SPSSAU问卷/量表 数据分析方法与应用

本内容来源于书籍: SPSSAU 科研数据分析方法与应用 已获得出版社和书籍作者内容授权

如需购买纸质书籍,可进入电子工业出版社京东官方旗舰店: https://item.jd.com/10096204208894.html 目录

第1章	SF	PSSAU 平台概述	5
1.1	S	PSSAU 平台简介	5
1.2	S	PSSAU 平台使用	6
1.3	菸	丧得帮助	11
第2章	数	y据探索及分析	14
2.1	5	♪析方法数据格式	14
2.2	挖	采索数据特征	15
	2.2.1	两种数据类型	15
	2.2.2	定类数据探索分析	
	2.2.3	定量数据探索分析	17
	2.2.4	小结	19
2.3	娄	牧据分布之正态性分析	20
	2.3.1	正态分布图示法	20
	2.3.2	正态分布检验法	22
	2.3.3	正态分布转换处理	23
	2.3.4	小结	24
2.4		常用分析方法选择	24
	2.4.1	定类或定量数据分析方法	24
	2.4.2	定类和定类数据分析方法	26
	2.4.3	定类和定量数据分析方法	27
	2.4.4	定量和定量数据分析方法	29
	2.4.5	小结	
第3章) []]卷数据清理	
3.1	娄	牧据标签设置	
3.2	娄	牧据编码	
3.3	톳	异常值处理	35
3.4	生	E成变量	36
	3.4.1	常用处理	37
	3.4.2	量纲处理	
	3.4.3	科学计算	40
	3.4.4	汇总处理	41
	3.4.5	日期相关处理	41
	3.4.6	其他	
3.5	枋	示题处理	43
第4章) []]卷数据差异关系研究	44
4.1	tz	检验	
	4.1.1	正态分布与方差齐性	45
	4.1.2	t 检验分析步骤	
	4.1.3	单样本 t 检验	47
	4.1.4	配对样本 t 检验	
	4.1.5	独立样本 t 检验	
4.2	ブ	方差分析	50

	4.2.1	方法概述	50
	4.2.2	方差分析类型的选择	
	4.2.3	单因素方差分析	53
	4.2.4	双因素及多因素方差分析	57
	4.2.5	简单效应分析	
4.3	-	卡方检验	61
	4.3.1	方法概述	61
	4.3.2	2×2 四格表卡方检验	
	4.3.3	R×C 列联表卡方检验与多重比较	
	4.3.4	fisher 卡方检验	67
	4.3.5	配对卡方检验	
	4.3.6	分层卡方检验	70
	4.3.7	卡方拟合优度检验	73
4.4	Ξ	非参数检验	74
	4.4.1	方法介绍	74
	4.4.2	单样本 Wilcoxon 检验	75
	4.4.3	两组独立样本 Mann-Whitney 检验	76
	4.4.4	多组独立样本 Kruskal-Wallis 检验	77
	4.4.5	配对样本 Wilcoxon 检验	79
	4.4.6	多样本 Friedman 检验	
第5章	Ì	司卷数据相关影响关系研究	
5.1	7	相关分析	
	5.1.1	相关关系概述	
	5.1.2	相关分析步骤	
	5.1.3	两个变量相关实例分析	
	5.1.4	偏相关实例分析	
5.2	4	线性回归	
	5.2.1	线性回归模型与检验	
	5.2.2	线性回归适用条件	
	5.2.3	线性回归的一般步骤	
	5.2.4	多重线性回归的实例分析	
	5.2.5	逐步线性回归的实例分析	
	5.2.6	有哑变量的线性回归	
5.3	L	Logistic 回归	
	5.3.1	方法概述	
	5.3.2	二元 Logistic 回归	
	5.3.3	多分类 Logistic 回归	
	5.3.4	有序 Logistic 回归	
第6章	Ì	可卷专属研究分析方法	
6.1	自	单选题与多选题分析	
	6.1.1	分析思路	
	6.1.2	频数统计实例分析	
	6.1.3	卡方检验实例分析	
6.2	圢	真空题分析	

		6.2.1	分析思路	
		6.2.2	实例分析	
	6.3	项	目分析	131
		6.3.1	原理介绍	131
		6.3.2	实例分析	
	6.4	效	度分析	135
		6.4.1	结构效度	135
		6.4.2	实例分析	136
	6.5	信息	度分析	138
		6.5.1	信度系数	138
		6.5.2	实例分析	139
	6.6	验记	证性因子分析	141
		6.6.1	方法概述	141
		6.6.2	验证性因子分析步骤	142
		6.6.3	验证性因子分析实例分析	146
	6.7	路征	经分析	151
		6.7.1	方法概述	152
		6.7.2	实例分析	153
	6.8	结相	构方程模型	157
		6.8.1	方法概述	157
		6.8.2	实例分析	159
		6.8.3	结构方程模型分析讨论	164
	6.9	中1	介效应分析	164
		6.9.1	中介变量与中介效应	165
		6.9.2	中介效应检验流程与实例	
		6.9.3	多重中介效应分析与实例	170
	6.10) 调1	节效应分析	172
		6.10.1	调节变量与调节效应	172
		6.10.2	简单斜率与斜率图	174
		6.10.3	调节效应分析步骤与实例	175
	6.11	. 有词	周节的中介分析	178
		6.11.1	方法概述	178
		6.11.2	有调节的中介作用实例	
参考	6文南	ŧ		

第1章 SPSSAU 平台概述

1.1 SPSSAU 平台简介

SPSSAU (Statistical Product and Service Software Automatically, 自动化统计产品和服 务软件) 是一款数据科学分析平台软件(官网: https://spssau.com), 其作为一款图形菜 单驱动的数据科学软件, 界面极其简单、友好, 输出结果全部采用规范化三线表格格式, 并且能提供分析建议, 并进行智能分析。SPSSAU 平台于 2016 年 6 月 18 日上线, 历经多 版本迭代, 截至本书出版时, 其版本号为 SPSSAU 23.0, 当前 SPSSSAU 平台已广泛应用于 自然科学、技术科学和社会科学的各个领域, 包括教育、师范、心理、医学、管理、经济 金融等。

SPSSAU 平台当前包括十三大模块,分别是通用方法、问卷研究、可视化、数据处理、 进阶方法、实验/医学研究、综合评价、计量经济研究、机器学习、 Meta 荟萃分析、空间 计量、文本分析和 Power 功效分析模块。其当

前可提供包括描述性统计、假设检验、聚类分析、相关回归分析、信度效度分析、数 据可视化、生存分析、综合评价、时间序列分析、Logistic 回归、曲线回归、计量经济模型、 机器学习等约 500 类算法。与此同时,其具有如图 1-1 所示的八大特点。



图 1-1 SPSSAU 特点

1.在线使用

传统统计软件,如 SPSS、SAS、R 等需要先下载再安装在电脑上,但 SPSSAU 平台为 在线使用(官网: https://spssau.com),不存在下载和安装软件的问题,作为 SaaS 服务软件,只需要在其平台上注册一个账号即可立即使用,且分析结果均保存在云端,在任意电脑和浏览器上均可使用。

2.智能分析

如果不理解各类分析指标的含义,以及分析方法的流程和步骤,那么 SPSSAU 平台能 提供分析建议和智能分析,协助理解各类指标的含义,并提供分析结果以解读建议,此功 能对于初学者较友好。

3.简单易用

上手一款专业数据分析软件并非易事,需要有一定的前期学习成本,但 SPSSAU 平台 上手非常简单,在操作时仅需搜索分析方法、拖曳分析项,便开始分析并得到结果,总共 仅三个步骤。该平台已具有 500 类分析检验和分析方法,但并不需要研究者逐一去查找, 仅需在"搜索方法"框中输入使用的分析方法,其支持对分析方法名称的全称、简称或拼音 等多种方法搜索,快速定位需要使用的分析方法。

4.服务详尽

多数情况下,研究人员需要理解各类指标的原理和意义,只有这样才能进行科学的分析,如果对于个别指标的含义有疑问,则可以在 SPSSAU 平台服务系统中寻找答案,其服务系统包括三类形式,分别是智能客服系统、帮助手册系统和人工服务系统,通过智能客服系统和帮助手册系统,可自助查阅资料,也可通过人工服务系统提交问题寻求帮助。除此之外,还可通过案例库系统找到分析灵感。

5.完备的帮助手册

在 SPSSAU 平台帮助手册中,能提供视频和文字两种形式,视频包括分析方法的解读 和疑难困惑问题,文字部分能提供原理解析、案例解读和疑难解惑三部分内容。在其帮助 手册中,可找到各类研究应用的说明和事例,同时能提供视频文案。另外,其帮助手册中 还能提供"数据分析知识"和"研究思考"等内容,研究人员可在其中找到分析帮助。帮助手册 的使用简单,可通过搜索关键词找到内容。除此之外,帮助手册可在 SPSSAU 微信小程序 和手机上进行查阅。

6.认可度高

通过中国知网搜索(搜索关键词为"SPSSAU"),结果显示已有超 6000 个引用文献,并 且包括较多知名期刊收录文献,如 SCI、EI、北大核心、CSSCI、CSCD 和 AMI 等。

7.快速更新

区别于传统统计软件, SPSSAU 平台具有快速迭代更新的特点, 通过其版本迭代历史 页面说明可查阅迭代更新记录, 并且与作者实际体验表现一致。

8.标准的输出格式

在科研写作过程中,通常需要将重要指标信息放置于同一表格中,并且使用三线表格进行规范,在 SPSSAU 平台中默认为三线表格结果,可实现一键复制和下载导出结果,以节约三线表格的制作时间,并且提供多类规范化格式供用户选择及使用。

1.2 SPSSAU 平台使用

SPSSAU 平台使用简单,首先上传数据、选择方法或功能、拖曳分析项(包括设置参数值),然后单击【开始分析】按钮即可得到分析结果,最后将分析结果导出,如图 1-2 所示。



图 1-2 SPSSAU 平台使用流程

1.进入 SPSSAU 平台

Q 搜索方法	×	筛选样本 100 ❷ € ♀ ♀	开始分析 📄 🔽
凸 通用	防法 🗸	【系统】网购满意度_定量	
频数	分类汇总	【系统】网购忠诚度_定量	
描述	交叉(卡方)	【系统】网购频率_定量	
相关	结性回归	【系统】网购金额_定量	
	**	【系统】退货次数_定量	
7左	1023H	【系统】投诉次数_定量	
単 年本 1 检验	配灯标验验	【系统】平台偏好_定类	
正态性检验	非参数检验	【系统】快递偏好_定类	
1月 问券	新売、	【系统】性别_定类	
		【系统】城市_定类	
2 可	视化 🗸	【系統】职业_定类	e
目 数据	動理 ✔		分析项(定类)

6 / 187

图 1-3 SPSSAU 平台界面

首先通过 SPSSAU 网址登录进入平台, SPSSAU 平台界面如图 1-3 所示, 其默认提供 系统数据,可对该数据直接操作并使用,但通常数据会自行上传。

SPSSAU 平台界面左侧为分析"仪表盘",其包括平台的十三大模块(各类分析方法)或 功能按钮;中间部分为"标题框",其能放置某"数据文档"的所有标题;右侧部分为"分析框", 将"标题框"中的标题拖曳到右侧"分析框"中即可进行分析。

2.上传数据

以上传研究者自己的数据为例,单击平台右上角的【上传数据】按钮后进入上传数据 界面,如图 1-4 所示。



图 1-4 上传数据界面

针对数据上传格式, SPSSAU 平台当前支持 Excel 数据格式、SPSS 软件数据格式、 Stata 软件数据格式和 SAS 软件数据格式,具体数据文件后缀分别为 xls、xlsx、csv、sav、 dta、sas7bdat 等。如果是 Excel 数据格式,需要确保数据不能出现合并单元格,第1行为 标题不能出现缺失,而且需要将数据放入 Excel 的第1个工作表 sheet 中。Excel 数据格式 如图 1-5 所示。

表格示例 Q										
第1列	第2列	第3列	第4列	第5列	第6列					
编号	性别	年齡	学历	年收入	日期					
1	1	3	2	2	2021/8/12					
2	1	2	3	2	2021/8/13					
3	2	2	4	2	2021/8/14					
4	2	1	1	3	2021/8/15					

图 1-5 Excel 数据格式

如果是 SPSS 软件数据格式、Stata 软件数据格式和 SAS 软件数据格式, SPSSAU 平台 会自动将其标题和数据标签信息等进行处理并解析至 SPSSAU 平台中。

操作时可通过单击"上传文件"或拖曳"上传数据"上传,上传成功后可直接进入 SPSSAU 平台界面中。上传成功后的数据文件,可在 SPSSAU 平台右上角"我的数据"中找到并进行 管理。数据文档管理界面如图 1-6 所示。

SPSSAU 科研数据分析方法与应用

2023-(-19	Ø	第7章	立即分析	查看	备份	下载	删除	分享
2022-	-17	Ø	新建 Microsoft Excel 工作表 (2)	立即分析	查看	备份	下载	删除	分享
2022-	-12	Ø	路径分析	立即分析	查看	备份	下载	删除	分享
2022-	-12	Ø	结构方程模型SEM	立即分析	查看	备份	下载	删除	分享
2022-	-12	Ø	rwg	立即分析	查看	备份	下载	删除	分享
2022-	-12	Ø	markov	立即分析	查看	备份	下载	删除	分享
2022-	-12	Ø	CFA	立即分析	查看	备份	下载	删除	分享
2022-	-30	Ø	rdddata(1)	立即分析	查看	备份	下载	删除	分享
2022-	-30	Ø	deming	立即分析	查看	备份	下载	删除	分享
2022-	-30		laverchi(1)	立即分析	杏石	备份	下载	删除	分享

图 1-6 数据文档管理界面

在"数据文档管理界面"中,可查看研究者的所有数据文档,包括上传时间和文档名称, 并且可对文档名称进行修改。数据文档管理包括立即分析、查看、备份、下载、删除和分 享,具体说明如表 1-1 所示。

功能	说明
立即分析	切换数据分析文档,如当前在分析A 文档,切换成分析B 文档
查看	查看原始数据,查看该文档的具体数据
备份	备份数据文档,将该文档备份成一份新文档
下载	下载数据文档,下载成 Excel 数据格式
删除	删除数据文档
分享	分享数据文档,如A研究者将某文档分享给B研究者使用

表 1-1 数据文档管理说明

【立即分析】按钮介绍:如果当前正对"A 文档"进行分析,可单击【立即分析】按钮 切换到对"B 文档"进行分析,即告诉 SPSSAU 平台当前需要对"B 文档"进行分析。

第1列	第2列	第3列	第4列	第5列	第6列
ID	性别	年龄	工作年限	A1休假制度	A2资金制
1	1	2	3	3	4
2	1	3	3	4	4
3	1	3	4	5	5
4	1	4	4	2	2
5	1	4	4	2	2
6	1	3	4	2	2
7	1	4	4	5	5
8	1	2	4	4	4
9	1	3	4	4	4
10	1	3	4	3	4
11	1	3	4	2	4
12	1	3	4	4	4
13	1	3	4	5	4
14	1	3	3	4	4
15	1	4	4	2	2
16	1	4	4	2	1
17	1	4	4	2	2
18	1	3	3	2	2
19	1	2	2	1	1
20	1	4	4	3	2

图 1-7 原始数据查看

单击【查看】按钮,可分页查看原始数据,并可通过单击【筛选查看】按钮设置及查 看某部分的行数据,单击【列表显示】按钮可设置及查看某部分的列数据,单击【文字格 式】按钮可查看带数据标签的数据信息。原始数据查看如图 1-7 所示。

单击【备份】按钮,可将该文档数据备份,单击【下载】按钮可将数据文档导出在本 地电脑上,单击【删除】按钮可将数据文档删除。最后,单击【分享】按钮会弹出【分享 数据文档】对话框,确认后会得到一个"分享链接",如图 1-8 所示。

分享数据文档	
□ 设置密码	
spssau.com/spssaudata.html?sharel	Data=!
确认并复制错	连接

将该"分享链接"粘贴至浏览器框后单击回车, "数据文档管理界面"中会出现该份数据 文档, "分享"功能用于快速将自己的数据分享给其他 SPSSAU 平台 用户。

3.选择方法或功能

Q_相关	×	筛选样本 100 🕢 📢 👰
查看更多"相关"相关文档 线性回归		相关 位于[通用方法]模块
相关 因子		研究定量数据之间的相关关系情况
ICC组内相关系数 偏相关	- -	

图 1-9 搜索分析方法或功能

将数据上传后,则默认进入了"SPSSAU 平台界面",在"SPSSAU 平台界面"研究者可按 需要选择分析方法或功能。如果研究者希望进行相关分析,那么可输入该方法的全称或简称,也可输入英文名称、拼音等,均可搜索到对应的分析方法,搜索"相关"二字,平台会 匹配可能的分析方法或功能,鼠标移动到某方法处会自动提示该文档所处的模块和概要说 明。搜索分析方法或功能如图 1-9 所示,单击该方法或功能,即可进入并使用。

筛选样本 200 @	<; Ω		开始分析 Pearson相关系数 V Q
【系统】网购频率_定量			【系统】网购满意度_定量 Spearman相关系数
【系统】网购金额_定量			【系统】网购忠诚度_定量 Kendall相关系数
【系统】退货次数_定量			
【系统】投诉次数_定量			
【系統】平台偏好_定类			
【系统】快递偏好_定类			分析版(字号)
【系统】性别_定类			7.3.101.292 (A.C.1002)
【系统】城市_定类		0	
【系統】职业_定类		Ð	

图 1-10 拖曳分析项及参数设置

图 1-8 【分享数据文档】对话框

4.拖曳分析项

选择好分析方法或功能后,结合研究目的,拖曳分析标题至右侧的【分析框】中,单 击【开始分析】按钮,会得到分析结果。拖曳分析项及参数设置如图 1-10 所示。

拖曳操作时,支持快捷键组合操作,分别说明:按住 Ctrl 键不会连续选择多个标题, 按住 Shift 键会连续选择多个标题。如果不按 Ctrl 或 Shift 键,一次仅能拖曳一个标题。将 标题拖曳至右侧的【分析框】后,可结合研究需要,通过【开始分析】按钮右侧的下拉列 表设置相应的参数值。例如,图 1-10 中可通过下拉列表选择【Spearman 相关系数】或 【Kendall 相关系数】等。单击【开始分析】按钮即可得到分析结果。

