如下。首先,研究将基于心理学中文核心期刊,检索应用贝叶斯统计方法的文献,并依据文献使用的贝叶斯统计方法与研究方向,对其进行多维度的分类,明确贝叶斯方法的应用范围与重点领域。其次,对分类后的文献展开详尽的评述,在重点领域讨论贝叶斯统计方法的具体研究方法与方向。随后,讨论贝叶斯统计方法相关文献在不同期刊的发表情况与趋势。最后根据分析结果,对未来研究方向提供参考性建议。

2 研究方法

文章以系统性评述为研究方法。系统性评述通过对文献检索、分类、再分析等方式,量化某一研究问题或方法在某一领域的发展与应用状况,是近年来综述类文章常用的一种研究方法。

2.1 文献检索

研究选取中国知网作为文献检索平台,聚焦近 20 年心理学领域的贝叶斯统计方法研究。文献检索的年限范围为 1999 年至 2019 年,检索学科领域为心理学。此外,为了确保所检索到的文献质量,研究仅以“中国社会科学引文索引”,即 CSSCI 期刊作为检索期刊。明确检索年限、期刊来源、学科领域后,确定检索词便是研究的关键点。检索词的选取

结合了国外相关研究的检索方式与我国贝叶斯统计方法研究应用实际,最终确定为“贝叶斯”、“后验”、“先验”、“MCMC”(蒙特卡洛马尔可夫链)、“Gibbs”、“WinBUGS”、“OpenBUGS”、“stan”、“JASP”。其中,MCMC 为贝叶斯框架下的参数抽样方法;Gibbs 抽样是 MCMC 算法的特例;WinBUGS、OpenBUGS 是专业贝叶斯统计推断软件;stan 是新兴的贝叶斯数据分析软件;JASP 是可视化贝叶斯统计分析软件,可以提供 APA 格式的结果输出图表。

具体检索步骤如下:(1)使用中国知网高级检索方法;(2)在文献分类目录下,选择哲学与人文社会科学下的心理学分类;(3)将上述检索词以“或含”的方式进行全文检索;(4)检索时间设置为 1999-2019;(5)期刊来源类别选择 CSSCI;共计检索出 1208 篇相关文献。

2.2 文献选择标准

依据检索结果对文献进行分类。在此过程中,作者发现许多文献虽然包括检索词,但与贝叶斯统计方法无关。例如,哲学心理学、现象心理学的文章关键词包含“先验”,但此处的“先验”并非贝叶斯先验;另外,存在文章参考文献中包含检索词,但这些文章本身与贝叶斯统计方法并无关系;此外还发现少量重复文章。在剔除以上几类文献后,得到 366 篇心理学使用贝叶斯统计方法的文献(如图 1)。

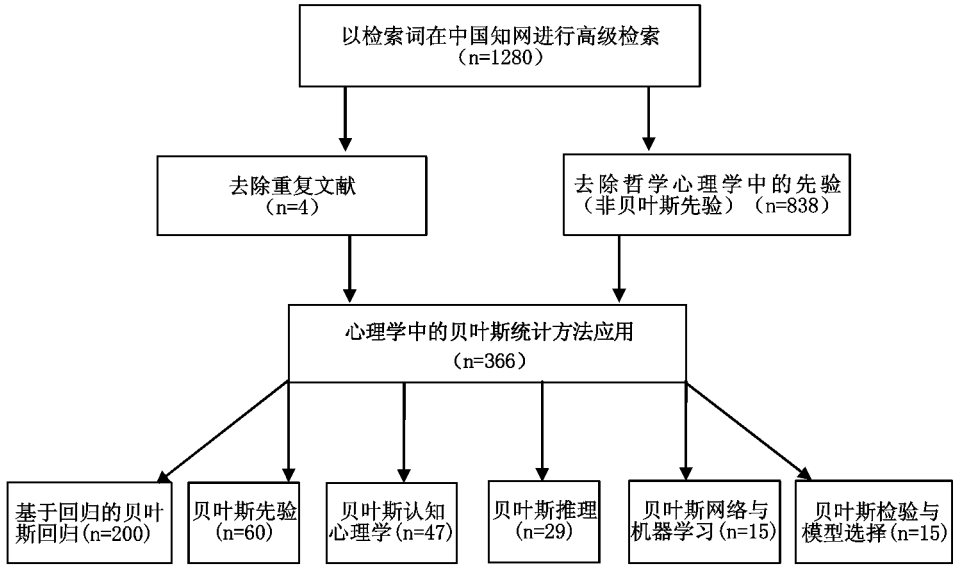


图 1 文献检索及筛选流程图

2.3 文献分类方式

在检索和鉴定符合条件的文章之后,根据贝叶斯方法统计的使用方式与具体领域,来分类识别所

选取的文章,具体依据是贝叶斯研究的方法与类型或者应用的子领域。文章采用系统性评述进行文献分类,具体步骤如下:以 366 篇文献为基础,先由一

人阅读每篇文献中关于贝叶斯统计方法的内容并对其归类,再由第二人进行第二次贝叶斯统计方法相关内容的阅读分类,以双重检查机制来确保研究的客观性,分类的准确性。

2.4 识别贝叶斯统计方法的研究与应用趋势

进行文献分类时,对从文章中提取的信息进行再次分析,以发现贝叶斯统计方法在心理学领域随时间变化的应用趋势。在分析过程中,虽然可以在文献摘要中粗略地识别趋势,但仍需要阅读数百篇文章的全文以充分描述这些趋势。由一人在 Excel 文件中汇总每篇文章的分类与趋势信息,如应用领域(例如,基于回归的贝叶斯、贝叶斯先验、贝叶斯认知心理学、贝叶斯推理、贝叶斯网络与机器学习、贝叶斯检验与模型选择)、研究类型(例如,模拟研究、实证研究、理论/方法研究、综述研究、贝叶斯元分析研究、操作教程)、发表时间、期刊名称。应用领域分类可以明确贝叶斯统计方法的应用范围和比重,帮助研究者了解贝叶斯统计方法在哪些领域已有成熟的应用;研究类型分类可以揭示已有贝叶斯统计方法与应用研究的着重点和薄弱处,为研究者进一步开发或使用贝叶斯统计方法提供支持;发表时间、期刊类别信息可以动态呈现贝叶斯统计方法的应用趋势与各期刊发表数,探索贝叶斯统计方法是否正逐渐被研究者和期刊所接受。

3 研究结果

以检索文献中贝叶斯统计方法的分类结果为基础,从贝叶斯统计方法在心理学领域的具体研究与应用方式出发,分类描述近 20 年来贝叶斯统计方法在国内心理学领域的具体研究方程与应用结果。包括使用的具体方法、研究类型、研究领域等。下文对各领域的具体情况做了详尽分析。

3.1 分类评述结果

表 1 呈现了 366 篇心理学领域贝叶斯统计方法相关文献的分类结果。依据具体使用的贝叶斯统计方法,366 篇文献被分为 6 类:基于回归的贝叶斯统计、贝叶斯先验(先验知识、先验概率)、贝叶斯认知心理学、贝叶斯推理、贝叶斯网络与机器学习、贝叶斯检验与模型选择(见表 1)。其中,基于回归的贝叶斯统计文章最多,达到 200 篇。贝叶斯先验的文章为 60 篇,贝叶斯认知心理学的文章为 47 篇,贝叶斯推理的文章 29 篇,贝叶斯网络与机器学习和贝叶斯检验与模型选择的文章各有 15 篇。此外,部分文章可能属于多个分类,这里由 3 名贝叶斯统计研究

者通过对文献的深度阅读,根据文章主要使用的贝叶斯统计方法,将其归为最接近的分类。

基于回归的贝叶斯统计方法相关文献为 200 篇,占比 54.2%,表明基于回归的贝叶斯统计方法在心理学领域的研究与应用最广泛。值得注意的是,这里的回归模型是一类模型的总称,包括:方差分析、结构方程模型、项目反应理论、计算机自适应测验、认知诊断、多水平模型等子类别。文章 3.2 和 3.3 节将分析上述模型中使用贝叶斯统计方法的具体情况。

表 1 心理学领域贝叶斯统计方法的应用分类

贝叶斯统计方法的分类	N	%
基于回归的贝叶斯统计	200	54.6%
贝叶斯先验(先验知识/先验概率)	60	16.4%
贝叶斯认知心理学	47	12.8%
贝叶斯推理	29	8%
贝叶斯网络与机器学习	15	4.1%
贝叶斯检验与模型选择	15	4.1%