5.分析结果

分析结果中可能包括表格、可视化图等,并且在表格或可视化图下方会展示"分析建议" 和"智能分析",可通过单击【复制】按钮将内容进行复制并使用。分析结果界面如图 1-11 所示。

区 相关 x				分析结果列表				
导出EXCEL表格导出PDF结果	导	出Word结果	导出所有图形 分字图片	分享結果 🖺 🔉 5				
		Pears	son相关-标准格式 🗉	Â				
	平均值	标准差	【系统】 网购满意度_定量	【系统】网购忠诚度_定量				
【系统】 网购满意度_定量	3.480	1.329	1					
【系统】网购忠诚度_定量	3.510	1.040	0.808**	1				
 分析建议 目 相关分析用于研究定量数据之间的关系情况,是否有关系,关系紧密程度情况等: 第一: 首先看YS为是否有显着关系: 第二: 該者分析相关关系为正向或负向,也可通过相关系数大小说明关系紧密程度: 第三: 对分析进行总结。 相关分析之前,可使用散点图观察、展示数据间的关联关系情况,也可使用正态图观察、展示数据正态性分布情况。 								
智能分析 目 从上表可知,利用相关分析去研究【系统】网购满意度_定量和【系统】网购忠诚度_定量之间的相关关系,使用Pearson相关系数去表示相关关系的强弱情况,具体分析可知:								

图 1-11 分析结果界面

图 1-11 的"分析结果列表"中展示了所有的分析结果, SPSSAU 平台会自动对分析结果 命名, 命名方式为"分析方法+编号", 编号会随着分析次数的增加而自动增加, 研究者可自 行修改某分析结果的名称(但编号不能修改), 并且可对其进行删除处理。

线性回归 78	×	正态性检验 77	×	正态性检验 76	×	正态性检验 75	×	正态性检验 74	×	\odot
正态性检验 73	×	单样本地检验 72	×	描述 71	×	描述 70	×	相关 69	×	\odot
相关 66	×	单样本地验 65	×	单样本t检验 64	×	单样本地验 63	×	单样本地验 62	×	Î
频数 61	×	相关 60	×	描述 59	×	描述 58	×	描述 57	×	
描述 56	×	描述 55	×	相关 54	×	相关 53	×	相关 52	×	
描述 51	×	描述 50	×	描述 28	×	描述 27	×	相关 26	×	
realdate										

图 1-12 分析结果列表

与此同时,当分析结果大于 30 个时, SPSSAU 平台会提供搜索历史结果功能,选择日 10 / 187 期搜索历史分析结果,操作共分为两步,第一步是单击【分析结果列表】的【…】图标, 第二步是搜索历史分析结果。分析结果列表如图 1-12 所示。

双击选择某日期, SPSSAU 平台会自动搜索该日期前后 7 天的分析结果,并且以列表 形式展示,如果该日期前后 7 天没有分析结果,则系统会提示"该日期跨度内暂无分析结 果"。分析结果搜索界面如图 1-13 所示。

QÉ	-月	-日				
2023年01月・ ↑ ↓						
_	Ξ	Ξ	四	五	六	日
26	27	28	29	30	31	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5
清除 今天						
站档力柱模型 1020						
结构方程模型 1019						
结构方程模型 1018						

分享数据分析结果
□ 设置密码
https://spssau.com/shareresults.html?shar
确认并复制链接

图 1-13 分析结果搜索界面

图 1-14 分享数据分析结果

6.导出分析结果

SPSSAU 平台能提供 7 种分析结果导出方式,分别是"复制结果"、"导出 Excel 表格"、 "导出 PDF 结果"、"导出 Word 结果"、"导出所有图形"、"分享图片"和"分享结果"。

"复制结果"可通过单击分析结果中表格标题右侧的【复制】按钮,一键进行内容复制, 该方式可将分析结果以 Excel、PDF 或者 Word 格式进行下载并导出。另外"导出所有图形" 指将分析结果中多张图批量导出到本地。"分享图片"指将分析结果存储为云端图片,并且 得到分享链接,使用该分享链接可查看图片化的分析结果。"分享结果"指平台会生成分享 链接,并可设置密码,打开该链接即可查看所有的分析结果。分享数据分析结果如图 1-14 所示。

1.3 获得帮助

数据分析过程中总会遇到各类问题,如出现上传数据提示出错、某指标的意义不能彻底理解和应用等。当遇到问题时,可通过 SPSSAU 平台的各类服务找到答案,分别是帮助手册、案例库、智能客服和人工客服。获得 SPSSAU 平台帮助如图 1-15 所示。

上述几类获得帮助的方式,可通过"SPSSAU 平台界面"上方的"客服中心"找到入口。 SPSSAU 平台客服中心如图 1-16 所示。



图 1-15 获得 SPSSAU 平台帮助 图 1-16 SPSSAU 平台客服中心

SPSSAU 平台帮助手册很强大,不仅能提供各类分析方法的帮助手册,而且能提供各类数据分析知识的内容手册等,其帮助手册目录如图 1-17 所示。

快速上手 上传数据 基本概念 分析思路模板 分析方法选择 通用方法 问卷研究 可视化 数据处理
通用方法 问卷研究 可视化 数据处理
 问卷研究 可视化 数据处理
可视化 数据处理
数据处理
`+=™- → `+
进阶方法
实验/医学研究
综合评价
计量经济研究
账号权益等
数据研究报告
SPSSAU学习
数据分析知识
研究思考
SPSSAU版本等
其他文档

图 1-17 SPSSAU 平台帮助手册目录

针对具体方法, SPSSAU 平台帮助手册能提供教学视频, 包括案例视频及疑难困惑类的内容视频, 并且能提供理论解析和案例解读、疑难解惑知识点等内容。SPSSAU 平台帮助手册如图 1-18 所示。

请输入搜索关键字				
因子分析				
▶ 分析方法视频解读: B站 优酷				
▶ <u>B站 优</u> 酷 因子分析的几种应用				
▶ <u>B站 优酷</u> 因子分析的分析步骤				
▶ B站 优酷 因子分析计算权重				
Disk 优酷 因子分析分析综合竞争力水平				
▶ <u>B站 优</u> 器 因子分析进行共同方法偏差				
案例数据下载 <u>下载</u> 数据格式说明 查看				
因子分析(探索性因子分析)用于探索分析项(定量数据)应该分成几个因子(变量),比如20个量表题项应该分成几个方面较为 合适,用户可自行设置因子个数如果不设置,系统会以特征根值大于1作为判定标准设定因子个数。 因子分析通常有三个步骤,第一步是判断是否适合进行因子分析,第二步是因子与题项对应关系判断,第三步是因子命名。				

图 1-18 SPSSAU 平台帮助手册

除此之外,用户还可在 SPSSAU 平台案例库中找到较多学习案例,或通过 SPSSAU 平台智能客服和人工客服寻求帮助,智能客服涵盖了平台中涉及专业知识和平台本身的问答内容,如果智能客服无法解决问题,则可通过人工客服(人工专业咨询)寻求答案。

第2章 数据探索及分析

在正式进行数据分析之前,通常需要对数据进行探索,简而言之即事先需要对数据的 基本情况有一定了解。例如,使用二元 Logistic 回归研究 X 对 Y 的影响时,需要了解 Y 值 是否为二分类数据、Y 值是否仅包括数字 0 和数字 1,或者 Y 值的分布情况是否有异常等。 当使用线性回归研究 X 对 Y 的影响时,Y 值的分布情况是否符合正态分布特征、在线性回 归分析之后对残差值进行正态性检验,诸如此类均需要对数据特征或数据分布情况进行探 索。

如果已经确定好使用的分析方法,那么要准备好正确的数据格式,以防出现后续无法 进行分析的尴尬,如重复测量方差分析的数据格式要求很严格,需要事先整理好分析软件 要求的格式。基于上述说明,本章分别从以下四个方面进行说明,分别为分析方法数据格 式、探索数据特征、数据分布之正态性分析,以及常用分析方法选择。

2.1 分析方法数据格式

数据格式是分析方法的基石,只有先准备好正确的数据格式才能完成分析,绝大多数 情况下数据格式中的 1 行代表 1 份数据, 1 列代表 1 个属性。例如,中国综合社会调查 CGSS 数据时,1 行代表 1 个被试,被试的特征称为属性,1 列代表 1 个属性,如性别、收 入情况和学历情况等,本书称其为"普通数据格式",常见且易懂。

除此之外,某些分析方法还会使用"加权数据格式",如总共有150个被试,男性为50个,女性为100个,如果使用"普通数据格式"表示,那么需要150行和1列进行标 识。如果使用"加权数据格式"表示,那么仅需要2行1列即可。"加权数据格式"事例1 如表2-1所示,1列为属性,另外1列为"加权项","加权项"能标识对应属性值的个数。

	加权项
男	50
女	100

表 2-1 ""加权数据格式"事例 1

有一项抽烟与是否患癌症之间关系的研究,共有 1920 个研究被试,抽烟共有两个属 性,分别是抽烟和不抽烟,是否患癌症有两个属性,分别是患癌症和身体正常,两个属性 均包括两个属性值,那么 2×2=4 个组合,因而可使用 4 行,两个属性分别单独 1 列,"加 权项"1 列共计 3 列进行数据标识。诸如此类均为"加权数据格式"。"加权数据格式"事例 2 如表 2-2 所示。

表 2-2 "加权数据格式" 事例 2

抽烟	是否患癌症	加权项
抽烟	患癌症	15
抽烟	身体正常	900
不抽烟	患癌症	5
不抽烟	身体正常	1000

综合上述说明,针对"普通数据格式"和"加权数据格式"进行汇总说明,如表 2-3 所示。

表 2-3 两种数据格式

类型	说明
普通数据格式	原始数据,适用于绝大多数分析方法
加权数据格式	汇总数据,针对某些分析方法

"普通数据格式"能记录原始数据,适用于绝大多数分析方法,并且简单易懂;但在指 定特定的分析方法时,可能出于数据简化、原始数据丢失或分析要求等原因,因而可能会 使用"加权数据格式"。"普通数据格式"携带信息量全面,可以转换成"加权数据格式",但是 无法把"加权数据格式"转换成"普通数据格式",即原始数据。

"加权数据格式"通常针对定类数据(定类数据是"性别"这样的类别数据,下述 2.2.1 节 有详细说明),只有定类数据才能使用个数来计量,其行数等于各个属性对应的属性值个数 相乘,如有 3 个属性,分别对应 3、4、5 个属性值,那么行数为 3×4×5=60 行,其列数等 于属性个数+1,1 为"加权项"列,如 3 个属性使用 4 列数据标识即可。"加权项"表示各个 属性值组合对应的样本个数。

"加权数据格式"通常仅适用于部分分析方法,该类分析方法通常针对定类数据进行分析,如卡方检验、Kappa 一致性检验、配对卡方、分层卡方、Poisson 回归、卡方拟合优度检验、Poisson 检验、对应分析等。

除上述的"普通数据格式"和"加权数据格式"外,某些分析方法还有特殊的数据格式要求, 如配对 t 检验、重复测量方差等,此类分析方法针对特殊的分析数据,因而数据格式较特 殊,此处不再详述,读者可参阅 SPSSAU 平台帮助手册中"常见研究方法数据格式说明"。

2.2 探索数据特征

正确的数据格式是完成科学分析的第一步,本节将从数据特征角度进行说明,探索已 有数据的基本特征情况,为后续深入分析做好准备,这也是科学分析的常见步骤,本节分 别从两种数据类型及其分别的探索分析情况进行说明,并且针对定类数据和定类数据、定 类数据和定量数据、定量数据和定量数据如何进行探索分析进行说明,并在最后总结。

2.2.1 两种数据类型

数据类型的分类标准较多,可分成分类数据和连续数据两类,也可分成名义数据、度 量数据和有序数据三类,本书按 SPSSAU 平台规则,将数据分为两类,分别是定类数据和 定量数据,如表 2-4 所示。

术语	说明	举例
定类数据	数字大小代表分类	性别(男和女),专业(文科、理科、工科)
定量数据	数字大小具有比较意义	GDP 数据,身高

表 2-4 两类数据类型

定类数据和定量数据的区别在于数字大小是否具有比较意义,如1 代表"男",2 代表 "女",数字1和数字2 仅为区分类别(提示:计算机存储数据时,其只会将数据存储为数 字,但是数字的实际意义,则由数据标签进行标识,见 3.1 节内容),而不能理解为"数值 越大越女性",因此性别为定类数据。例如,GDP 数据,数字越大表示 GDP 越高,则其为 定量数据。 有时候会出现定类数据和定量数据模糊不清的情况,如收入分组数据,其分为 4 项, 1 代表"5000 元以下",2 代表"5000 元~1 万元",3 代表"1 万元~2 万元",4 代表"2 万元 以上"。此数据可理解为数值越大收入水平越高,但也可看作定类数据,相当于将样本人群 分成 4 个不同的类别。此种情况可结合使用场景和分析方法等综合判定数据为定类数据还 是定量数据。

针对上述收入分组数据,如果希望查看详细信息,则将收入分组数据看成定类数据, 计算各个收入类别的百分比进行分析;如果仅希望使用一个平均值表示整体收入情况,那 么可将收入分组数据看成定量数据。在进行线性回归分析时,收入分组数据通常会被看成 定量数据,其可便捷地分析,如"收入越高越如何"这类关系,如果将收入分组数据看成定 类数据进行线性回归分析,那么其需要进行哑变量处理(见 2.4 节内容)后放入模型,此 时会得到具体各项的详细分析信息,与此同时也会让分析变得更复杂。

数据类型非常重要,其与 2.1 节所述的数据格式有紧密关系,并且数据类型会影响甚 至决定数据探索的分析思路,以及本书后续章节中分析方法的应用等。接下来会分别针对 定类数据和定量数据如何探索分析进行阐述。

2.2.2 定类数据探索分析

在数据分析前根据需要对数据进行探索分析,其目的在于了解当前数据。通常情况下, 会批量式对所有定类数据进行探索以了解数据的基本特征情况,并且探索数据分布情况是 否正常,如果异常则需要进一步处理。针对定类数据,可分成以下几个方面进行探索分析, 包括数据标签、频数分析、描述分析、可视化图和数据编码共 5 项,如表 2-5 所示。

项	分析或处理	作用
1	数据标签	标识数字代表的意义
2	频数分析	探索各个类别的频数分布情况等
3	描述分析	查看数据是否存在异常值等
4	可视化图	使用可视化图展示各类别分布
5	数据编码	针对分布不均的类别进行组合,成为新的类别

表 2-5 定类数据探索分析

在计算机的世界里,其只认识数字信息,但是定类数据的数字有实际意义,如男和女, 计算机会将男存储为数字,并且使用标签标识,如数字1代表男,数字2代表女。那么针 对定类数据,其第一步即进行数据标签,即标识数字代表的意义,具体内容详见3.1节。

通过频数分析可以了解数据是否存在缺失,如研究数据共 100 个,但是当某个数据出 现缺失时,有效样本会变少。与此同时,可以通过频数分析了解数据的分布情况,如男和 女的分布情况、各项比例是多少。我们还可以使用描述分析了解数据的最小值或最大值情 况,如性别,理论上只有数字 1 和数字 2,那么描述分析时,如果出现最小值为 0,或最 大值为 3 时,此时需要查看原始数据是否出现问题。如果原始数据出现了问题,则需要检 查具体原因(包括是否原始数据出错,或存在异常值),如果原始数据出错,则处理正确后 再进行分析即可;如果存在异常值,则可进一步对异常值进行处理(见 3.3 节内容)。

另外,还可使用一些可视化图,如饼图、柱形图等直观地展示各个类别的分布情况, 其仅是频数分析的另外一种展示效果,更加直观。如果发现某个类别的样本量很少,如专 业分为文科、理科和工科,文科为100个样本,理科为100个样本,工科为10个样本, 此时可考虑将理科和工科合并成一个类别进行后续分析,数据类别的合并或处理,需要使 用数据编码功能,具体可见3.2节内容。

2.2.3 定量数据探索分析

针对定量数据探索分析,通常涉及异常值探索、数据分布探索和数据处理等。定量数 据探索如表 2-6 所示。

表 2-6 定量数据探索

项	分析或处理	作用
1	描述分析	查看数据是否存在异常值
2	可视化图	探索数据分布情况
3	数据处理	针对异常数据进行处理

首先可以使用描述分析,查看定量数据的最小值和最大值,如身高数据,正常成年人的身高介于 1.5~2.0 米,如果发现最小值小于 1.5 米,或最大值大于 2.0 米,那么意味着数据可能存在异常值。此情况需要确认原始数据是否正常,如果是原始数据的问题,则直接将原始数据修改成正确数据;如果是异常值,则可使用异常值处理方式进行处理,具体详见 3.3 节内容。除此之外,还可通过"3 倍标准差法则"查看数据中是否存在异常值,如果平均值为 1.7 米,标准差为 0.1,那么超出 3 倍标准差范围 (1.7-0.1×3, 1.7+0.1×3),即 1.4~2.0 米之外的数据通常被认为是异常数据,应该对该类数据进行处理。与此同时,通过查看数据平均值与中位数,可了解数据的整体分布情况,如果平均值与中位数差别很大,如平均值为 1.7 米,但是中位数为 1.5 米,就意味着可能存在异常数据,平均值明显大于中位数,意味着有个别很大的异常值拉升了平均值;反之,平均值明显小于中位数,意味着有个别很大的异常值拉升了平均值;反之,平均值明显小于中位数,意味

在使用描述分析了解数据基本特征后,还可使用图形直观地查看数据分布,常见的可 视化图共四类,分别是直方图、核密度图、箱线图和小提琴图,如表 2-7 所示。

项	可视化图	特点
1	直方图	简洁易懂,能直观地查看数据分布
2	核密度图	类似直方图,但以曲线形式展示分布特征
3	箱线图	通过展示不同分位数信息等查看数据分布
4	小提琴图	箱线图和核密度图的整合,深入查看数据分布

表 2-7 定量数据探索——可视化图

直方图简单易懂,其通过柱子高低来描述数据区间的分布多少,柱子高低代表真实的 数据分布量情况,其使用较广泛。核密度图与之类似,但核密度图使用曲线高低来描述数 据分布情况,并且核密度图的曲线高低仅具有分布相对大小的意义。箱线图的原理是利用 分位数,包括 25%、50%和 75%分位数来展示数据分布情况,小提琴图是将箱线图与核密度 图进行综合,绘制在一起,可更加深入地查看数据分布情况。在实际研究中,多数情况下 只需要使用直方图和核密度图探索数据分布特征。