由表 1 可知,使用贝叶斯先验的心理学研究论文共有 60 篇,占比达到 16.3%。这类文章在数据分析中利用了历史数据、专家经验等先验信息。对此类文章进一步分析显示,目前心理学领域的贝叶斯先验相关研究主要体现在先验知识上,共计 53 篇,占贝叶斯先验研究的 84%。此类研究多是基于先验知识、先前经验的问题解决(张锦坤,刘晓倩,杨丽娟,2018)。此外,贝叶斯先验概率的研究有 7 篇,分别应用在不同领域。

表 1 显示贝叶斯统计方法在认知心理学领域共计 47 篇,占比 12.8%。系统性评述结果表明,认知心理学领域的贝叶斯统计方法应用主要分为如下两种:一种将贝叶斯统计方法作为计算模型对认知任务进行建模,主要使用的是贝叶斯优化模型或贝叶斯计算模型来预测和解释推理过程(刘雁伶,黄仁辉,胡竹菁,2015),这类文章共计 33 篇,占认知心理学中贝叶斯统计方法文献的 71.7%。另一种仅将贝叶斯统计模型作为解释人类知觉或推理的模型,这类文章共 14 篇。

贝叶斯推理分类包含研究或应用贝叶斯推理方法的文章,这类文章主要关注贝叶斯推理问题的影响因素,以及贝叶斯推理在心理学实证研究中的应用(杨莉,胡竹菁,2010)。此类文章共有 29 篇,占心理学领域贝叶斯统计方法研究的 7.9%。

贝叶斯网络与机器学习的文章主要是应用贝叶斯网络、机器学习、语言或图像识别方法的文章,共

计 15 篇。其中,4 篇文章将贝叶斯网络应用于认知诊断领域(喻晓锋,丁树良,秦春影,陆云娜,2011)。语言/图像识别的贝叶斯统计方法相关研究 4 篇,机器学习的研究 2 篇。

系统性评述发现,国内心理学领域近十年才有少量关于贝叶斯检验与模型选择的文章,共计 15 篇。其中仅 5 篇文章使用贝叶斯因子检验实证研究中的心理学理论(王效广,胡祥恩,2019), p 值依旧是当今心理学数据分析中检验假设的主要工具。此外,还包括 1 篇贝叶斯因子的介绍与软件应用类论文。贝叶斯统计推断的研究有 7 篇,主要是关于贝叶斯假设检验的评论类文章,讨论贝叶斯假设检验应用的可能性及必要性,但无具体实证研究应用。

综合系统性评述的结果,研究发现基于回归的贝叶斯统计文章最多,应用范围最广,更能反映贝叶斯统计在心理学领域的主要应用方式。因此,下文的 3.2 和 3.3 节对其研究类型、研究方法、应用方式等进行了更为详尽的阐述。

3.2 基于回归的贝叶斯统计方法具体研究方法

基于回归分析的贝叶斯统计方法文章,根据研究方法可分为六种不同的子类型。表 2 中列出了 200 篇贝叶斯回归文章的分类结果。其中,“模拟研究”指通过设定模型与参数真值生成模拟数据,然后对模拟数据进行分析来验证方法的有效性,或比较不同方法的优劣。“实证研究”指在实际数据分析中使用贝叶斯统计方法来回答实质性研究问题的文献。“综述研究”指使用综述的方式介绍某一具体贝叶斯统计方法的文献。“理论/方法研究”指仅使用公式或数学证明的文献,其技术细节较多,针对的是贝叶斯统计方法研究者。“贝叶斯元分析”指使用贝叶斯统计方法开展元分析研究。“操作教程”指贝叶斯统计方法的介绍文献,或者是贝叶斯软件的操作指南,主要针对初步接触贝叶斯统计方法的研究者。基于回归的贝叶斯统计方法应用类别广、文献数量多,梳理并分析不同研究方法的文章,有助于了解不同方法所解决的研究问题。

表 2 基于回归的贝叶斯统计方法的研究类型

基于回归的贝叶斯统计	N	%
模拟研究	107	53.5%
实证研究	42	21.0%
综述研究	41	20.5%
理论/方法研究	6	3.0%
贝叶斯元分析	3	1.5%
操作教程	1	0.5%

模拟研究 这类文章共计 107 篇,占贝叶斯回

归文章的 53.5%,是目前心理学领域贝叶斯回归模型主要的研究方法。其中,使用贝叶斯统计方法进行数据模拟、参数估计、模型比较的文章较多(刘玥,刘红云,2012)。另有部分研究将贝叶斯算法与其他算法进行比较(韩雨婷,高旭亮,汪大勋,蔡艳,涂冬波,2018),或将 MCMC 算法用于缺失值处理。整体来看,认知诊断领域的模拟研究较多,达到 36 篇,占模拟研究 1/3 以上。主要是由于认知诊断的实证研究面临诸多困难,如题目属性的标定。而 MCMC 算法的参数先验分布模拟及后验概率估计解决了这一问题。此外,计算机自适应测验的模拟研究 29 篇,项目反应理论的模拟研究 26 篇,结构方程模型 7 篇,方差分析的模拟研究 7 篇,同时存在少量其他领域的模拟研究。模拟研究中采用 WinBUGS、OpenBUGS 贝叶斯软件的文章仅 6 篇。系统性评述发现,研究者倾向使用 R 语言自编程序开展模拟研究。因此,导致贝叶斯统计软件在模拟研究中应用较少。

实证研究 主要指使用贝叶斯统计数据分析方法来解决实质性问题的研究,共有 42 篇,占有贝叶斯回归研究的 21%。实证研究中,结构方程模型的相关研究最多,达到了 13 篇,大多关于结构方程模型的贝叶斯参数估计(安芹,客文静,2018)。项目反应相关的实证研究共计 12 篇,计算机自适应测验的实证研究有 7 篇,认知诊断实证研究 5 篇,都主要涉及贝叶斯参数估计。此外,实证研究中使用 WinBugs 和 JAGS 进行数据处理的文章仅 2 篇,表明贝叶斯统计软件在心理学实证研究中的应用十分有限。

综述研究 主要指使用综述形式介绍贝叶斯统计方法的文章,共有 41 篇。其中,项目反应理论相关的综述研究 14 篇,结构方程模型相关的综述研究 13 篇。这些文章多是在具体领域的综述中提及贝叶斯统计方法的应用,如项目反应理论的参数估计方法综述中包含贝叶斯参数估计(刘慧,简小珠,张敏强,熊悦欣,2012),结构方程模型的综述中包含贝叶斯参数估计(王婧,唐文清,张敏强,张文怡,郭凯茵,2017)。但也有少量指向特定模型的贝叶斯综述文章(张沥今,陆嘉琦,魏夏琰,潘俊豪,2019)。

理论或方法研究 主要是关于贝叶斯统计方法的理论或方法研究,涉及公式推导或数学证明的文献,共 7 篇。这类文献统计知识较多,主要面向贝叶斯统计方法研究者和心理统计学研究者。其中结构方程模型的研究 4 篇,以贝叶斯估计方法在结构方

程模型里的应用研究为主(方杰,温忠麟,张敏强,2017)。

贝叶斯元分析 指使用贝叶斯元分析方法开展的研究,这类文章仅3篇,无法进行系统性评述。

操作教程 关于如何使用贝叶斯统计方法的详细介绍,或者是贝叶斯软件的操作指南,主要针对初步接触贝叶斯统计的研究者,这类文章仅1篇。缺乏相关统计方法介绍及统计软件的操作指南,可能导致研究者即使想采用贝叶斯统计方法,仍面临技术门槛的问题。

3.3 基于回归的贝叶斯统计方法具体研究方向

基于回归模型的贝叶斯统计方法研究在心理学领域应用最广,且分布在不同的研究方向。图2的研究方向分类结果显示,基于回归模型的贝叶斯统计方法主要分布在8个子领域。其中,项目反应理论相关文章54篇,认知诊断相关文章48篇,计算机自适应测验相关文章40篇,结构方程模型相关文章38篇。此外,方差分析、多层线性模型、MCMC及元分析领域也有少量应用。研究将对具体研究方向内的贝叶斯统计方法进行详细评述。