关于定量数据探索分析, 接下来以案例形式具体说明, 当前有一份某专业 70 名学生的身高数据(见"例 2-1.xls"), 首先对其进行探索分析。第一步进行描述分析, 具体 SPSSAU 操作路径为【通用方法】→【描述】, 分析结果如表 2-8 所示。

名称	样本量	最小值	最大值	平均值	标准差	中位数	25%分位数	75%分位数	IQR 值
身高	70	1.410	2.140	1.642	0.159	1.640	1.490	1.765	0.275

表 2-8 身高描述分析结果

表格显示: 样本量为 70, 意味着并没有缺失数据。平均值为 1.642, 中位数为 1.640, 平均值与中位数基本相等, 意味着不存在异常数据。标准差为 0.159, 如果按偏离平均值 3 倍标准差作为标准, 此时数据并没有异常, 但如果按偏离平均值 2 倍标准差作为标准, 即 数据介于 1.324~1.96, 那么可能存在异常数据, 因为最大值为 2.140 明显高于 1.96。

表格中展示了 25%分位数、75%分位数和 IQR 值, 其意义分别为 25%的身高低于 1.490 米(或 75%的身高高于 1.490 米), 75%的身高低于 1.765 米(或 25%的身高高于 1.765 米), IQR 值为 75%分位数-25%分位数, 表示集中于中间的 50%身高数字范围的落差值, 该值越 大意味着数据越分散; 反之, 该值越小意味着数据越集中, 讲述箱线图时会涉及。接下来 使用直方图、核密度图、箱线图和小提琴图进行探索, 该四项均在 SPSSAU 平台【可视化】 模块中, 结果分别如下。



身高直方图如图 2-1 所示。柱子高度代表数据分布数量,身高的数据分布范围主要集中于 1.5~1.9 米。1.5 米以下或 1.9 米以上这两个区间的样本群体明显较少,意味着如果在后续分析中需要对数据进行分组,可能分为 3 个组别较适合,分别是 1.5 米以下、1.5~1.9 米、1.9 米以上。

与此同时,还可使用核密度图查看身高分布情况,如图 2-2 所示。核密度图曲线对应 的 X 轴数字为具体身高数字,Y 轴代表分布相对大小(核密度估计值),从核密度图可以看 出,身高数据范围集中于 1.5~1.8 米,并且从 1.8 米之后分布快速下降,这一结论与直方 图结论基本一致。直方图与核密度图的数学原理不同,但二者的实际功能或意义一致,因 而可以得到基本一致的结论,研究者以实际使用的便捷性和个人偏好,任选其一即可。



箱线图能展示数据的 25%分位数、50%分位数和 75%分位数,以及上限值和下限值。身高箱线图如图 2-3 所示。图中最上方横线表示上限值,该值=75%分位数+1.5IQR 值,图中最下方横线表示下限值,该值=25%分位-1.5IQR 值。上限值的意义为数据可接受的最大值,如果大于该值,意味着数据为异常数据。下限值的意义为数据可接受的最小值,如果小于该值,意味着数据为异常数据。可以看出,上限值明显较高,意味着可能存在较大的异常数据。



图 2-3 身高箱线图

小提琴图是将箱线图与核密度图进行整合。身高小提琴图如图 2-4 所示。其左右两侧 的弧线为核密度图的"纵向"展示,且完全对称。其中间展示了 5 个值并且使用线条连接, 从上到下依次为上限值(75%分位数+1.5IQR 值)、75%分位数、50%分位数、25%分位数和下 限值(25%分位数-1.5IQR 值)。相对于箱线图,小提琴图中使用了曲线展示数据分布情况。 具体使用时,可结合研究者的偏好选择。



图 2-4 身高小提琴图

通过对身高数据的探索分析,可以了解身高的整体分布情况,做到分析前"心中有数", 包括数据是否有缺失、是否有异常值、数据整体分布情况等,以及后续如果需要对数据进 行分组处理,大概需要分成几组合适、具体如何分组等。

2.2.4 小结

数据分为两种类型,分别是定类数据和定量数据,在实际研究中,首先应该分别针对 19/187 两种数据进行探索,了解数据中是否存在缺失数据或异常数据,并且对数据分布情况做到 "心中有数",如果存在缺失数据或异常数据,那么需要查看原因及确定是否有必要采取相 应的处理措施。通过对数据分布进行探索,对研究数据有了基本认知和了解,可为后续正 式分析做好准备。

2.3 数据分布之正态性分析

在统计研究中,非常多的研究方法均需要数据满足正态分布的特质,如方差分析时要 求因变量 Y 满足正态分布,或者线性回归分析时要求残差满足正态分布。因而研究数据是 否满足正态分布成为一件重要的事情。正态分布检验方式很多,包括图示法和检验法,图 示法包括直方图、PP 图、QQ 图、核密度图等,检验法包括 Kolmogorov-Smirnov 检验、 Shapiro-Wilk 检验和 Jarque-Bera 检验等。图示法和检验法各有优缺点,如表 2-9 所示。

表 2-9 正态分布检验方式汇总

项	检验方式	特点
1	图示法	简单且直观,但精确度相对较低
2	检验法	检验严格且精确度高

图示法比较直观,可通过观察图形特征情况进行判断,其在实际研究中使用广泛,至 于使用广泛的原因,作者认为可能因为"实际研究数据较难存在理论上的正态分布",尤其 是在数据量较小时,因而针对较稳健的分析(如方差分析)时,数据接近正态分布且与实 际情况更加吻合,通过图示法观察数据接近正态特质即可。检验法较严格,在数据样本较 大时使用检验法相对较多。具体研究时,通常可结合两种方法进行综合判断和科学抉择。 与此同时,如果在分析时不满足正态分布,那么此时可考虑先对数据进行转移处理, 使转换后的数据满足正态分布,再进行分析。

2.3.1 正态分布图示法

正态分布图示法具体可分为以下四种,分别是直方图、PP 图、QQ 图、核密度图,如表 2-10 所示。

项	图示法	特点
1	直方图	检验各分组数据的分布特征,并进行判断
2	PP 图	查看真实数据与理论正态分布数据的累计概率是否一致,并进行判断
3	QQ 图	查看真实数据与理论正态分布数据的分位数是否一致,并进行判断
4	核密度图	直观地查看数据的分布特征,并进行判断

表 2-10 正态分布图示法

直方图是一种数据分布的可视化展示,其数学原理为:首先将数据分为多个组别,然 后累计汇总各个组别的样本数量,并且进行展示。如果数据具有正态性,那么应该呈现"钟 形、两头低、中间高"的分布形状,如果直方图基本呈现此特征则说明数据具有正态性。类 似地,核密度图是对直方图的进一步抽象,其特征如果与正态分布基本一致,则说明数据 具有正态性。关于直方图和核密度图的进一步说明,详见 2.2.3 节内容。

关于 PP 图和 QQ 图的原理及实现说明, PP 图的原理是, 对比真实数据与理论正态分 布数据分别的累计概率值, 如果数据具有正态性, 那么真实数据的累计概率值应该与理论

正态分布数据的累计概率值基本保持一致,反之,如果数据不满足正态性,那么真实数据 的累计概率值与理论正态分布数据的累计概率值会有很大差别。用真实数据的累计概率值 与理论正态分布数据的累计概率值绘制散点,如果数据符合正态性,那么散点应该呈现一 条对称线;反之,散点应该明显偏离对称线。

与 PP 图类似, QQ 图的原理也是对比真实数据与理论正态分布数据是否具有一致性, 但 QQ 图对比的是分位数点 (PP 图对比的是累计概率值)。关于 QQ 图检验数据是否具有 正态性, 其与 PP 图基本一致, 如果散点呈现一条对称线, 那么说明数据具有正态性; 反 之, 如果散点明显偏离对称线, 则说明数据不具有正态性。PP 图或 QQ 图可通过 SPSSAU 平台【可视化】模块中的 PP/QQ 图绘制。以"例 2-1.xls", 即身高数据为例绘制 PP 图或 QQ 图。

图 2-5 所示为身高数据的 PP 图,图中横坐标为真实身高数据的累计概率(如小于 1.8 米的样本百分比),纵坐标为理论正态分布身高数据的累计概率(如理论上小于 1.8 米的样 本百分比)。图中的散点基本在对称线上,意味着真实身高数据的累计概率与理论正态分布 身高数据的累计概率基本一致,因而说明数据具有正态性;反之,如果多数散点远离对称 线,则说明数据不具有正态性。



图 2-6 身高数据的 QQ 图

图 2-6 所示为身高数据的 QQ 图,图中横坐标为真实身高数据的观察值(真实身高数据),纵坐标为期望正态值,即理论正态分布身高数据的分位数值(关于分位数值:如真实

数据中有 30%的人身高小于 1.7 米, 那么这 30%的数据应该对应的理论数据是 1.72 米, 此 处 1.7 米为真实身高数据的观察值即分位数值, 1.72 米为理论正态分布身高数据的分位数 值)。图中散点基本在对称线上, 意味着真实身高数据的观察值与理论正态分布身高数据的 分位数值基本一致, 因而说明数据具有正态性; 反之, 如果多数散点远离对称线, 则说明 数据不具有正态性。但从图中还可以看到真实身高为 2.140 米时, 该点明显偏离对称线, 如果将该点剔除, 则数据会展示出更好的正态性。

2.3.2 正态分布检验法

正态分布检验法通常包括以下 3 种,分别是 Kolmogorov-Smirnov 检验、Shapiro-Wilk 检验和 Jarque-Bera 检验,如表 2-11 所示。

项	检验	特点
1	Kolmogorov-Smirnov 检验	通常适用于大样本
2	Shapiro-Wilk 检验	通常适用于小样本
3	Jarque-Bera 检验	通常适用于大样本

表 2-11 正态分布检验法

一般来讲,如果样本较大(如样本大于 50),那么使用 Kolmogorov-Smirnov 检验较 适合;如果数据量较小(如样本小于 50),那么使用 Shapiro-Wilk 检验较适合。此处样本 的大小标准并不是数学角度的划分标准,建议研究者可结合检验法进行综合判断。Jarque-Bera 检验是针对偏度和峰度数据进行拟合优度判断的,由于偏度和峰度数据通常只能在大 样本数据中才能被很好地识别正态性特质,因此 Jarque-Bera 检验只适用于大样本数据。 关于上述 3 种检验,可通过在仪表盘中依次单击【通用方法】→【正态性检验】模块得到, 以下以身高数据为例进行说明。

Kolmogorov-Smirnov 检验的原假设为数据具有正态性,因而若接受原假设即 p>0.05,则说明数据具有正态性;反之,若 p<0.05,则说明数据不具有正态性。Kolmogorov-Smirnov 检验和 Shapiro-Wilk 检验如表 2-12 所示。当使用 Kolmogorov-Smirnov 检验时,p=0.198> 0.05,说明身高数据具有正态性,这一结论与图示法一致。与此同时,Shapiro-Wilk 检验的原假设为数据具有正态性,此处 p=0.010<0.05,即拒绝原假设,按 Shapiro-Wilk 检验标准,身高数据并不具有正态性,因而在分析时需要注意检验的适用性,如果 Shapiro-Wilk 检验适用于小样本(如把样本小于 50 作为标准),那么就不应该关注该检验 指标。

						Kolmogorov-Smirnov 检验 Shapiro-Wilk 材			Vilk 检验
名称	样本量	平均值	标准差	偏度	峰度	统计量	0	统计量	2
						D值	P	W值	ρ
身高	70	1.642	0.159	0.395	-0.105	0.088	0.198	0.953	0.010*

表 2-12 Kolmogorov-Smirnov 检验和 Shapiro-Wilk 检验

注: **p<*0.05。

Jarque-Bera 检验的原假设为数据具有正态性,若接受原假设即 p>0.05,则说明数据 具有正态性,反之,若 p < 0.05 则说明数据不具有正态性,如表 2-13 所示。表格中 p = 0.399 > 0.05,说明身高数据具有正态性,这一结论与图示法一致。

表 2-13 Jarque-Bera 检验

名称	样本量	χ^2	df	p
身高	70	1.839	2	0.399

由于正态分布的检验方式多样,并且不同检验方式在适用性上不同,因而在实际研究 中,建议综合图示法和检验法,多次对比进行科学抉择。如果图示法和检验法均显示数据 不具有正态性(并且研究方法要求数据具有正态性),则可考虑对数据进行正态分布转换处 理(或者用其他研究方法),接下来将阐述正态分布转换的几种方式。

2.3.3 正态分布转换处理

在实际研究中,理论上的正态分布并不存在,但是可以将数据进行转换,使其尽量满 足正态分布。正态分布转换可以在一定程度上让数据变得正态,正态分布转换通常有 3 种 处理方式,分别是数据压缩法、Box-Cox 变换和 Johnson 转换,这 3 种方式,均可通过在 仪表盘中依次单击【数据处理】→【生成变量】模块得到,如表 2-14 所示。

表 2-14 正态分布转换

项	处理方式	特点
1	数据压缩法	改变数据的单位大小,如求自然对数、开根号等
2	Box-Cox 变换	一种广义幂变换方式
3	Johnson 转换	Johnson 分布族变换

数据压缩法是指对数据进行一些数学变换处理,如取自然对数、开根号等。通常情况下,此种处理方式适用于数字较大时。当数字较大时进行取自然对数或开根号处理,此时处理后的数字会相对较小(如 100 这个数字取自然对数后为 4.605),相对于压缩数据分布的峰度和偏度,使数据尽可能地符合正态分布。需要注意的是,如果原始数据中有负数,则不能进行取自然对数或开根号处理。与此同时,如果数字非常小,则可以考虑对数据进行取指数次方,使其相对数字变得更大,以尽量满足正态分布的特质。数据压缩的方式有多种,常用于取自然对数或开根号,其他的一些处理方式,如取数据的平方、正余弦等,可结合实际数据进行尝试和对比。

数据压缩法是一种探索性处理方式,该处理方式相对简单方便,但会丢失原有数据的 实际意义,仅余下原有数据的相对大小意义,如身高为 1.6 米,求自然对数后变成 0.47 米, 实际意义已经消失,仅余下数字大小的相对意义。

除数据压缩法外,还可使用 Box-Cox 变换和 Johnson 转换。Box-Cox 变换是一种更为 复杂的数理转换方式,其原理与数据压缩法类似。Box-Cox 变换的数学公式如下:

$$y_new = \{\frac{y^{\lambda} - 1}{\lambda}, \lambda \neq 0$$
$$\ln y, \lambda = 0$$

式中, y 为原始数据; y_new 为新生成得到的数据; λ值为一个固定参数值, 计算机会 使用极大似然法求解λ值, 待得到该值后, 利用公式对数据进行转换。如果λ值为 0, 则完 全要对数据进行取自然对数处理; 如果λ值不为 0, 则先取数据的λ次方-1, 然后除以 值, 得到新数据。类似数据压缩法, Box-Cox 变换会让数据更可能变得正态, 但并非进行 Box-Cox 变换后, 数据一定会满足正态分布的特质。

Johnson 转换利用分布族数学原理进行转换, 其转换公式相对复杂, 有兴趣的读者可

23 / 187

查阅相关数理书籍。与此同时,与数据压缩法和 Box-Cox 变换类似, Johnson 转换后的数 据会相对更加满足正态分布,但是并非转换后数据一定会满足正态分布,而且会丢失数据 的原始意义,仅余下数据的相对意义。

在实际数据分析过程中,理论上的正态分布很难存在,首先使用图示法进行探索,如 果数据基本满足正态分布,也可以考虑使用部分替换方法,如方差分析时要求因变量 Y 满 足正态分布,如果因变量 Y 不满足正态分布,则改用非参数检验方法即可。如果某研究方 法对于数据满足正态分布有严格要求,则可考虑使用正态分布转换进行处理,并且结合多 种方式尝试对比,找出最为适合的处理方式,但需要注意的是,正态转换后数据原有数字 的实际意义会消失,仅余下数字大小的相对意义。

2.3.4 小结

非常多的分析方法或检验,均要求数据满足正态分布,本节首先针对正态分布的检验 方式,包括图示法和检验法进行说明;然后阐述当数据不满足正态分布时,使用正态分布 转换进行转换,使数据尽可能地满足正态分布,以满足分析方法的要求。

2.4 常用分析方法选择

上述 3 节内容分别从数据类型的区分、数据探索,以及数据正态分布进行说明。本节 将讲述数据的两种类型,包括定类数据和定量数据,进一步剖析基于数据类型下的常用分 析方法,包括定类或定量数据分析方法、定类和定类数据分析方法、定类和定量数据分析 方法、定量和定量数据分析方法,从4 个角度进行说明。本节更多的是阐述各种数据类型 组合下的分析方法的选择和基本应用,具体各类分析方法的事例和使用在第二篇对应章节 中有具体阐述,本节共分为五部分。

2.4.1 定类或定量数据分析方法

在 2.2 节已经讲述过两种数据类型及对应的探索分析,本节将进一步具体说明。定类 或定量数据分析方法如表 2-15 所示。

数据类型	分析方法
宁米	✔ 频数分析、饼图、圆环图、柱形图、条形图、帕累托图
龙关	✓ 卡方拟合优度检验
今 早	✔ 描述分析、直方图、核密度图、箱线图、小提琴图
疋里	✓ 单样本 t 检验、单样本Wilcoxon 检验

表 2-15 定类或定量数据分析方法

首先针对定类数据,可以使用频数分析了解所求数据的分布情况,使用饼图或圆环图 直观地展示数据分布,使用柱形图或条形图展示数据分析的相对大小情况。如果希望分析 "二八原则",那么可使用帕累托图进行展示。如果希望分析各类别的分布是否与预期有明 显差异,如男性和女性的预期分布是 50%和 50%,事实上的数据分布是否与预期有差异呢? 此时可使用卡方拟合优度检验进行分析。

针对定量数据,通常使用描述分析了解数据的平均值、最大值、最小值、标准差、方

差及百分位数等。如果希望使用图示法查看数据分布情况,则可使用直方图、核密度图、 箱线图或小提琴图。如果希望研究定量数据的平均值是否与预期有明显差异,如身高的预 期是 1.7 米,事实上真实身高是否与 1.7 米有明显差异呢?则可使用单样本 t 检验进行分析, 并且单样本 t 检验理论上要求身高满足正态分布,如果不满足,此时可使用单样本 Wilcoxon 检验进行分析。