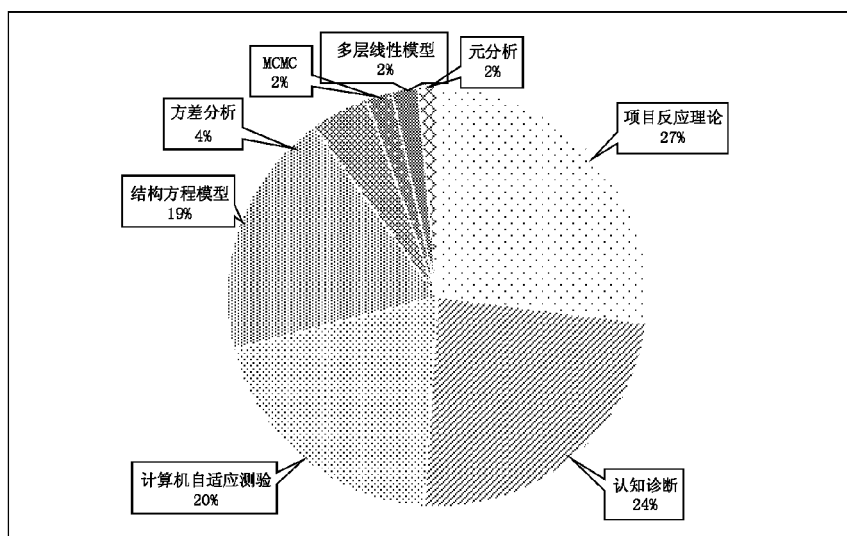


图2 基于回归的贝叶斯统计的研究方向

项目反应理论 系统性评述结果显示,贝叶斯统计方法常用于项目反应理论的数据模拟、参数估计、模型拟合,共计54篇文献。其中,45%的文章使用贝叶斯参数估计方法,包括使用MCMC抽样算法估计被试能力的文章(丁树良,罗芬,戴海琦,朱玮,2007)。比较贝叶斯参数估计方法与其他参数估计方法(如极大似然估计)的研究占比15%(涂冬波,蔡艳,戴海琦,丁树良,2011)。此外还有部分应用后验概率、先验分布开展模拟研究的文章,和少量缺失值处理的文章。其中,项目反应理论的贝叶斯统计方法研究有27篇集中在模拟研究领域,实证研究的文章12篇,贝叶斯统计方法的研究仅有1篇。国内项目反应理论的相关研究主要采取模拟研究这一方式,或许是造成这一现象的原因。

认知诊断 认知诊断作为新一代测量理论的代表,近年来也受到研究者的广泛关注。贝叶斯统计方法用于认知诊断的文献共48篇,主要是运用后验概率、极大后验估计法、期望后验估计法和贝叶斯判别等方法,估计属性掌握模式或对属性掌握模式进

行判别(涂冬波,蔡艳,2015)。同时模拟研究高达35篇,占认知诊断文章的73%,这与目前国内外认知诊断的传统研究方法是一致的。

计算机自适应测验 计算机自适应测验作为一种新兴的测验形式,已应用于Scholastic Aptitude Test (SAT)、Test of English as a Foreign Language (TOEFL)等国际测验中。由于目前计算机自适应测验是基于项目反应理论或者认知诊断来实施的,所以MCMC算法在计算机自适应测验中的应用方式与前两者基本一致。系统性评述结果显示这类研究共有40篇,主要是使用贝叶斯最大后验概率和边际后验概率估计被试参数,区别在于后验加权在计算机自适应测验选题策略中的应用(陈平,李珍,辛涛,2011)。其中30篇集中在模拟研究领域,实证研究领域仅有5篇。

结构方程模型 结构方程模型是分析心理学潜变量间关系最强有力的统计方法。研究检索出38篇使用贝叶斯统计方法对结构方程模型进行参数估计或模型选择的文章,主要涉及变量间的中介效应

和调节效应分析(方杰,张敏强,顾红磊,梁东梅,2014)。此外,有10篇贝叶斯参数估计与传统分析方法的比较研究,2篇MCMC抽样算法用于处理缺失值的研究,2篇MCMC算法用于区间估计和点估计的研究。与上述几个研究方向不同的是,贝叶斯结构方程模型的实证研究(13篇)多于模拟研究(8篇)。可能的原因是结构方程模型已在心理学各领域,如社会心理学、教育心理学、发展心理学等,得到广泛应用。但相对来说,贝叶斯结构方程模型的应用依然较少。

其他 此外,贝叶斯统计方法在方差分析的应用集中在概化理论,即使用MCMC算法分析方差分量,这类研究共6篇,皆为模拟研究,占方差分析研究的67%(黎光明,张敏强,黄宪,王旭,2013)。贝叶斯统计方法在多层线性模型、元分析的文章,以及将MCMC作为实证研究数据分析方法的文章数量较少,无法描绘其应用趋势,故不多赘述。

综合3.2和3.3节对基于回归的贝叶斯统计方法的分析可知,目前模拟研究依旧是贝叶斯统计方法研究的主流,实证研究及理论\方法类的研究较少,贝叶斯统计软件的介绍文章也很少。从具体应用方式上来看,在项目反应理论、认知诊断、计算机自适应测验及结构方程模型等主要测量方向中,以使用贝叶斯统计方法进行参数估计的文章为主,而使用贝叶斯假设检验并运用于实证的研究很少。

3.4 期刊及年限

表3给出了近20年来贝叶斯统计方法相关文章在心理学领域核心期刊的发表数量,同时还展示

了贝叶斯统计方法文章占各期刊总发表数量的比例。由表3可知,《心理学报》的贝叶斯统计方法相关文章占比最高,达到3.4%。但是,《心理科学》的贝叶斯统计方法文章发表数量最多,达到84篇,占比为3.2%。《心理科学进展》和《心理学探新》的贝叶斯统计方法文章发表数量也相对较多,分别为71篇和48篇。其余几个期刊相关论文发表数量较少。

表3 近20年心理学核心期刊的
贝叶斯统计方法论文数量

期刊名称	N	%
心理学报	77	3.4%
心理科学	84	3.2%
心理科学进展	71	2.9%
心理学探新	48	1.9%
应用心理学	13	1.6%
心理与行为研究	9	0.9%
心理发展与教育	9	0.7%
中国临床心理学杂志	9	0.5%
中国心理卫生杂志	1	0.1%

同时,图3根据系统性评述绘制了近20年来贝叶斯统计方法在心理学核心期刊的发表趋势。由图3可知,2010年前九大核心期刊贝叶斯统计方法文章年发表量大多低于5篇,峰值也仅为5篇。2010年后,《心理学报》、《心理科学》及《心理科学进展》的贝叶斯统计方法年发表量迎来一个小高峰,峰值超过10篇。但是其他期刊的贝叶斯统计文章发表数量依旧较低。综上可知,贝叶斯统计方法近10年来逐渐受到心理学研究领域的关注和应用,文章数量开始稳步上涨。