除此之外,在了解数据分布之后或者在具体分析之前,还可以对数据进行处理,如针 对定类数据进行数据标签设置、数据编码等。定类或定量数据处理如表 2-16 所示。

数据类型		数据处理		事例
	~	数据标签	~	数字 1 表示男性, 数字 2 表示女性
<u>⇔</u> ₩	✓	数据编码	~	数字 1 表示初中, 数字 2 表示高中, 数字 3 表示大专, 数字 4 表示本
龙尖				科,希望组合成本科以下和本科共两组
	~	哑变量	~	回归分析时定类数据处理
	✓	数据编码	✓	如将成绩 0~60 分区间组合成低分组, 61~80 分区间组合成及格组,
定量				80分以上组合成高分组
	~	生成变量	~	如求对数、z标准化、中心化、正向化、逆向化处理等

表 2-16 定类或定量数据处理

针对定类数据,可对其进行数据标签设置,以标识数字的实际意义,如数字1的实际 意义为男性,数字2的实际意义为女性。与此同时,如果定类数据需要合并为几个类别, 那么可使用数据编码功能,如数字1~4分别表示初中、高中、大专和本科,现希望将数 据合并成本科以下和本科共两组,即意味着数字1、2和3可编码成数字1,数字4可编码 成数字2,从而得到新的数据,数字1表示本科以下,数字2表示本科(具体数字的意义 可使用数据标签进行设置)。在进行影响关系研究时,定类数据需要设置成哑变量,如专业 分为三类,分别是文科、理科和工科。哑变量处理说明如表2-17所示。

样本编号	专业	是否文科	是否理科	是否工科
1	文科	1	0	0
2	理科	0	1	0
3	理科	0	1	0
4	文科	1	0	0
5	工科	0	0	1
6	文科	1	0	0
7	工科	0	0	1

表 2-17 哑变量处理说明

哑变量处理的原则是,有多少个类别,即得到多少列新数据(列名为类别的名称),每 列只有1和0两个数字,数字1表示隶属该属性,数字0表示并非隶属该属性,如样本编 号为1时,该样本为文科,其对应的"是否文科"为数字1,对应的"是否理科"和"是否工科" 这两项为数字0。样本编号为5时,该样本为工科,那么其对应的"是否工科"为数字1,对 应的"是否文科"和"是否理科"这两项为数字0,新得到的3列统称为哑变量。

在具体研究时,如进行线性回归研究专业对某项的影响时,首先需要对专业进行哑变 量处理,得到3列;然后将其中任意两列放入模型中(注意是任意两列,余下1列作为参 照项不能放入模型中,如果3列均放入模型中就会出现数学逻辑问题,如将文科作为参照 项,那么理科和工科这两个哑变量需放入模型中)。关于哑变量处理,可通过在仪表盘中依 次单击【数据处理】→【生成变量】模块得到。

针对定量数据,如将成绩 0~60 分组成低分组,61~80 分组成及格组,80 分以上组成高分组,那么可使用数据编码功能(具体为范围编码功能)。成绩进行范围编码如图 2-7

所示。

		范围	銅编码	~ 🕜			Q
75	范围介	Ŧ		编码数	7	标签[可选]	
0	到	60	\rightarrow	1	+ -	低分组	
61	到	80	\rightarrow	2	+ -	及格组	
81	到	100	\rightarrow	3	+ -	高分组	

图 2-7 成绩进行范围编码

针对定量数据,如希望对某项进行取对数、开根号,或者对其进行 z 标准化、中心化、 正向化、逆向化处理等,均可通过在仪表盘中依次单击【数据处理】→【生成变量】模块 得到,具体内容可见 3.4 节。

2.4.2 定类和定类数据分析方法

定类数据和定类数据之间的关系研究,基本上是基于差异关系的研究,统称为卡方检 验,但卡方检验又细分出较多的检验方式。除此之外,还可使用对应分析研究两个或多个 定类数据之间的关联关系,在确认数据之间具有差异关系后,可进一步进行对应分析并可 视化展示两个或多个定类数据之间的关联情况,具体定类和定类数据分析方法如表 2-18 所示。

类型	分析方法	说明	SPSSAU 实现路径
	交叉 (卡方)	✓ 两个定类数据之间的交叉分析	【通用方法】→【交叉(卡方)】
		✓ 提供Pearson卡方检验	
		 ✓ 提供 Pearson 卡方检验、Yates 校正卡方检 	
	卡方检验	验和 Cochran-Armitage 趋势卡方检验等	【实验/医学研究】→【卡方检验】
		✔ 提供卡方检验事后多重比较	
差异关系		✓ 提供 <i>m</i> × <i>n</i> 表格的fisher 卡方检验结果	
	fisher 卡方	✓ 提供 Pearson 卡方检验和 Yates 校正卡方	【实验/医学研究】→【fisher 卡方】
		检验结果	
	配对卡方	 ✓ 包括 McNemar 检验和 Bowker 检验 	【实验/医学研究】→【配对卡方】
	分尾卡方	✔ 研究两个定类数据之间的差异关系时,	【实验/医学研究】→【分目卡方】
	カムトカ	将第三个类别因素纳入考虑范畴	【关题]因于明九】 「【方层下方】
土旺土玄	对应分析	✓ 包括简单对应和多元对应分析	【词类研究】→【对应分析】
八帆大尔	AT 12 11 11	✔ 图示化展示定类和定类数据的关系情况	▶□ 19711111111111111111111111111111111111

表 2-18 定类和定类数据分析方法

通常情况下只需要使用交叉(卡方)研究两个定类数据和定类数据之间的差异关系, 也称交叉分析,其具体检验方式为 Pearson 卡方检验。如果样本量较少, Pearson 卡方检 验结果可能不够准确,则可使用 Yates 校正卡方或者 fisher 卡方进行检验;如果定类数据 中出现序定类数据(如病情严重程度分为轻度、中度和重度),则可使用 Cochran-Armitage 趋势卡方进行检验。具体关于 Pearson 卡方检验、Yates 校正卡方检验和 fisher 卡方检验的选择和实例,详见 4.3 节内容。 如果研究数据为配对实验数据,则可使用配对卡方,配对卡方包括两种检验,分别是 McNemar 检验和 Bowker 检验。除此之外,如果在研究两个定类数据之间的差异关系时, 可将第三个类别(通常称作混杂因素)纳入考虑范畴,此时应该使用分层卡方。关于配对 卡方、分层卡方的具体选择、使用及实例可见 4.3 节内容。

对应分析是一种多元统计分析方法,其目的在于将具体类别以散点形式呈现在图中, 通过观察点与点之间的距离来分析关联情况。如果分析时仅包括两个定类数据,则为简单 对应分析;如果分析时超过两个定类数据,则为多元对应分析。具体实例应用可见 6.3 节 内容。

在分析方法使用前,需要准备好正确的数据(关于两种数据格式,可见 2.1 节内容), 针对交叉(卡方),其仅适用于"普通数据格式"。如果是卡方检验,SPSSAU 平台支持"普通 数据格式"和"加权数据格式";如果是 fisher 卡方,SPSSAU 平台支持"交叉汇总格式",即将 两个定类数据的具体两交叉类别样本个数呈现在表格中。卡方检验-交叉汇总格式如表 2-19 所示。

名称	治愈	显著	无效
A 药	35	32	12
B 药	16	23	14

表 2-19 卡方检验-交叉汇总格式

如果是配对卡方或卡方检验, SPSSAU 平台支持"普通数据格式"和"加权数据格式"。如 果是对应分析, SPSSAU 平台也支持"普通数据格式"和"加权数据格式"。

2.4.3 定类和定量数据分析方法

定类数据和定量数据之间的关系研究,多数情况下是研究差异关系,如研究不同专业 群体工资的差异性,差异关系研究包括多种方法,具体如表 2-20 所示。

分析方法	说明	SPSSAU 实现路径
	✔ 定类和定量数据的差异关系	【通用方法】→【方差】
古羊公拆	✔ 如不同专业群体生活费的差异性	【进阶方法】→【双因素方差】
力左方初	✔ 包括单因素方差、双因素方差、三因素方	【进阶方法】→【三因素方差】
	差和多因素方差	【进阶方法】→【多因素方差】
	✔ 单因素方差分析后的两两组别比较	
事后多重比较	✔ 如研究发现专业与生活费有差异后,可进	【进阶方法】→【事后多重比较】
	一步分析文科生和理科生工资的差异	
	✓ 定类(且为两个类别)与定量数据的差异	
独立样本 t 检验	关系	【通用方法】→【独立 t 检验】
	✓ 如男生和女生生活费的差异	
	✓ 定类(且为两个类别)与定量数据的差异	
Mann-Whitney 检验	关系(定量数据不满足正态分布前提时)	【通用方法】→【非参数检验】
	✔ 如男生和女生的生活费	
	✓ 定类和定量数据的差异关系(定量数据不)	
Kruskal-Wallis 检验	满足正态分布前提时)	【通用方法】→【非参数检验】
	✔ 如不同专业群体生活费的差异	
而才,於政	✔ 配对数据的差异关系	【通田方注】→【配对・检验】
日山八丁 1 1 1 2 3 2	✔ 如实验前和实验后,学生成绩的差异	
	✓ 配对数据的差异关系,且配对数据不满足	
配对Wilcoxon 检验	正态分布前提时	【实验/医学研究】→【配对wilcoxon】
	✔ 如实验前和实验后,学生成绩的差异	

表 2-20 定类和定量差异关系方法

如果研究不同专业(且专业包括两个以上类别,如文科、理科和工科)与每月生活费 的差异关系,专业为定类数据,每月生活费为定量数据,此时通常使用单因素方差分析进 行研究。此处,定类数据仅为一项即专业,如果在某些研究中,定类数据的个数大于 1, 如两个则为双因素方差,如果三个则为三因素方差,甚至更多时则统称为多因素方差。单 因素方差简单易懂,因而使用最多。如果发现不同专业群体在生活费上呈现差异性,那么 具体是哪两个专业之间呈现差异性呢?此时可使用事后多重比较进一步研究。如果专业仅 包括两个类别,分别是文科和理科,则研究专业与生活费的差异时,一般使用独立样本 t 检验。

在方差分析理论上要求因变量 Y,如生活费需要满足正态分布,如果其不满足(或者 说严重不满足时),则可使用非参数检验方法 Kruskal-Wallis 检验研究差异关系。类似地, 独立样本 t 检验也要求因变量 Y(如生活费)满足正态分布,如果不满足,则可使用非参 数检验方法 Mann-Whitney 检验进行研究。

如果研究数据为配对实验数据,如研究实验前和实验后的成绩差异,那么可使用配对 t 检验。理论上要求实验前和实验后的数据都满足正态分布,如果不满足,则考虑使用配 对 Wilcoxon 检验。

如果是研究影响关系,如 X 对于 Y 的影响,则要结合 Y 的数据类型,选择正确的分析 方法。若 Y 为定量数据,则可使用线性回归进行研究;若 Y 为定类数据,则可使用 Logistic 回归进行研究。具体线性回归可进一步分为线性回归、分层回归和分组回归, Logistic 回归可进一步分为二元 Logistic 回归、多分类 Logistic 回归和有序 Logistic 回归。 定类和定量影响关系方法如表 2-21 所示。

分析方法	说明	SPSSAU 实现路径
线性同归	 ✓ 线性回归、分层回归、分组回归 ✓ 加 X 对 Y 的影响, Y 一定是定量数据, X 可 	【通用方法】→【线性回归】 【讲阶方法】→【分层回归】
	为定类数据或定量数据	【计量经济研究】→【分组回归】
Logistic 回归	 ✓ 二元 Logistic 回归、多分类 Logistic 回归、 有序 Logistic 回归 ✓ 如 X 对 Y 的影响,Y 一定是定类数据,X 可 为定类数据或定量数据 	【进阶方法】→【二元Logit】 【进阶方法】→【多分类Logit】 【进阶方法】→【有序Logit】

表 2-21 定类和定量影响关系方法

当研究 X 对于 Y 的影响时,如果 Y 是定量数据,那么可使用线性回归,此时 X 中如果 包含定类数据,那么可对 X 进行哑变量处理后放入模型中(关于哑变量内容可见 2.4.1 节 内容)。如果 X 是定量数据,则不需要处理可直接将其放入模型中。线性回归有两种变换 形式,分别是分层回归和分组回归。分层回归,如线性回归模型 1 里有 X1,模型 2 在模 型 1 基础上加入 X2 (模型 2 里有 X1 和 X2),或者模型 3 在模型 2 基础上加入 X3 (模型 3 里面有 X1、X2 和 X3),这 3 个模型(模型 1、模型 2 和模型 3)均是线性回归,但 3 个模 型合并在一起后,可以查看新增加的 X 对模型解释力度的变化及帮助等,分层回归实质上 还是线性回归,其常用于模型稳健性检验。分组回归是指研究 X 对于 Y 的影响时,区分不 同组别进行研究,如组别为男性时,X 对于 Y 的影响,组别为女性时,X 对于 Y 的影响, 分组回归实质上也是线性回归,其常用于模型稳健性检验、异质性检验等。

如果研究 X 对于 Y 的影响时, Y 的数据类型为定类数据, 不论 X 是定类数据还是定量数据, 那么此时可使用 Logistic 回归。如果 X 是定类数据, 则可对其进行哑变量处理后放入模型中; 如果 X 是定量数据, 则不需要处理可直接放入模型中。以 Y 的具体数据类型为标准, 可进一步分成三种类型的 Logistic 回归, 如表 2-22 所示。

类型	说明
二元Logistic 回归	Y 仅包括两个类别,如是和否、购买和不购买
多分类Logistic 回归	Y包括至少三个类别,如专业分为文科、理科和工科
有序Logistic 回归	Y 为有序定类数据,如病情严重程度分为轻度、中度和重度

表 2-22 三种类型的 Logistic 回归

如果 Y 仅包括两个类别,如购买和不购买,此时可使用二元 Logistic 回归,并且通常 情况下 Y 的两个类别需要分别使用数字 1 和数字 0 进行标示,如数字 1 表示愿意购买,数 字 0 表示不愿意购买。如果 Y 包括至少三个类别,并且三个类别之间有相对可比较的大小 关系,如病情严重程度分为轻度、中度和重度,虽然看成是三个类别,但三个类别之间可 进行程度大小对比,此时可使用有序 Logistic 回归,也可使用多分类 Logistic 回归。如果 类别之间只有分类关系,如专业分为文科、理科和工科,且类别之间不能进行程度大小对 比,此时只能使用多分类 Logistic 回归。

2.4.4 定量和定量数据分析方法

如果是研究定量数据和定量数据之间的关系,首先可使用散点图查看两个定量数据之间的 基本关系情况,然后使用相关分析查看两个定量数据之间的相关系数情况。如果两个定量 数据均满足正态分布,那么使用 Pearson 相关系数即可;如果两个定量数据不满足正态分 布,则可使用 Spearman 相关系数得到更科学的结论。与此同时,如果研究 X 对于 Y 的影 响,只要 Y 为定量数据,则可使用线性回归分析进行研究,线性回归可进一步分为线性回 归、分层回归和分组回归,如表 2-23 所示 (包括 2.4.3 节部分内容)。

分析方法	说明	SPSSAU 平台实现路径
散点图	✔ 直观地查看定量和定量数据的关系	【可视化】→【散点图】
相关公析	✔ 两个定量数据之间的相关关系	【通田支注】→【相关】
相大力切	✓ 包括Pearson 相关和Spearman 相关	【通用力弦】 一【相关】
		【通用方法】→【线性回归】
线性回归	✔ 线性回归、分层回归、分组回归	【进阶方法】→【分层回归】
		【计量经济研究】→【分组回归】

表 2-23 定量和定量数据分析方法

除此之外,定量数据还可使用其他分析方法,如主成分分析、因子分析和聚类分析。 定量数据信息浓缩及聚类方法如表 2-24 所示。

表 2-24 定量数据信息浓缩及聚类方法

分析方法	说明	SPSSAU 实现路径
主成分分析	✔ 信息浓缩,如将 20 列数据浓缩成 3 个成分	【进阶方法】→【主成分】
因子分析	✔ 信息浓缩,如将 20 列数据浓缩成 3 个因子	【进阶方法】→【因子】
		【进阶方法】→【聚类】
聚类分析	✓ 包括 <i>K</i> -means 聚类、 <i>K</i> -prototype 聚类和分层聚类	【进阶方法】→【聚类】
		【进阶方法】→【分层聚类】

如果研究时包括多个定量数据,如 20 个定量数据,并且这 20 个定量数据之间有一定的相关关系,此时可对这 20 个定量数据进行信息浓缩,提炼成较少的几个成分(或因子),使用较少的几个成分(或因子)去揭示这 20 个信息,简而言之,将 20 个定量数据浓缩成

29 / 187

几个关键词,使用该关键词表示 20 个定量数据的信息。信息浓缩可使用的分析方法分别 是主成分分析和因子分析,在实际研究中,主成分分析和因子分析均基于定量数据,如果 包括定类数据,则考虑首先对定类数据进行哑变量处理,然后将哑变量数据放入分析中, 更多说明和事例可见第 6 章内容。

结合研究目的,还可对定量数据进行聚类,聚类包括两类,分别是样本聚类和指标聚 类。样本聚类指,如有 100 行样本,那这 100 行样本可以聚类成几个类别。指标聚类指, 如有 20 列指标数据,那这 20 列指标数据可以聚成几个类别。具体到分析方法上,如果是 样本聚类,那么通常使用 K-means 聚类方法,K-means 聚类方法仅针对定量数据,如果 数据中包括定类数据,那么此时可使用 K-prototype 聚类方法。如果是指标聚类,则通常 使用分层聚类方法完成分析。

定量数据可适用的分析方法有很多,如可针对定量数据进行权重计算、预测分析、优 劣决策分析、综合评价等。在实际研究中应以研究目的作为标准进行选择和使用,具体可 见第三篇内容(数据综合评价及预测)。如果研究数据是通过问卷收集得到的,可参考本书 第四篇内容(问卷数据分析)。如果是医学相关领域收到的数据,可参考本书第五篇内容 (医学数据分析)。