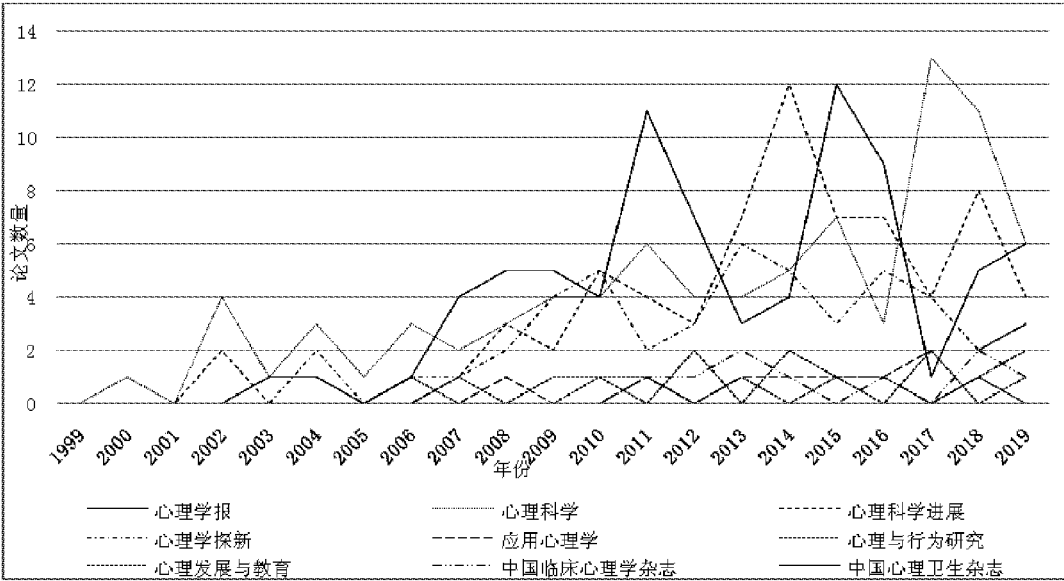


图3 近20年心理学核心期刊贝叶斯统计方法文章的发表趋势

4 总结与展望

研究以系统性评述的方式梳理了贝叶斯统计在心理学领域的应用与现状,对心理学实证与方法研究者有以下几点帮助。首先,文章有助于实证研究者了解贝叶斯统计方法的应用情况,为研究者掌握新方法与新技术提供了帮助。例如,系统性评述发现在项目反应理论、认知诊断、计算机自适应测验、结构方程模型等研究方向已有较多文献讨论与应用贝叶斯统计方法。同时给出了各方向的代表性文献,方便研究者查看贝叶斯统计方法的应用模式与步骤。此外,评述还发现贝叶斯统计方法近十年来的应用比例逐步增加,相关研究也逐渐被主流期刊所接受。这一结果在一定程度上会增加研究者对贝叶斯统计方法的信心。

其次,文章有助于贝叶斯统计方法研究者深入了解实证研究的实际需求。例如,系统性评述表明,贝叶斯文章的研究方法主要是模拟研究,在实证研究中的应用相对较少,说明贝叶斯统计方法的实际应用还有欠缺。笔者认为可能的原因是缺少指导性的综述文章。检索出的综述大多是在具体领域内提及贝叶斯统计方法的应用,针对贝叶斯统计方法本身的综述很少。增加指导性的综述文章能够让实证研究者更全面的了解贝叶斯统计方法的优势与使用步骤等。此外,评述发现贝叶斯统计方法缺少用户友好的软件及相关介绍,大多数文献的贝叶斯分析需在编程软件中实现。开发操作便捷的贝叶斯统计方法分析工具,并提供教程对推广贝叶斯统计方法十分重要。

最后,研究基于文献分析指出了目前贝叶斯统计方法的应用不足。尽管贝叶斯论文发表数量逐年上升,但其在心理学期刊发表总量中的占比仍然较小。相比传统频率统计,许多研究者对贝叶斯统计方法的相关概念(先验、后验、可信区间、贝叶斯因子等)与统计推断步骤还很陌生。究其原因,贝叶斯统计方法的相关知识尚未进入心理统计课程,阻碍了研究者学习和应用贝叶斯统计方法。因此,在心理统计领域普及贝叶斯统计知识对进一步推广贝叶斯统计方法尤为重要。

从国际上心理学研究数据分析方法的发展趋势来看,贝叶斯统计方法逐渐成为主流(van de Schoot,2017)。鉴于贝叶斯统计方法的诸多优势,推进贝叶斯统计方法应用有利于心理学研究的发展。贝叶斯统计方法应当成为心理学研究者手中广

为使用的数据分析工具,帮助他们探索新发现、回答新问题。

参考文献

- 安芹,客文静.(2018).父母教养方式与研究生职业延迟满足:自我分化的中介作用及城乡差异.《中国临床心理学杂志》,26(6),97-101.
- 陈平,李珍,辛涛.(2011).认知诊断计算机化自适应测验的题库使用均匀性初探.《心理与行为研究》,9(2),125-132.
- 丁树良,罗芬,戴海琦,朱玮.(2007).多题多做测验模型及其应用.《心理学报》,39(4),730-736.
- 方杰,温忠麟,张敏强.(2017).类别变量的中介效应分析.《心理科学》,40(2),217-223.
- 方杰,张敏强,顾红磊,梁东梅.(2014).基于不对称区间估计的有调节的中介模型检验.《心理科学进展》,22(10),1660-1668.
- 韩雨婷,高旭亮,汪大勋,蔡艳,涂冬波.(2018).多级评分项目的多维CAT选题策略开发.《心理科学》,41(6),222-229.
- 胡传鹏,孔祥祯,Wagenmakers, E. J., Alexander, L., 彭凯平.(2018).贝叶斯因子及其在JASP中的实现.《心理科学进展》,26(6),951-965.
- 黎光明,张敏强,黄宪,王旭.(2013).概化理论方差分量变量估计方法.《心理学探新》,33(3),239-245.
- 刘雁伶,黄仁辉,胡竹菁.(2015).因果强度推理的贝叶斯模型综述.《心理学探新》,35(5),418-424.
- 刘玥,刘红云.(2012).贝叶斯题组随机效应模型的必要性及影响因素.《心理学报》,44(2),263-275.
- 刘慧,简小珠,张敏强,熊悦欣.(2012).多水平IRT的发展与应用述评.《心理科学进展》,20(4),627-632.
- 涂冬波,蔡艳.(2015).基于属性多级化的认知诊断计算机化自适应测验设计与实现.《心理学报》,47(11),1405-1414.
- 涂冬波,蔡艳,戴海琦,丁树良.(2011).多维项目反应理论:参数估计及其在心理测验中的应用.《心理学报》,43(11),1329-1340.
- 王孟成,邓俏文,毕向阳.(2017).潜变量建模的贝叶斯方法.《心理科学进展》,25(10),1682-1695.
- 王效广,胡祥恩.(2019).简单非一致:以再认启发式使用为例.《心理学探新》,39(2),38-44.
- 王婧,唐文清,张敏强,张文怡,郭凯茵.(2017).多阶段混合增长模型的方法及研究现状.《心理科学进展》,25(10),1696-1704.
- 吴凡,顾全,施壮华,高在峰,沈模卫.(2018).跳出传统假设检验方法的陷阱——贝叶斯因子在心理学研究领域的应用.《应用心理学》,24(3),195-202.
- 杨莉,胡竹菁.(2010).贝叶斯推理中的情感启发式.《心理学

- 探新,30(3),18-21.
- 喻晓峰,丁树良,秦春影,陆云娜.(2011).贝叶斯网在认知诊断属性层级结构确定中的应用.心理学报,43(3),338-346.
- 张沥今,陆嘉琦,魏夏琰,潘俊豪.(2019).贝叶斯结构方程模型及其研究现状.心理科学进展,27(11),1812-1825.
- 张锦坤,刘晓倩,杨丽娴.(2018).背景图相关度与知觉流畅性对意义材料学习的影响.心理与行为研究,16(6),757-762.
- 钟建军,Dienes,Z.,陈中永.(2017).心理研究引入贝叶斯统计推断的必要性、应用思路与领域.心理科学,40(6),1477-1482.
- Clewley,N.,Sherry,Y. C.,& Xiaohui,L.(2011).Mining learning preferences in web-based instruction:Holists vs. serialists. Educational Technology & Society,14(4),266-277.
- Gu,X.,Mulder,J.,Dekovic,M.,& Hoijtink,H.(2014).Bayesian evaluation of inequality constrained hypotheses. Psychological Methods,19(4),511-527.
- König,C.,& vande Schoot,R.(2017).Bayesian statistics in educational research:A look at the current state of affairs. Educational Review,70(4),486-509.
- Mcneish,D.(2016).On using Bayesian methods to address small sample problems. Structural Equation Modeling:A Multidisciplinary Journal,23(5),750-773.
- Scheibehenne,B.,& Pachur,T.(2015).Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. Psychonomic Bulletin & Review,22(2),391-407.
- Sims,C. R.,Neth,H.,Jacobs,R. A.,& Gray,W. D.(2013).Melioration as rational choice:Sequential decision making in uncertain environments. Psychological Review,120(1),139-154.
- vande Schoot,R.,Winter,S. D.,Ryan,O.,Zondervan-Zwijnenburg,M.,& Depaoli,S.(2017).A systematic review of Bayesian articles in psychology:The last 25 years. Psychological Methods,22(2),217-239.
- Wagenmakers,E. J.,Marsman,M.,Jamil,T.,Ly,A.,Verhagen,J.,Love,J.,...& Morey,R. D.(2018).Bayesian inference for psychology. Part I:Theoretical advantages and practical ramifications. Psychonomic Bulletin & Review,25(1),35-57.

Application and Development of Bayesian Statistical Methods

Li Guiyu¹, Gu Xin²

(1. The Institute of Curriculum & Instruction, East China Normal University, Shanghai 200062;

2. Department of Educational Psychology, East China Normal University, Shanghai 200062)

Abstract: Bayesian statistics has become a popular alternative to the traditional frequentist statistical analysis. This paper comprehensively reviews the application and the direction of the Bayesian methods in psychology. The review shows that the Bayesian articles mainly focus on simulation studies, and the application directions are item response theory, cognitive diagnosis, computer adaptive test, and structural equation models. In addition, it is found that Bayesian statistics is gradually accepted by Chinese scholars in psychology. Finally, this paper discusses the limitations and possible reasons of the application of Bayesian statistics in psychological research, and suggests that there is a need of user-friendly Bayesian software and Bayesian knowledge in psychology courses.

Key words: Bayesian statistics; prior; posterior; regression model; systematic review