2.4.5 小结

本章从数据类型视角,分别阐述了定类或定量数据的常用分析方法选择和使用,并分 别阐述了定类和定类数据的分析方法选择、定类和定量数据的分析方法选择,以及定量和 定量数据的分析方法选择,以便读者初步理解各类常用分析方法的基本要求和使用意义, 更多关于分析方法的具体使用和案例内容可见后续章节。

第3章 问卷数据清理

在正式进行数据分析前,首先要对数据进行探索分析,在对数据特征和基本分布情况 有了了解后,再对数据进行清理,或者在进行某分析的同时,对数据进行处理。通常情况 下,数据清理工作包括数据标签设置、数据编码、异常值处理、生成变量和标题处理等。 数据标签针对定类数据时,标识数字代表的是实际意义;数据编码可对原始数据进行重新 编码,或者重新组合类别进行处理;如果探索分析时发现异常值,则可使用异常值处理功 能;如果在分析前需要对数据进行处理,如取其对数值、对数据进行标准化处理等,则可 使用生成变量功能;对分析项的名称(标题)进行修改,可使用标题处理功能。本章分别 阐述 SPSSAU 平台中数据清理的操作和原理,共分为5节。

3.1 数据标签设置

在计算机的世界里,其只会对数字进行处理,文字信息会被处理成数字,通过数据标签进行标识。例如,性别分为男和女,计算机存储的时候会将男和女分别存储为数字(如数字1表示男,数字2表示女),并且使用数据标签对其进行标识,标识数字代表的意义则称为数据标签。例如,要输出男和女的比例,原理上计算机是先计算数字1和数字2的比例,然后在输出的时候将数字1和数字2分别替换成已有的数据标签男和女,从而实现计算机与人的交互(如果没有数据标签,则默认显示数字1和数字2分别的比例)。此处男和女分别使用什么样的数字去标识是由研究数据决定的,研究者只需要理解数字代表的实际意义即可。SPSSAU 平台数据标签功能如表 3-1 所示。

项	说明
数据标签	✓ 针对选中项标识数据标签
上住坛交	✓ 上传数据文档时要标识好标签信息
工程你觉	✓ 上传数据如果为文字, SPSSAU 平台会自动按标签处理
下载标签	✓ 下载所有的标签信息

表 3-1 SPSSAU 平台数据标签功能

	0	P	C	D	E.	F	C	TT	т
	A	D D			E	Г	6	H	1
1	编兮	性别_数字格式	性别	性别_随意格式	A1	A2	A3	A4	字历
2	1	1	男	1	1	3	3	1	1
3	2	1	男	男	3	1	2	1	1
4	3	1	男	男	2	1	5	1	2
5	4	1	男	男	4	4	5	4	2
6	5	1	男	男	5	5	4	4	2
7	6	2	女	2	2	2	2	2	2
8	7	2	女	女	5	1	4	1	2
9	8	2	女	女	1	3	5	1	2
10	9	2	女	女	4	1	4	4	2
11	10	2	女	女	1	3	5	1	3
12	11	2	女	女	3	3	2	5	3
13	12	2	女	女	1	3	4	1	3
14	13	2	女	女	3	1	1	3	3
15	14	2	女	女	1	5	5	2	4
16	15	1	男	男	1	2	5	3	4

图 3-1 数据标签案例部分数据

在仪表盘中可通过依次单击【数据处理】→【数据标签】模块来完成对数据标签的设 置,与此同时,如果原始数据格式中已处理好数据标签格式,那么上传数据文档时平台会 自动解析数据标签信息。如果上传数据为文字格式,那么平台会自动将文字编码成数字, 并且会自动打上数据标签。与此同时,在平台中已有的数据标签信息,可自行下载及使用, 包括单独下载数据标签信息、或者下载数据文件时附带下载数据标签信息。

接下来以案例形式阐述 SPSSAU 平台中的数据标签功能(见例 3-1.xls)。数据标签案 例部分数据如图 3-1 所示。

数据中 B 列(性别_数字格式)只有两个数字分别是 1 和 2, 上传到 SPSSAU 平台中时, 平台并不知晓数字代表的意义,此时可进行设置,首先单击【性别 数字格式】标题项,然 后在右侧输入框中分别对应输入【男】和【女】,最后确认即可。数据标签案例部分数据操 作如图 3-2 所示。

批量选中处理 1 ②	数据标签 💿 🧕 🔮
≫编号	
◎ 性别_数字格式	1 表示 男
♥ 性别_文字格式	2 表示 女
▶ 性别_随意格式	
SA1	
A2 A	
SA3	
SA4	
◎ 学历	
◎ 身高	
● 是否愿意购买	

4	٨	P	C
	A	В	U
1	标题	数字	标签
2	性别_数字格式	1	男
3		2	女
4	学历	1	中学
5		2	大专
6		3	本科
7		4	研究生

图 3-2 数据标签案例部分数据操作 图 3-3 数据标签的格式说明

如果有多个标题项,如图中 A1、A2、A3 和 A4 共四项,这四项的数据标签信息完全 一致,那么可先按住【Ctrl】或【Shift】组合键选中这四项,然后设置标签,即可批量对这 四项进行处理。

如果需要设置的标签项非常多,更为方便的设置方式为,在上传数据文档中先加入一 个工作表,名称为"tags",然后将所有的标签信息放入该工作表中,上传后平台会自动解析 数据标签。工作表"tags"中具体放置数据标签的格式说明如图 3-3 所示。

使用一个单独的工作表,名称为"tags",表中的 A、B、C 列分别代表标题、数字和标 签。标题列为具体数据列的标题信息,标题列中的具体标题名称需要与原始数据保持一致, 并且仅在该标题的第 1 个标签数字展示,该标题的其余标签项留空即可,数字和标签分别 为数字与标签的对应关系。当正确数据上传后,平台会自动解析已设置好的数据标签信息。

上述已经提及,计算机只认识数字信息,如果是文字信息,那么 SPSSAU 平台如何处 理呢?文字信息的数据标签处理如图 3-4 所示。C 列(性别_文字格式)全部是文字,并且 只包括两个文字,分别是"男"和"女",而 D 列(性别_随意格式)是数字和文字信息的混合。

如果某列数据全部是文字,那么平台会自动将文字编码成数字,编码规则按拼音优先 排序,从数字 1、2、3 依次编码,由于 C 列仅有两个文字,并且"女"的拼音在先,因此 "女"编码为数字 1,"男"编码为数字 2,并且平台会自动设置好数据标签。如果某列数据既 包括数字又包括文字,那么平台的处理规则为保留数字信息,将文字信息置为 null 值。图 中 D 列的两个数字会保留,但是其余的文字信息全部会被置为 null 值。

在实际研究中、正确的数据格式为、要么某列全部为数字、要么某列全部为文字。如

果是文字和数字的混合,此种情况仅保留数字,因而研究者需要确保数据格式正确后再将 数据上传到平台中。设置完成数据标签后、后续进行分析时、只要涉及数字实际意义的呈 现时,平台就会呈现其标签信息,而不会呈现数字;反之,如果没有设置好标签,平台会 呈现如1、2此类的数字信息。

在完成数据标签设置后、如果需要查看具体的数据标签信息、可通过在仪表盘中依次 单击【数据处理】→【数据标签】模块,即可轻松查看该标签信息。如果需要下载标签信 息、共有两种方式、第一种方式是下载数据文档、通过在仪表盘中依次单击平台右上方的 【我的数据】→【下载】模块,下载原始数据,即可在工作表"tags"中查看标签信息。第二 种方式是通过在仪表盘中依次单击【数据处理】→【数据标签】模块后,单击【下载】按 钮进行下载。下载数据标签如图 3-5 所示。

	А	В	С	D
1	编号	性别_数字格式	性别_文字格式	性别_随意格式
2	1	1	男	1
3	2	1	男	男
4	3	1	男	男
5	4	1	男	男
6	5	1	男	男
7	6	2	女	2
8	7	2	女	女
9	8	2	女	女
10	9	2	女	女
11	10	2	女	女
12	11	2	女	女
13	12	2	女	女
14	13	2	女	女

	数据标签	0 Q	¢	4	_
1	表示	男			
2	表示	女			

图 3-4 文字信息的数据标签处理 图 3-5 下载数据标签

3.2 数据编码

某些分析方法仅对数据格式有特殊要求,如采用二元 Loaistic 回归时因变量只能为数 字 1 和数字 0, 但是原始数据包括数字 1 和数字 2, 因此可将数字 1 和数字 2, 分别替换 成数字 1 和数字 0. 以便满足数据分析方法的基本要求。在探索分析时发现学历分为中学、 专科、本科和研究生四类,但是中学样本过少,现希望将中学和专科合并为"本科以下"; 或者将身高数据分成两个组别,如按1.7米作为划分标准,分成"较低身高"和"较高身高"两 组、诸如上述需求、需要使用数据编码功能完成。

SPSSAU 平台中提供了 3 种数据编码功能,分别是数字编码、范围编码和自动分组。 上述采用二元 Logistic 回归分析时的数字编码、学历分组均可通过数字编码功能完成。如 果是身高数据分组,则可使用范围编码完成。如果是希望按照一定规则(如按平均值或分 位数)将数据自动分成几个组别,则可使用自动分组功能。SPSSAU 平台数据编码功能如 表 3-2 所示。

功能	说明
数字编码	✔ 针对具体数字进行编码
范围编码	✔ 将某范围内的数据进行编码
自动分组	✔ 平台自动对数据分组编码并打上数据标签

数据编码可通过在仪表盘中依次单击【数据处理】→【数据编码】模块进行设置,以 案例形式阐述 SPSSAU 平台中的数据编码功能(见例 3-1.xls),数据编码案例部分数据如 33 / 187 图 3-6 所示。

针对学历,其数字1代表中学,数字2代表大专,数字3代表本科,数字4代表研究 生。现希望将中学和大专合并成一个组合即大专以下,此时,可将数字1和数字2编码成 数字1,数字3编码成数字2,数字4编码成数字3。确认处理后,平台会默认生成一个新 标题"New_学历"用于标识新得到的学历数据,新数据中的数字1代表大专以下,数字2代 表本科,数字3代表研究生。如果勾选图中的【覆盖】复选框,则会在原始数据基础上进 行替换而不会生成新标题。学历进行数字编码的操作如图3-7所示。

I	L	K			
学历	身商	是否愿意购买			
1	1.832387983	1			
1	1.937183688	1			
2	2.30444508	1		数字编码 🖌 🕜	覆盖 🗌 🕜
2	1.982761766	1			
2	1.746921073	1	1	\rightarrow 1	中学
2	1.850742267	1	2		+=
2	1.940240607	2	2		~~
2	1.952610427	2	3	\rightarrow 2	本科
2	1.806957131	2			_
3	1.728906006	2	4	\rightarrow 3	研究生
2	4 (70220404	2			

图 3-6 数据编码案例部分数据 图 3-7 学历进行数字编码的操作

针对身高数据,如果希望将其分成两组,1.7 米以下为"较低身高"组别,1.7 米以上为 "较高身高"组别,则可使用范围编码。身高进行范围编码如图 3-8 所示。确认处理后,平 台会默认生成一个新标题"New_身高",用于标识新得到的身高数据,新数据中的数字 1 代 表"较低身高",数字 2 代表"较高身高"。如果勾选图中的【覆盖】复选框,则会在原始数据 基础上进行替换而不会生成新标题。

如果希望对身高自动分组, SPSSAU 平台能提供 4 种方式, 分别是按 27%分位数和 73% 分位数分成 3 组 (小于 27%分位数、27%~73%分位数、大于 73%分位数)、按 50%分位数分 成 2 组、按平均值分成两组、按 25%、50%和 75%分位数分成 4 组 (小于 25%分位数、 25%~50%分位数、50%~75%分位数、大于 75%分位数)。自动分组仅方便用于快速分组处理, 其可以完全被范围编码替换。分组编码操作如图 3-9 所示。



图 3-8 身高进行范围编码 图 3-9 分组编码操作

在数据编码时, SPSSAU 平台能提供两个参数, 分别是【覆盖】和【标签同步】。 SPSSAU 平台数据编码参数值说明如表 3-3 所示。在进行数字编码和范围编码后, 如果勾 选【覆盖】复选框, 则平台会直接对编码标题项的原始数据进行覆盖; 如果不勾选【覆盖】 复选框, 则平台会生成新标题(前缀为"New_")来标识新得到的编码数据。【标签同步】 意义为, 数字编码时, 如果数据已有标签信息, 如数字 1 为男, 数字 2 为女, 此时可将数 34/187 字 1 编码为 5, 数字 2 编码为 10, 那么得到新数据时其数字 5 和 10 的标签分别为男和女。

表 3-3 SPSSAU 平台数据编码参数值说明

参数	说明		
覆盖	✓ 数字编码和范围编码后,直接覆盖原始数据		
标签同步	✓ 编码标题有标签信息,会自动同步到新标题中		

3.3 异常值处理

如果在探索分析时发现数据中存在异常值,可通过在仪表盘中依次单击【数据处理】 →【异常值】模块进行异常值设置和处理。异常值设置包括判断标准和异常值处理共两项, 如表 3-4 所示。

表 3-4 SPSSAU 异常值设置

项	说明	
判断标准	设定异常值的标准,如身高大于 2.2 米	
异常值处理	针对符合的异常值,进行对应处理	

关于异常值判断标准,如果在探索分析时发现身高大于 2.2 米,那么可在判断标准时 设置"数字>2.2",判断标准的设置还包括"缺失数字",其指原始数据中某项为 null 值。与此 同时,判断标准还包括等于某个数字、小于某个数字或数字偏离标准差幅度(通常是 3 倍 或 2 倍标准差,可见 2.2.3 节内容)。判断标准共有 5 个条件,5 个条件是或者的关系,选 择任意一个或多个判断标准时,只要选中的任意一个判断标准满足即可。异常值设置如图 3-10 所示。

批量选中处理 1 📀	异常值设置 💿 Ω
编号 性别_数字格式 性别_文字格式 性别_随意格式 A1 A2 A3 A4 学历	判断标准 ■缺失数字 □数字 = 0 □数字 < 0 □数字 > 999
身高 是否愿意购买	数字 ± 3 标准差 异常值处理 @设置为Null 「填补 平均值 ◆ 「插值 线性插值 ◆

图 3-10 异常值设置

设置好异常值判断标准后,需要对异常值功能进行设置,如表 3-5 所示。异常值处理 共分为 3 种方式,分别是设置为 Null、填补和插值,3 种方式为互斥关系,只能选择其中 任意一种方式。

表 3-5 SPSSAU 异常值功能设置

异常值处理方式	说明	
设置为 Null	将满足条件的异常值设置为 Null 值	
填补	包括平均值、中位数、众数、随机数、数字 0 和自定义共 6 种方式	
插值	包括线性插值和该点线性趋势插值两种方式	

设置为 Null 这种处理方式使用较多,可直接将满足条件的异常值设置为 Null 值,但此种方式处理后,分析的有效样本量会减少,因而其适用于样本量相对较大时。如果采用填补处理方式,其包括 6 种方式,其中平均值、中位数或众数填补,指先将除异常值外的数据取平均值(中位数或众数),然后将异常值全部替换成该平均值(中位数或众数);如果是随机数填补,平台会将满足条件的异常值替换成一个随机数字;数字 0 填补指平台将满足条件的异常值替换成数字 0;自定义填补指平台将满足条件的异常值替换成主动设置的一个具体数字。如果采用插值处理方式,共包括两种,分别是线性插值和该点线性趋势插值,如表 3-6 所示。

编号	原始数据	线性插值	该点线性趋势插值
1	1	1	1
2	3	3	3
3	4	4	4
4	异常值	5	5.46511
5	6	6	6
6	9	9	9

表 3-6 插值处理方式

线性插值指将离异常值最近的前面 1 个点和后面 1 个点,如表格中编号为 4 的异常值, 其前面一个点编号为 3 原始数据为 4,即(3,4),其后面一个点编号为 5 原始数据为 6,即(5,6),将这两个点连成一条线时会得到坐标公式,代入异常值的编号最终会计算得到 线性插值。例如,表格中(3,4)和(5,6)两个点连成直线时,该直线的坐标公式为 y =x+1,代入异常值对应的编号 4,得到线性插值的具体值为 4+1=5。另外,编号指原始数 据的编号,如有 100 个样本,编号则从 1、2、3 一直递增到 100。

该点线性趋势插值指先将除异常值外的其他正常数据进行线性回归拟合(其中 Y 为原 始数据, X 为编号),得到线性回归方程公式(关于线性回归,可参考第 5 章内容);然后 将异常值对应的编号作为 X 代入,计算得到的 Y 值即为该点线性趋势插值。除异常值外, 共有 5 行数据,该编号作为 X,原始数据作为 Y,进行线性回归拟合得到线性回归方程公 式,即 Y =-0.30233+1.44186X,代入异常值编号 4,即该点线性趋势插值等于-0.30233+ 1.44186×4=5.46511。

异常值处理方式还可使用缩尾法和截尾法,通常在经济学相关领域中使用较多,具体可见 3.4.6 节内容。

3.4 生成变量

数据清理时经常会使用生成变量功能, SPSSAU 平台提供的生成变量功能, 共包括常用、量纲处理、科学计算、汇总处理、日期相关处理和其他共六类, 可通过在仪表盘中依次单击【数据处理】→【异常值】模块找到。SPSSAU 平台生成变量功能如图 3-11 所示。
SPSSAU 科研数据分析方法与应用

生成变量 🖉 🧕 高级公式 🔮	
平均值	
请输入搜索	
常用	
平均值	
求和	
虚拟(唑)变重 7标准化(S)	
中心化(C)	
乘积(交互项)	
自然对数(Ln)	

图 3-11 SPSSAU 平台生成变量功能

接下来将生成变量功能进行汇总,如表 3-7 所示,并且分别进行阐述。

表 3-7 生成变量功能汇总

项	说明
一世日	包括 8 项,分别是平均值、求和、虚拟(哑)变量、z 标准化、中心化、乘积(交互项)、
带用	自然对数(ln)和 10 为底对数
皇炯协理	包括 15 项,分别是归一化、均值化、正向化、逆向化、适度化、区间化、初值化、最小值
里纳处理	化、最大值化、求和归一化、平方和归一化、固定值化、偏固定值化、近区间化、偏区间化
科学计算	包括 7 项,分别是平方、根号、绝对值、倒数、相反数、三次方和取整
汇总处理	包括 4 项,分别是最大值、最小值、中位数和计数
日期相关处理	包括 5 项,分别是日期处理、日期相减、滞后处理、差分处理和季节差分
甘仙	包括 10 项,分别是样本编号、Box-Cox 变换、秩、缩尾处理、截尾处理、Johnson 转换、排
共祀	名、相除、相减、非负平移

3.4.1 常用处理

关于常用功能中虚拟(哑)变量的原理说明,可见 2.4 节内容。平均值、求和、乘积 (交互项)的说明如表 3-8 所示。

表 3-8	生成变量之常用功能	1

X1	X2	平均值	求和	乘积(交互项)
1	2	1.5	3	2
2	3	2.5	5	6
3	4	3.5	7	12
3	5	4.0	8	15

表格中共有两项, X1 和 X2, 此时求这两项的平均值或求和, 分别为第 3 列和第 4 列。 如果是乘积(分析上通常称其为交互项),则对选中项进行相乘,见表格中第 5 列。除此 之外,常用功能还包括 z 标准化、中心化、自然对数和 10 为底对数,如表 3-9 所示。

表 3-9 生成变量之常用功能 2

X1	z 标准化	中心化	自然对数	10 为底对数
1	-1.3056	-1.2500	0.0000	0.0000
2	-0.2611	-0.2500	0.6931	0.3010
3	0.7833	0.7500	1.0986	0.4771
3	0.7833	0.7500	1.0986	0.4771

先将 X1 进行 z 标准化,处理后,数据会变成平均值为 0,标准差为 1,关于 z 标准化, 其公式如下:

$$z$$
标准化 = $\frac{x - \overline{x}}{Std}$, \overline{x} 表示平均值, Std 表示标准差

中心化处理后,数据会变成平均值为 0,如表 3-9 所示的第 3 列,其公式如下: 中心化 = $x - \overline{x}, \overline{x}$ 表示平均值

通常情况下,如果数字非常大,则可使用自然对数或10为底对数,将数据进行压缩。

3.4.2 量纲处理

量纲处理指将数据压缩在一定范围内,但在部分量纲处理时,还可以对数据的方向进 行统一(如违约率数字越小越好,使用逆向化处理将其转换成数字越大越好),量纲处理是 一系列处理的统称, SPSSAU 平台共包括 15 项。生成变量功能之量纲处理如表 3-10 所示。

项	说明	注意
归一化	处理后数字介于[0,1]	无
均值化	处理后数字大小表示平均值的倍数	通常仅针对大于 0 的数据
正向化	处理后数字介于[0,1],并且数字越大越好	无
逆向化	处理后数字介于[0,1],并且数字越大越好	无
适度化	处理后越接近某个数字越好	无
区间化	处理后数字介于设定的固定区间内	无
初值化	处理后数字大小代表初值(第1个数字)的倍数	通常仅针对大于 0 的数据
最小值化	处理后数字大小代表最小值的倍数	通常仅针对大于 0 的数据
最大值化	处理后数字大小代表最大值的倍数	通常仅针对大于 0 的数据

表 3-10 生成变量功能之量纲处理

归一化时,分子为 x 值与最小值的差值,分母为最大值与最小值的差值且为一个固定 值。分子取最大值时其与分母相等,此时归一化值为 1;分子的最小值为 0 时,此时归一 化值为 0,因而进行归一化处理后,数据被压缩在[0,1]。

归一化:
$$\frac{x - x_{\text{Min}}}{x_{\text{Max}} - x_{\text{Min}}}$$

均值化时,分子为 x 值,分母为 x 值的平均值,其意义为平均值的倍数,均值化通常 仅针对全部大于 0 的数字。

均值化:
$$\frac{x}{\overline{x}}$$
, \overline{x} 表示平均值

正向化时,分子为 x 值与最小值的差值,分母为最大值与最小值的差值且为一个固定 38/187 值。当 x 取最大值时,此时分子最大即归一化值最大,为 1;当 x 取最小值时,此时分子 最小即归一化值最小,为 0,因而归一化处理后,数据被压缩在[0,1],正向化与归一化的 公式完全一致,其实际意义为,将数据压缩在[0,1],并且保持其数字的相对大小意义不变 化。

正向化:
$$\frac{x - x_{Min}}{x_{Max} - x_{Min}}$$

逆向化时,分子为最大值与 × 值的差值,分母为最大值与最小值的差值且为一个固定 值。分子取最大值时归一化值为 1,分子取最小值即为 0 时,此时归一化值为 0,因而归 一化处理后,数据被压缩在[0,1],逆向化的实际意义为,将数据压缩在[0,1],并且调换数 字的相对大小意义,如原始数字越大越差(类似负债),处理后数字变成越大越好。因而在 实际研究中,在对逆向化指标进行逆向化处理后,其会变成正向指标。

逆向化:
$$\frac{x_{\text{Max}} - x}{x_{\text{Max}} - x_{\text{Min}}}$$

适度化时, k 值为一个输入参数值, 如 k=1, 其意义为, 数字越接近于 1, 适度化后数 字越大, 适度化处理后数字均小于等于 0, 但越接近于 0 说明其离 k 值越近。

适度化: - |x - k|

区间化时, a 值和 b 值均为输入参数值, 如 a=1 且 b=2, 其意义为将数据压缩在[1,2], 区间化是归一化的通用化公式,将数据压缩在设置的范围内,并且保持其数字的相对大小 意义不变化。

区间化:
$$a + (b - a) * \frac{(x - x_{\text{Min}})}{x_{\text{Max}} - x_{\text{Min}}}$$

初值化时, 分母 x0 为原始数据的第 1 个值, 如 2000 年的 GDP 数据, 将 2000 年后的 GDP 数据与 2000 年的 GDP 数据进行对比,处理后意义为, 2000 年的 GDP 数据的倍数, 初值化通常仅针对全部大于 0 的数字。

初值化:
$$\frac{x}{x_0}$$
, x_0 表示初值

最小值化时, 分母 xMin 为原始数据的最小值, 其处理后意义为, 数据是最小值的多 少倍, 最小值化通常仅针对全部大于 0 的数字。

最小值化:
$$\frac{x}{x_{Min}}$$

最大值化时, 分母 xMax 为原始数据的最大值, 其处理后意义为, 数据是最大值的多 少倍, 最大值化通常仅针对全部大于 0 的数字。

最大值化:
$$\frac{x}{x_{Max}}$$

求和归一化时,分母为所有数据的和,其处理后意义为,数据是各数字的占比,求和 归一化通常仅针对全部大于0的数字。

求和归一化:
$$\frac{x}{\sum_{i=1}^{n} x_i}$$

平方和归一化时,分母为所有数据平方的和再开根号,其意义为,得到数字的相对大小,通常情况下,处理后数字一定小于1,平方和归一化仅针对全部大于0的数字。

平方和归一化:
$$\frac{x}{\sqrt{\sum_{i=1}^{n} x_{i}^{2}}}$$

固定值化时, FixedValue 是一个输入参数值,如其为 10,式中的分母为一个固定值, 其表示所有数据离 10 的最远距离。固定值化的实际意义为,离 10 的相对距离(处理后数 字越大越接近,数字越小越远离),处理后数据介于[0,1],0 代表远离 10,1 代表刚好为 10。

固定值化:
$$x_i = 1 - \frac{|x_i - \text{FixedValue}|}{max |x - \text{FixedValue}|}$$

偏固定值化时, FixedValue 是一个输入参数值,如其为 10,式中的分母为一个固定值, 其表示所有数据离 10 的最远距离。偏固定值化的实际意义为,离 10 的相对距离(处理后 数字越大越远离,数字越小越接近),处理后数据介于[0,1],0 代表刚好为 10,1 代表远离 10。

偏固定值化:
$$x_i = \frac{|x_i - \text{FixedValue}|}{max | x - \text{FixedValue}|}$$

近区间化时, p 和 q 是两个输入参数值, 如 p=10, q=20, 如果数据在[10,20]内, 那 么说明在该区间, 处理后数据为 1。如果数据不在[10,20]内, 处理后数字越大意味着越接 近该区间, 数字越小意味着越远离该区间, 且处理后数据介于[0,1]。

近区间化:
$$x_i = \{ \begin{array}{c} 1 - \frac{\max(p - x_i, x_i - q)}{\max(p - \min(x), \max(x) - q)} &, x_i \notin [p, q] \\ 1 &, x_i \in [p, q] \end{array}$$

偏区间化时, p 和 q 是两个输入参数值, 如 p=10, q=20, 如果数据在[10,20]内, 那 么说明在该区间, 处理后数据为 0。如果数据不在[10,20], 处理后数字越大意味着越远离 该区间, 数字越小意味着越接近该区间, 且处理后数据介于[0,1]。

3.4.3 科学计算

除上述 15 项量纲处理外, SPSSAU 平台还能提供常用的科学计算, 包括取数据的平方、 根号、绝对值、倒数、相反数、三次方和取整, 与此同时, 还可使用 SPSSAU 平台提供的 "高级公式"完成更多的科学计算, 可通过在仪表盘中依次单击【数据处理】→【生成变量】 模块找到。

3.4.4 汇总处理

生成变量之汇总处理如表 3-11 所示。X1、X2 和 X3 这三项的编号为 1 时对应的最大值为 5, 最小值为 1, 中位数为 2。其实际意义为, 如取 3 门课程的最高分、最低分及中间分。

	编号	X1	X2	Х3	最大值	最小值	中位数	计数
Ĩ	1	1	2	5	5	1	2	1
	2	2	3	3	2.5	5	6	0
	3	3	4	2	3.5	7	12	0
	4	3	5	3	4	8	15	0

表 3-11 生成变量之汇总处理

如果是计数,计算 X1、X2、X3 这三项中出现某个值(该值为输入参数值)的次数, 如计算 X1、X2、X3 三项中出现 1 的次数,那么编号为 1 时 3 个数字分别是 1、2 和 5,出 现 1 的次数为 1。其实际意义为,如对错题,若分别使用数字 1 表示正确、数字 0 表示 错误,那么统计出现 1 的次数即为统计选对的次数。

3.4.5 日期相关处理

日期相关处理包括取出日期数据的基本信息,包括年、月、日等,还涉及日期相减、 滞后处理、差分处理和季节差分等。生成变量功能之日期相关处理如表 3-12 所示。

表 3-12 生成变量功能之日期相关处理

项	说明
日期处理	取出日期数据的基本信息,包括年、月、日、周或季度
日期相减	两个日期数据相减
滞后处理	将时间序列进行滞后处理
差分处理	将时间序列进行差分处理
季节差分	将时间序列进行季节性差分处理

表 3-13 生成变量功能之日期相关处理事例

编号	日期 1	日期 2	日期 1-年	日期 1-月	日期 1- 日期 2	时间序 列数据	滞后 1 阶	差分 1 阶
1	2023-3-1	2022-11-22	2023	3	99	100	null	null
2	2023-2-28	2022-11-23	2023	2	97	98	100	-2
3	2023-2-27	2022-11-24	2023	2	95	95	98	-3
4	2023-2-26	2022-11-25	2023	2	93	97	95	2
5	2023-2-25	2022-11-26	2023	2	91	78	97	-19
6	2023-2-24	2022-11-27	2023	2	89	76	78	-2

生成变量功能之日期相关处理事例如表 3-13 所示,针对日期 1 取其年份和月份数据, 见第 4 列和第 5 列。并且日期 1 与日期 2 相减后得到两个日期的差值天数,见第 6 列。针 对时间序列数据(第 7 列),其滞后 1 阶为第 8 列,即当前日期数据是上一个日期的数据, 滞后 2 阶指当前日期数据是上两个日期的数据。差分 1 阶(第 9 列)指当前日期数据减去 上一个日期的数据,如表格中编号为 2 时,对应 98-100=-2。如果是季节周期性数据,可 先设置季节周期值后再进行差分处理,其原理与普通差分处理类似,如季节差分 1 阶,季 节周期值为 4 (一年 4 个季度则季节周期值为 4),差分是指当前日期数据-前一个季节周 期值对应的数据。

3.4.6 其他

SPSSAU 平台生成变量中还能提供包括样本编号、Box-Cox 变换、秩、缩尾处理、截尾处 理等共计 10 项功能,如表 3-14 所示。

项	说明
样本编号	如 100 个样本,编号为 1~100,提供顺序编号和随机编号两种方式
Box-Cox 变换	非正态数据转换处理的一种方式
秩	数据的秩
缩尾处理	一种异常值的处理方式
截尾处理	一种异常值的处理方式
Johnson 转换	非正态数据转换处理的一种方式
排名	数据的排名情况
相除	两个数据相除

表 3-14 生成变量功能之其他处理

例如, 共有 100 个样本, 针对该样本设置一个编号, 可以从 1~100 顺序递增(顺序 编号), 也可以随机编号 1~100, 实际研究中仅希望分析前 50 个样本, 此时可先设置样本 编号, 然后复选编号小于 50 的样本进行分析。Box-Cox 变换和 Johnson 转换均是非正态 转换的处理方式, 具体可见 2.3 节内容。

秩是指数据的排名,其与"排名"功能非常类似,但区别在于当有几个相同的排名时, 秩会取排名的平均值。如果前 3 个数字均相同,那么排名上均为 1 (升序时),而秩会取排 名的平均值,即 0.33333333。

通常在计量研究时会对异常值进行处理, 缩尾处理包括双向缩尾、上侧单向缩尾和下 侧单向缩尾。双向缩尾指, 如可将小于 2.5%分位数值设置为 2.5%分位数值, 将大于 97.5% 分位数值设置为 97.5%分位数值。上侧单向缩尾只能将大于 97.5%分位数值设置为 97.5%分 位数值。下侧单向缩尾只能将小于 2.5%分位数值设置为 2.5%分位数值(注: 此处参数默认 值为 0.05 即 5%, 研究人员可对其进行设置, 平台处理时以小于该参数值/2, 或大于 1-该 参数值/2 作为上侧或下侧的临界值标准)。

相除是指将两项数据进行相除,选择被除数和除数项即可,类似地还有相减功能。非 负平移指若数据出现小于等于 0 时,则全部加上一个"平移值",该"平移值"=数据最小值的 绝对值+参数值(注:参数值默认为 0.01),其意义为让数据全部大于 0(并且大于等于参 数值)。如果数据全部大于 0,则平台不会进行非负平移。生成变量之非负平移如表 3-15 所示。

编号	X1	X2	X1 非负平移 0.01	X2 非负平移 0.01
1	-1	2	0.01	2
2	2	3	3.01	3
3	3	4	4.01	4
4	3	5	4.01	5

表 3-15 生成变量之非负平移

表格中 X1 中出现了-1, 即小于等于 0 的数字, 且 X1 最小值为-1, 绝对值为 1, 那么

平移值=1+0.01 (0.01 为参数值),此时非负平移后的数据见第 4 列。由于 X2 全部大于 0,因此非负平移并不生效,其处理后的数据保持一致 (见第 5 列)。

3.5 标题处理

通常在数据编码、生成变量等处理之后,标题信息并不规范,如可能出现"你的性别是?" 这样的标题名称,可对其进行编辑,修改成"性别"。在 SPSSAU 平台中可单击标题进行修 改,也可批量对标题进行修改,批量修改时,可直接将已处理好的标题名称进行粘贴,一 次性完成对标题的修改。可通过在仪表盘中依次单击【数据处理】→【标题处理】模块完 成对标题名称的修改,也可通过标题框中的【…】按钮选择批量修改标题。除此之外,还 可对不需要使用的标题进行删除处理。SPSSAU 平台标题处理如图 3-12 所示。

X2 标题处理 数据标签 X3 数据编码 生成变量 X4 无效样本 异常值 X3 X4	X1		数据处理 🗸 💙	■ 批量修改
X3 数据编码 生成变量 × X2 X4 无效样本 异常值 × X3	X2	标题处理	数据标签	× X1
X5 无效样本 异常值 × X3 xx4	X3 X4	数据编码	生成变量	× X2
	X5	无效样本	异常值	× X3 × X4

图 3-12 SPSSAU 平台标题处理

第4章 问卷数据差异关系研究

在科研数据分析中,差异比较是常见的一类分析方法,如两组或多组均数、率,以及 分布的差异比较,一般采用假设检验类方法进行统计分析。在统计学中,通过检验样本统 计量之间的差异做出一般性结论,判断总体参数之间是否存在差异,这种推论过程被称作 假设检验。

本章主要介绍常用的差异比较方法,包括 t 检验、方差分析、卡方检验,非参数检验。 这四类分析方法的分析目的、对数据类型的要求及研究设计方法的注意事项如表 4-1 所示。

差异比较方法	分析目的	数据类型	研究设计方法
<i>t</i> 检验	两组样本数据均数差异比较	分组变量: X(定类) 因变量: Y(定量)	成组设计或配对设计
方差分析	两组或多组样本均数差异比较	分组变量: X(定类) 因变量: Y(定量)	完全随机、随机区组、 析因设计、正交设计等
卡方检验	两组或多组率、分布差异比较/关联性分析	行变量: X(定类) 列变量: Y(定类)	成组设计或配对设计
非参数检验	一般在以上三类方法使用条件不满足时使用	3	

表 4-1 常用差异比较方法

假设检验类分析方法分为参数检验和非参数检验两种。本章介绍的 t 检验与方差分析, 这两个方法基于总体分布为正态分布、总体方差相等的前提对总体均数差异进行检验。在 总体分布类型已知(如正态分布)的条件下,对其未知参数进行检验,这类方法称为参数 检验。

与之相反,总体分布未知或已知总体分布与检验所要求的条件不符,经数据转换也不 能使其满足参数检验的条件时,需要采用一种不依赖总体分布形式的检验方法,这种方法 不是对已知参数进行检验,而是检验总体分布位置是否相同,称为非参数检验。卡方检验、 非参数秩和检验属于非参数检验。

假设检验的基本思想是,先提出原假设(一般假设两个统计量相等,本节以均值为例, 即假设均值相等,记为 H0)和备择假设(两个均值不等,记为 H1),要求原假设与备择假 设互斥;然后利用小概率事件原理对原假设进行反证,当概率 p 小于显著性水平 α 时则拒 绝原假设,从而选择备择假设。统计学上 α 通常取 0.05,有时也可以取 0.01 或 0.001。在 结果解释上,当拒绝原假设选择备择假设(p<0.05)时,通常解释为两组均值的差异具有 统计学意义,或简写为差异显著;反之,当接受原假设(p>0.05)时,通常解释为两组均 值无差异或差异不显著。

4.1t 检验

t 检验用于两组定量数据资料的均数差异比较,在一般背景资料分析、基线分析中应 用较广泛。数据资料要求为二水平定类数据以用于分组,因变量要求为定量数据(连续型 数据)。例如,某研究欲比较不同性别新生儿体重有无差别,此时性别为定类数据,其用于 分组,体重为定量数据,作为目标变量或因变量。

由于 t 检验要求数据具有正态性,因此在 t 检验前应当先检验正态分布的情况。一般 来说,只要不是严重偏态,t 检验就基本适用,否则可考虑对数据进行转换使其近似正态 后再做检验,或者使用对应的非参数秩和检验方法,这在本章 4.4 节将具体介绍。 t 检验包括 3 种类型: 单样本 t 检验、配对样本 t 检验, 以及独立样本 t 检验。3 种 t 检验 在选择时的注意事项如表 4-2 所示。

t 检验	研究目的	适用条件	常见应用场景举例	SPSSAU 平台分析路径
单样本 <i>t</i> 检验	一组数据的均数 是否与给定常数之 间存在差异	单样本服从正 态分布	研究人员考察某地 新生儿的平均体重与 常模体重的差异	【通用方法】→【单样本 t 检验】 【实验/医学研究】→【概 要 t 检验】
配对样本	两组配对数据均	差值数据服从	体育疗法前后测定	【通用方法】→【配对 t
t 检验	数的差异比较	正态分布	的舒张压有无差异	检验】
独立样本 t 检验	两组独立数据均 数间的差别比较	两样本分别服 从正态分布且要 求方差齐性	干预组与对照组人 群的心肌血流量有无 差别	【通用方法】→【t 检验】 【实验/医学研究】→【概 要 t 检验】

表 4-2 3 种 t 检验在选择时的注意事项

在 SPSSAU 平台的【通用方法】【实验/医学研究】模块下,共提供了 4 个独立的模块 来完成各种类型的 t 检验。

4.1.1 正态分布与方差齐性

t 检验要求进行差异比较的数据必须是定量数据,且要求数据服从或近似于正态分布, 独立样本 t 检验还要求两组数据的方差齐性。

1.正态分布

数据服从或近似于正态分布,是 t 检验的基础条件。实际分析时可以用图示法或显著 性检验法判断该条件是否满足。图示法通过绘制直方图、正态 PP 图、QQ 图、核密度图等 对正态分布进行直观判断,此类方法判断标准较宽松,有一定的主观性;显著性检验法则 主要包括 Kolmogorov-Smirnov 检验、Shapiro-Wilk 检验和 Jarque-Bera 检验等。正态分布 检验具体介绍和分析见本书 2.3 节内容。

样本数据完全服从正态分布过于理论化,在实际分析中数据出现一定偏态较常见。 t 检验、方差分析等参数检验方法对正态分布要求相对稳健,这些方法本身对轻微的偏态具 有一定"抗性",不会影响其结果。

2.方差齐性

方差齐性也称方差齐次或简称为方差齐, 该检验假设两组数据的总体方差相同, 在 SPSSAU 平台中利用 F 检验进行统计推断。按 α=0.05 显著性水平, 当 F 检验的概率 p 大于 0.05 时认为两组数据具有方差齐性, 当 p 小于 0.05 时认为两组数据的方差不齐。

独立样本 t 检验要求数据具有正态性,同时还要求两组数据具有方差齐性,方差不齐时需要对 t 检验的结果进行校正。因此进行 t 检验时,应根据方差齐性检验的结论,选择 t 检验或校正后的 t 检验显著性结果。

通过在仪表盘中依次单击【通用方法】→【t 检验】模块进行独立样本 t 检验,平台会 自动判断方差齐性条件,并输出对应的 t 检验结果。如果用户需要单独做方差齐性检验, 可通过在仪表盘中依次单击【通用方法】→【方差】模块来实现。

本节结合具体案例进一步介绍方差齐性检验的操作与结果解读。

【例 4-1】某研究收集到 100 例儿童的腰围(厘米)数据,其中肥胖组 50 例,用数字 1 表示组别,另外 50 例为非肥胖组,用数字 0 表示组别,试分析两组数据的方差是否相等。案例数据为模拟获得,仅用于示范分析方法的应用,数据文档见"例 4-1.xls"。

1) 方差齐性检验

数据读入平台后,在仪表盘中依次单击【通用方法】→【方差】模块,将【组别】变 量拖曳至【X(定类)】分析框中,将【腰围】变量拖曳至【Y(定量)】分析框中,选择下拉列 表中的【方差齐检验】。方差齐检验操作界面如图 4-1 所示,最后单击【开始分析】按钮。

2) 结果分析

两组数据的方差齐检验分析结果如表 4-3 所示。F=2.576, p=0.112>0.05, 表明两组数据具有方差齐性。

□ 通用	方法 🗸	开始分析	前 方差齐检验 ∨ ②	Ō
频数	分类汇总	组别	方差齐检验 X(Welch anova	定类)
描述	交叉(卡方)	腰围	Brown-Forsythe anova	
相关	线性回归		Y	(定量)
方差	t检验			
单样本t检验	配对t检验	e		

图 4-1 方差齐检验操作界面

表 4-3 肥胖与非肥胖组方差齐检验分析结果

ாக	组别(标准	F	n	
坝	非肥胖组(n=50)	肥胖组(<i>n</i> =50)	Г	ρ
腰围	3.95	5.41	2.576	0.112

4.1.2 t 检验分析步骤

在统计学中, t 检验的步骤是:研究者首先提出原假设(H0,两均值相等)和备择假设(H1,两均值不相等);然后用数据构造并计算 t 统计量,并通过 t 分布计算概率 p;最后利用 p 推断原假设是否成立,从而完成整个假设检验流程。在实际科研数据分析中,要综合 t 检验类型、适用条件,总结 t 检验分析方法应用的分析思路。t 检验一般分析步骤如图 4-2 所示。



图 4-2 t 检验一般分析步骤

(1) 根据科研设计与研究目的,选择单样本 t 检验、配对样本 t 检验、独立样本 t 检 验中的一种,如配对设计选择配对样本 t 检验,成组设计选择独立样本 t 检验。

(2) 对两组数据进行正态分布检验,满足条件时使用 t 检验,非正态分布(或严重偏离正态分布)时考虑先进行数据转换再做 t 检验或采用非参数秩和检验方法(单样本 t 检验和配对样本 t 检验对应的是 Wilcoxon 检验、独立样本 t 检验对应的是 Mann-Whitney 检验)。

(3) 如果研究目的是比较成组设计的两组均数的差异,即选择独立样本 t 检验时,还 应对两组数据是否具有方差齐性进行检验。

(4) 由统计软件工具执行 t 检验, 如遇到方差不齐时, 则对 t 检验进行校正, 校正后的结果一般由统计软件或统计平台自动完成。

(5)对结果进行分析, 撰写结论。按 α=0.05 水平, 当 t 检验概率 p<0.05 时, 认为 两组均数差异有统计学意义;反之, p>0.05 说明两组均数无差别。

4.1.3 单样本 t 检验

分析一组数据的均数是否与给定常数存在差异时,要用随机抽取的单个样本均数和已 知总体均数进行比较,观察该组样本与总体均数的差异性。例如,研究人员考察某地新生 儿的平均体重与同类大型研究常模体重的差异,要求该组样本数据服从正态分布,且没有 明显的异常数据。

【例 4-2】已知某地区 12 岁男孩平均身高为 142.5 厘米 (常模数据)。1973 年某市测量了 120 名 12 岁男孩的身高,试分析该市 12 岁男孩与该地区 12 岁男孩身高的平均值是否相等。案例数据来源于卢纹岱 (2006),数据文档见"例 4-2.xls"。

1) 数据与案例分析

数据文档中的"身高"为定量数据,用于记录 120 名男孩的身高数据,研究目的是比较 120 个身高数据的平均值与"142.5"的差异,从分析目的上选择单样本 t 检验。

2) 正态分布检验

在仪表盘中依次单击【通用方法】→【正态性检验】模块, 在输出结果中选择 Shapiro- Wilk 检验, 得出 p=0.93>0.05, 说明该组数据服从正态分布, 满足单样本 t 检验 对数据正态性的要求。

3) t 检验

数据文档读入平台后, 在仪表盘中依次单击【通用方法】→【单样本 t 检验】模块, 将【身高】变量数据拖曳至右侧【分析项(定量)】分析框中, 在分析框顶部的数值框输入 对比数字【142.5】。单样本 t 检验操作界面如图 4-3 所示, 最后单击【开始分析】按钮。

口 通用	月方法 🗸	开始分析	142.5
方差	t检验	自高	
单样本地检验	配对检验	210	分析项(定量)
正态性检验	非参数检验		

图 4-3 单样本 t 检验操作界面

4) 结果分析

120 名男孩的平均身高单样本 t 检验分析结果如表 4-4 所示。第 2~6 列为常见统计指标, 第 7、8 列为 t 统计量和概率 p 值。

表 4-4 120 名男孩的平均身高单样本 t 检验分析结果

名称	样本量	最小值	最大值	平均值	标准差	t	р
身高	120	125.900	160.900	143.048	5.821	1.032	0.304

120 名 12 岁男孩的身高平均值为 143.048 厘米, 与该地区 12 岁男孩身高的一般水平 142.5 厘米相比, t=1.032, p=0.304>0.05, 差异无统计学意义。

4.1.4 配对样本 t 检验

配对设计常见的形式是自身配对或非自身配对。例如,对同一研究对象在试验干预前、 干预后的某指标数据进行比较,同一研究对象在两个不同部位所测定的某指标数据进行比 较,或者同源配对或条件相近者异体配对数据结果的比较。

配对设计所获两组数据,因其本身具有相关性,所以需要使用针对性的分析方法。配 对样本 t 检验可用来检验有配对或相关关系的两组数据均数是否相等。统计分析时取两组 配对数据的差值与数字"0"进行差异比较。使用条件上,要求配对的两组样本数据差值服从 正态分布。

【例 4-3】10 名高血压患者在实施体育疗法前后测定舒张压,请判断体育疗法对降低 血压是否有效。案例数据来源于卢纹岱(2006),数据文档见"例 4-3.xls"。

1) 数据与案例分析

实施体育疗法前、后的舒张压数据分别为"治疗前""治疗后"两个定量数据,本例属于自身前后配对设计,综合考虑采用配对样本 t 检验进行差异比较。

2) 正态分布检验

通过在仪表盘中依次单击【数据处理】→【生成变量】→【相减(Minus)】模块,计算得 到 差 值 数 据,读取【通 用 方 法】→【正 态 性 检 验】 模 块 的 Shapiro-Wilk 检 验, p=0.225>0.05,说明差值数据服从正态分布,满足 t 检验对正态分布的要求。

□ 通用	方法 🗸		开始分析	<u></u>	
频数	分类汇总		治疗后		
描述	交叉(卡方)				
相关	线性回归				
方差	t检验				配对1(定量)
单样本地检验	配对检验	e			
正态性检验	非参数检验				
目间类	研究 🗸				
	W120 +		治疗前		記数2(完局)
信度	效度				同しからそくがに回り

图 4-4 配对 t 检验操作界面

3) 配对 t 检验

数据文档读入平台后,在仪表盘中依次单击【通用方法】→【配对 t 检验】模块,将

【治疗后】变量拖曳至右侧【配对 1(定量)】分析框中,将【治疗前】变量拖曳至右侧【配 对 2(定量)】分析框中。配对 t 检验操作界面如图 4-4 所示,最后单击【开始分析】按钮。

4) 结果分析

10 名患者实施体育疗法前后舒张压数据配对 t 检验分析结果如表 4-5 所示。表格中第 2~3 列为配对前后对舒张压的统计, 第 4 列为配对前后的差值, 最后两列为 t 统计量和 p 值。 在进行体育疗法前, 患者的舒张压为(119.50±10.07), 在进行体育疗法后, 患者的 舒张压为(102.50±11.12), 显然在进行体育疗法后, 患者的舒张压出现了下降。

表 4-5 10 名患者实施体育疗法前后舒张压数据配对 t检验分析结果

冬玫	配对(平均位	羊店	+	2	
白柳	配对 1	配对 2	一左區	Ĺ	μ
体育疗法后 配对 体育疗法前	102.50 ± 11.12	119.50 ± 10.07	-17.00	-5.639	0.000**

注: **p<0.01。

差值为-17.00, 说明治疗后舒张压下降了 17 个单位, 如果差值与数字"0"相比差异显 著,则说明配对的两个样本均值差异有统计学意义。t=-5.639, p<0.01, 按 α=0.01 水平, 认为体育疗法治疗后高血压患者的舒张压下降明显, 与治疗前相比差异有统计学意义, 该 疗法对降低高血压患者的舒张压有效。

4.1.5 独立样本 t 检验

独立样本 t 检验采用成组设计来检验两组数据的各自总体均数是否相等。例如, 研究 人员想比较不同性别人群的抑郁评分是否有显著差异, t 检验仅可对比两组独立成组数据 的差异, 如果为三组或多组, 则使用方差分析。

独立样本 t 检验要求两组数据分别服从正态分布,且两组数据具有方差齐性。所谓方 差齐性,即要求两组数据的总体方差相同。此处注意,如果方差出现不齐的情况, SPSSAU 平台会自动完成对方差不齐的校正,可输出校正后的 t 检验结果,所以遇到方差不齐时仍 可继续读取 t 检验结果。

【例 4-4】沿用【例 4-1】的数据, 某研究收集到 100 例儿童的腰围(厘米)数据, 其中肥胖组 50 例, 另外 50 例为非肥胖组。数据文档见"例 4-4.xls", 试分析肥胖组与非肥 胖组儿童的腰围有无不同。

1) 数据与案例分析

"组别"变量编码 1 表示肥胖组,编码 0 表示非肥胖组,"腰围"记录两组儿童的腰围数据,为定量数据。本例为成组设计,目的是比较两组儿童腰围的差异,应采用独立样本 t 检验。

2) 正态分布与方差齐性条件判断

通过在仪表盘中依次单击【通用方法】→【正态性检验】模块,在输出结果中选择 Shapiro-Wilk 检验,得出两组数据的显著性 p 值分别为 0.11、0.971,均大于 0.05,说明数 据服从正态分布。方差齐性已在【例 4-1】中讨论过,F=2.576, p=0.112>0.05,表明两组 数据具有方差齐性。

3) t 检验

数据文档读入平台后,在仪表盘中依次单击【通用方法】→【t 检验】模块。将【组 别】变量拖曳至右侧【X(定类,仅两组)】分析框中,将【腰围】变量拖曳至右侧【Y(定量)】 分析框中。独立样本 t 检验操作界面如图 4-5 所示,最后单击【开始分析】按钮。

□ 通用	方法 🗸		开始分析	<u>î</u> Q
频数	分类汇总		组别	X(定类,仅两组)
描述	交叉(卡方)		腰围	
相关	线性回归			Y(定量)
方差	t检验			
单样本t检验	配对t检验	Ð		

图 4-5 独立样本 t 检验操作界面

4) 结果分析

肥胖与非肥胖组腰围数据独立样本 t 检验分析结果如表 4-6 所示。

表 4-6 肥胖与非肥胖组腰围数据独立样本 t 检验分析结果

组别(平均值土标准差)	腰围
非肥胖组(n=50)	52.52 ± 3.95
肥胖组(n=50)	61.02 ± 5.41
t	-8.965
p	0.000**

注: ** p<0.01。

肥胖组儿童的腰围为(61.02±5.41) 厘米,非肥胖组儿童的腰围为(52.52±3.95) 厘 米。t=-8.965, p<0.01, 按 α=0.01 显著性水平,有理由认为肥胖组与非肥胖组儿童的腰围 差异显著。

4.2 方差分析

t 检验针对的是两组数据均值的比较,如果是三组及更多组数据均值的比较,则需要 采用方差分析。方差分析的基本思想是对数据的总变异进行分解,将各部分的方差与误差 相比较,从而判断因素或交互作用的统计学意义。

本节主要介绍只有一个因变量 Y 的方差分析。只有一个自变量 X 的方差分析称为单因 素方差分析,有两个自变量 X 的方差分析则称为双因素方差分析,当自变量因素个数超过 两个时,其方差分析统称为多因素方差分析。

4.2.1 方法概述

1.原理与概念

方差分析的基本思想是对误差的分解,总误差被分解为组间误差和组内误差,组内误 差用于估计抽样的随机误差,组间误差可能是由抽样的随机误差造成的,更多的是组间自 身差异的系统误差导致。我们需要证明的是系统误差不等于 0,所以组间误差除以组内误 差大到一定程度时可认为组间效应显著。为消除个案数的影响,给分母和分子同时引进自 由度,构造 F 统计量,根据 F 分布计算概率 p,利用 p 推断原假设是否成立。所以,方差 分析等价于 F 检验, 中间计算的统计量为 F 统计量。

方差分析中的常用术语概念。

(1)因素:方差分析是通过组间差异比较来推断自变量对因变量的影响的,此处的自 变量通常也称作因素或因子。例如,超市某商品销售量的影响因素有商品价格、货架摆放 位置、广告宣传等。

(2) 水平: 方差分析中的因素要求是定类数据, 因素的不同类别取值称作水平。例如, 商品在货架的摆放位置有低、中、高三个水平, 商品价格有原价、促销活动价两个价格水 平。

(3)组合或单元: 各因素不同水平的交叉称为组合或单元, 如某商品摆放在货架中间 层且以促销价格销售时更有利于售卖。

(4)主效应:某一因素各单独效应的平均效应,即某一因素各水平之间的平均差别。 一个因素的主效应显著,意味着该因素的各个水平在其他因素的所有水平上的平均数存在 差异。

(5) 交互作用:反映两个或两个以上因素相互依赖制约、共同影响因变量的变化。如 果一个因素对因变量的影响会因另一个因素的水平不同而有所不同,则可以说这两个变量 之间具有交互作用。

2.方差分析的适用条件

方差分析要求因变量 Y 为定量数据(连续型数据), 自变量 X 为定类数据(分类数据)。 如果需要控制混杂干扰,则要加入协变量,协变量也要求为定量数据。

进行方差分析需要数据满足以下两个基本前提。

- (1) 各观测变量数据总体要服从正态分布。
- (2) 各观测变量数据总体要满足方差齐性。

理论上,数据需要满足以上两个条件才能进行方差分析。在实际分析时,可根据试验 设计方法、数据样本量、行业一般要求,以及检验方法的结果综合讨论并对两个条件做出 判断,通常方差分析本身较稳健,因此可适当放宽两个条件,具体结合文献资料而定。

对于非正态分布数据,可以考虑做正态转换使数据满足正态要求后继续进行方差分析, 或当严重偏离正态分布时采用非参数秩和检验作为替代;对于方差不齐的情况,可使用非 参数秩和检验,同时可选择使用更为稳健的 Welch 方差分析或 Brown-Forsythe 方差分析, 其中 Welch 方差分析较常用。

正态分布条件的判断,可参考第 2 章的内容。非参数检验可查阅本章 4.4 节的相关内容。

3.事后多重比较

单因素方差分析的 F 检验 p<0.05,即总体上组间差异显著时,需要继续对因素各水平间的差异进行两两配对的多重比较,双因素或多因素方差分析在交互作用不显著的前提下,也需要继续针对有显著影响的因素主效应进行多重比较。

例如,对四所中学(全国重点中学、市重点中学、区重点中学、一般中学)某年级数 学统一测试成绩进行方差分析,结果显示不同中学间成绩有显著差异。这是总体结论,究 竟哪些中学之间有差异?如全国重点中学与市重点中学间成绩有无差别,市重点中学和一 般中学间成绩有无差别,需要进一步分析。这一过程所采用的方法就称为多重比较,通过 对总体均值之间的两两配对进行比较进一步分析到底哪些均值间存在差异。

事后多重比较的方法有多种, SPSSAU 平台【事后多重比较】模块中共提供了 LSD、 Scheffe、Tukey、Bonferroni 校正、Sidak、Tamhane T2、SNK、Duncan 等 8 种常用方法, 如表 4-10 所示。

多重比较方法	适用场景	其他说明
LSD	一般在事前有明确要比较的两组差异时使用	对差异较敏感
LSD	比较次数过多时不建议使用	容易犯I类错误
Scheffe	各组样本量不等时使用	相对较保守
Tukey	各组样本量相等时使用	校正I类错误
Bonferroni 校正	较流行,待比较的组别数量较少时使用 校正方式: a/比较次数,如 3 次比较,则 p 与 0.05/3=0.0167 进行比较	组数较大时结果相对较 保守
Sidak	是对 LSD 法的校正,对比组别较少时使用	比 LSD 法保守
Tamhane T2	如果方差不齐但希望进行多重比较,则使用此方法	方差不齐时使用
SNK	是对Tukey 法的修正,应用广泛	不及 LSD 法灵敏
Duncan	是对 SNK 法的修正	在农业研究中使用广泛

表 4-10 常用多重比较方法

冯国双指出,如果比较组数不是很多(如3组),则 Tukey 法和 Bonferroni 校正法均 可作为首选。如果比较组数较多(如4组以上),建议首选 Tukey 法(包括各组例数不等 的情况,因为大多数软件在例数不等时给出的是 Tukey-Kramer 法)。张文彤和邝春伟指出, 如果需要进行任意两组之间的比较而各组样本含量又相同时,可选用 Tukey 法;当样本含 量彼此不同时,可选用 Scheffe 法。

4.差异比较结果的字母标记

差异的字母标记法是指使用不同字母标记并注释多重比较的差异结果,如果对比的组 之间有相同字母即表示差异不显著,如果对比的组之间字母不同则表示两组间差异显著。 SPSSAU 平台默认提供 0.01 和 0.05 两种不同显著性水平时的字母标记法结果, 0.05 水平时 使用小写字母标识, 0.01 水平时使用大写字母标识,具体结果通过在仪表盘中依次单击 【进阶方法】→【事后多重比较】模块来完成。

4.2.2 方差分析类型的选择

在进行数据分析时采用何种类型的方差分析,可参考试验设计方法与研究分析目的综 合决定。

1.试验设计方法与方差分析类型

通常情况下, 方差分析与科学试验设计相辅相成。在考虑使用方差分析方法时, 应结 合研究分析目的及科学试验所用的设计方法进行综合考虑。常见试验设计方法与方差分析 类型如表 4-11 所示。

试验设计方法	方差分析类型选择
完全随机设计	单因素方差分析
随机区组设计	双因素方差分析
析因设计	双因素或多因素方差分析(可考察交互作用)
拉丁方设计	三因素方差分析(不考察交互作用)
正交设计	一般为多因素方差分析(可考察交互作用)

表 4-11 常见试验设计方法与方差分析类型

(1) 完全随机设计有时也被称为单因素设计,是将被试随机化分配到各处理组中,仅 考察一个处理因素对试验指标的影响。统计方法采用单因素方差分析,如果该因素有统计 学意义,则继续进行各组均数间的多重比较。

(2)随机区组设计先将被试划分为区组,再将每一区组的被试随机分配到各个处理组中,每个区组的样本量等于处理组的个数。随机区组设计可采用双因素方差分析进行结果分析,一般认为区组因素是次要因素,通常不需要考察区组与处理因素间的交互作用。

(3) 析因设计也叫作全因子设计,是试验中所涉及的全部因素的各水平全面组合形成 的不同试验组合,每个试验组合要进行两次或两次以上的独立重复实验。它不仅可检验每 个因素各水平间的差异,而且可检验各因素间的交互作用。统计方法采用双因素方差分析 或多因素方差分析,可考察因素间的交互作用。

(4) 拉丁方设计是随机区组设计的扩展,如果研究涉及一个处理因素和两个需要控制的区组因素,每个因素的水平数相等,则可采用拉丁方设计。统计分析方法采用多因素方 差分析,但需要注意拉丁方设计不能考察交互作用。

(5)试验因素较多且希望考察交互作用,用较少的试验次数获得较佳的试验结果,此 时可采用正交设计。统计方法采用多因素方差分析,且根据正交试验方案的不同可考察因 素间的交互作用。

2.研究分析目的与方差分析类型

科研数据分析中的差异关系研究是错综复杂的,单个因素的分析往往作为基础性工作, 更多情况下需要综合考虑多个因素的关系或联系。方差分析可以是一个因素,也可以是多 个因素,还可以分析因素间的交互作用,因此应用范围较广泛。

根据研究分析目的的不同,可使用单因素方差分析、双因素方差分析、多因素方差分 析及协方差分析。有两个及以上因素时,又可分为有交互作用的方差分析和无交互作用的 方差分析。研究分析目的与方差分析类型如表 4-12 所示。

研究分析目的	方差分析类型	SPSSAU 平台中的分析路径
单个白娈昰对田孪昰的影响	单因麦方差分析	【通用方法】→【方差】
十十日文星州四文星的影响	十四家乃左方仍	【进阶方法】→【事后多重比较】
西久白亦是对田亦是的影响	无交互作用的双因素方差分析	【进阶方法】→【双因素方差】
网丨日文里刈囚文里的影响	有交互作用的双因素方差分析	【进阶方法】→【事后多重比较】
夕太白亦昌对田亦昌的影响	无交互作用的多因素方差分析	【进阶方法】→【三因素方差】
多十日文里对囚文里的影响	有交互作用的多因素方差分析	【进阶方法】→【事后多重比较】
有协变量的影响	协方差分析	【进阶方法】→【协方差】

表 4-12 研究分析目的与方差分析类型

SPSSAU 平台在【通用方法】【进阶方法】两个功能下,共提供了 6 个独立的模块来完成各种类型的方差分析。

4.2.3 单因素方差分析

只有一个分类型自变量 X 和一个因变量 Y 的方差分析称为单因素方差分析,如要检验 四种饲料喂猪对猪体重增加值的均值是否相等,只涉及"饲料类型"一个类别因子,就是单 因素方差分析。从试验设计的角度,一般采用的是完全随机设计。

1.分析思路

单因素方差分析思路如图 4-8 所示。

1) 适用条件判断

一般在单因素方差分析前,应检验各组数据是否满足正态分布的要求,以及是否满足 方差齐性的要求。

2) 分析方法选择

根据上一步正态分布和方差齐性检验的结果,选择合适的分析方法。如果同时满足正态分布与方差齐性,则进行单因素方差分析;如果方差不齐,则选择 Welch 检验或 Brown-Forsythe 检验对方差分析结果进行校正;如果数据非正态分布,则考虑转换数据使 其满足正态分布或当严重偏态时考虑采用 Kruskal-Wallis 非参数秩和检验作为替代分析方法。

方差分析、Welch 检验及 Kruskal-Wallis 非参数秩和检验在结果解释和分析时, 若检验的概率 p<0.05, 则认为差异显著或因素对因变量的影响有统计学意义; 反之, 若 p>0.05, 则认为因子的影响无意义。



图 4-8 单因素方差分析思路

3) 多重比较

方差分析单因素主效应结果有显著性后,可根据具体情况选择合适的多重比较方法进行组间的两两比较。常见的多重比较方法 LSD、Scheffe、Tukey、Bonferroni 校正、Sidak 等方法可用于方差齐性的情形。当方差不齐时,可使用 Tamhane T2、Games-Howell 等方法。如果采用 Kruskal-Wallis 非参数秩和检验,则可采用 Nemenyi 法进行多重比较。

对照(赋值 0)	低剂量(赋值 1)	高剂量(赋值 2)
6.82	5.66	2.13
5.73	4.82	2.71
6.46	5.53	2.50
7.36	4.98	2.67
7.62	4.40	3.60
7.77	4.18	3.36
7.90	4.07	2.33
7.89	4.11	2.85
7.60	5.00	3.12
6.95	4.30	2.42

表 4-13 不同剂量 DON 染毒后幼鼠的 || 型胶原含量

2.实例分析

【例 4-7】研究显示 DON(脱氧雪腐镰刀菌烯醇)可能对幼鼠关节软骨代谢产生影响。 将 30 只健康幼鼠完全随机分配至对照组、低剂量组和高剂量组,每组 10 只。高、低剂量 组分别给予 0.25 微克/克、0.06 微克/克的 DON,对照组给予相同容量生理盐水灌胃,连 续 80 天后用免疫组法检测幼鼠软骨内的 II 型胶原含量,含量降低提示关节软骨损伤。不 同剂量 DON 染毒后幼鼠的 II 型胶原含量如表 4-13 所示。试分析 DON 对关节软骨代谢是 否存在影响。案例数据来源于李晓松(2017),并对数据进行了修改及编辑,数据文档见 "例 4-7.xls"。

1) 数据与案例分析

数据文档中的"group"为分组组别,编码 0 表示对照组,编码 1 表示低剂量组,编码 2 表示高剂量组。本例为完全随机设计,多组均值的差异比较应考虑使用单因素方差分析。

2) 正态分布与方差齐性条件判断

通过在仪表盘中依次单击【通用方法】→【正态性检验】模块进行正态检验,三组胶 原含量数据服从整体分布 (p 均大于 0.05)。通过依次单击【通用方法】→【方差】→【方 差齐检验】模块进行方差齐性检验, F=1.160, p=0.329>0.05, 说明数据满足方差齐性要求。 三组幼鼠胶原含量方差齐性检验如表 4-14 所示。

表 4-14 三组幼鼠胶原含量方差齐性检验

īБ		F	n		
坝	对照 (n=10)	低剂量(n=10)	高剂量(n=10)	ľ	p
胶原含量	0.71	0.58	0.47	1.160	0.329

本例数据满足正态分布、方差齐性要求,根据上述分析思路,接下来进行单因素方差分析。

3) 单因素方差分析

在仪表盘中依次单击【通用方法】→【方差】模块,将组别变量【group】拖曳至右侧【X(定类)】分析框中,将因变量【胶原含量】拖曳至右侧【Y(定量)】分析框中。选择下 拉列表中的【方差分析】,最后单击【开始分析】按钮。单因素方差分析操作界面如图 4-9 所示。

口 通月	用方法 🗸	开始分析	方差分析 >
频数	分类汇总	droup	¥/宁米)
描述	交叉(卡方)	gioup	AVEX/
相关	线性回归	胶原含量	
方差	t检验		Y(定量)

图 4-9 单因素方差分析操作界面

单因素方差分析结果如表 4-15 所示,其中平方和、自由度、均方为 F 检验中间计算 的过程值,通常可不做解读分析,主要关注的结果是 F 统计量值与显著性 p 值。

表 4-15 单因素方差分析结果

项	差异	平方和	自由度	均方	F	р
	组间	99.152	2	49.576	139.375	0.000
胶原含量	组内	9.604	27	0.356		
	总计	108.756	29			

在科研写作时,常见用于结果报告的表。对照、低剂量、高剂量组胶原含量的差异比较如表 4-16 所示。

group(平均值土标准差)	胶原含量/(微克/克)
对照 (n=10)	7.21 ± 0.71
低剂量(n=10)	4.71±0.58
高剂量(n=10)	2.77±0.47
F	139.375
p	0.000**

表 4-16 对照、低剂量、高剂量组胶原含量的差异比较

注: ** p<0.01。

三组胶原含量依次为:对照组(7.21±0.71)微克/克、低剂量组(4.71±0.58)微克/克、 高剂量组(2.77±0.47)微克/克。方差分析结果显示,F=139.375,p<0.01,在 α=0.01 水 平下,表明不同剂量 DON 分组的胶原含量差异具有统计学意义,或者说不同剂量 DON 对 关节软骨损伤有影响。

4) 事后多重比较

单因素方差分析在主效应显著的前提下,即当 F 检验的 p 值小于 0.05 时要继续考察因 素各水平之间对因变量影响的差异,可以通过【事后多重比较】模块来完成。

在仪表盘中依次单击【进阶方法】→【事后多重比较】模块,将组别自变量【group】 拖曳至右侧【X(定类)】分析框中,将因变量【胶原含量】拖曳至右侧【Y(定量)】分析框中。 选择本例多重比较方法下拉列表中的【Bonferroni校正】,同时勾选【字母标记法】复选框。 如果在研究中还需要报告多重比较结果的效应量,则要勾选【效应量】复选框,平台会输 出 Cohens'd 值效应量,一般可根据研究需要进行输出和分析解释,本例不做选择。事后 多重比较操作界面如图 4-10 所示,最后单击【开始分析】按钮。

∅ 进阶	方法 🗸	开始分析	亩	Bonferroni校正 ✓ マ 字母标记法
逐步回归	二元Logit	group		X(定类)
多分类Logit	有序Logit	· · · · · · · · · · · · · · · · · · ·		
事后多重比较	偏相关	<u></u> 取 尿 占 重		Y(定量)
典型相关	双因素方差			

图 4-10 事后多重比较操作界面

【事后多重比较】模块输出的结果比较丰富,包括方差分析及多重比较结果。本节主要介绍多重比较结果。三组幼鼠胶原含量差异检验事后多重比较(Bonferroni 校正)如表 4-17 所示。

表 4-17 三组幼鼠胶原含量差异检验事后多重比较(Bonferroni 校正)

项	(1) 名称	(J)名称	(1) 平均值	(J)平均值	差值(1-J)	р
	对照	低剂量	7.210	4.705	2.505	0.000**
胶原含量	对照	高剂量	7.210	2.769	4.441	0.000**
	低剂量	高剂量	4.705	2.769	1.936	0.000**

注: ** *p<*0.01

高剂量组的胶原含量显著低于其他两组(Bonferroni p<0.01),低剂量组的胶原含量显 著低于对照组(Bonferroni p<0.01)。 幼鼠胶原含量两组间差异比较的字母标记结果如表 4-18 所示。表中包括显著性水平 α=0.05 和 α=0.01 两个标准的字母标记。

分析项	选项	平均值	字母标记 (0.05 水平)	字母标记 (0.01 水平)
	对照	7.21	а	А
胶原含量	低剂量	4.71	b	В
	高剂量	2.77	с	С

表 4-18 幼鼠胶原含量两组间差异比较的字母标记结果

注: 0.05 水平时使用小写字母标记, 0.01 水平时使用大写字母标记。

若各水平字母相同则表示差异不显著,若各水平字母不同则表示差异有统计学意义。 例如,在 α=0.05 水平下,对照组标记字母 a,而其他两组标记字母 b 和 c,3 个不同字母 表明对照组、低剂量组、高剂量组的胶原含量两两之间的差异有统计学意义。

4.2.4 双因素及多因素方差分析

有两个及以上分类型自变量 X 和一个因变量 Y 的方差分析称为双因素或多因素方差分 析,它比单因素方差分析考虑了更多因素对结果的影响,相对而言比较贴合实际的使用需 求。例如,研究文章主题与生字密度对小学生阅读理解的影响,此时有文章主题和生字密 度两个因素,双因素方差分析可同时分析文章主题和生字密度对阅读理解的影响,还可以 分析两者的交互作用对阅读理解是否有影响。

1.分析思路

常见的随机区组设计、两因素析因设计在数据分析时,可采用双因素方差分析。双因素方差分析的分析思路如图 4-11 所示,同样也适用于多因素方差分析。



图 4-11 双因素方差分析的分析思路

1) 适用条件判断

双因素及多因素方差分析要求各组数据具有正态分布及方差齐性,在有重复试验数据 资料的情形下,一般正常进行正态分布和方差齐性的检验。此外的一些情形往往和试验设 计方法紧密联系。在实际研究分析中,要结合方差分析自身的稳健性和试验设计,一般会 假设数据基本满足正态分布和方差齐性的要求。

2) 是否考察交互作用

从专业经验及研究目的出发,决定试验设计是否考察交互作用。在双因素方差分析中, 根据是否考察交互作用,其对统计结果的解读也相应不同。 3) 效应显著性检验

对因素主效应、交互作用的效应进行显著性检验,如果有考察交互作用,则优先解读 交互作用是否有统计学意义。当 F 检验的 p 值小于 0.05 时,认为主效应或交互作用对因变 量的影响有统计学意义。

4) 进一步分析

交互作用显著时,应进一步考察在控制 A 因素后 B 因素不同水平对因变量 Y 的影响差 异,或反过来研究控制 B 因素后 A 因素不同水平对因变量 Y 的影响差异,这一过程叫作简 单效应分析。

交互作用不显著或试验设计中不考察交互作用时,若主效应有显著性则继续对其组间 差异进行多重比较、操作及结果解释,和前面单因素方差分析多重比较一致。

2.实例分析

【例 4-8】为了研究一种含有荷尔蒙的新药是否真能缓解抑郁症状,每组随机选择 6 名抑郁症患者(男女各 3 名):第 1 组患者接受安慰剂,第 2 组患者接受中等剂量新药, 第 3 组患者接受大剂量新药,3 组患者接受治疗后的抑郁水平如表 4-19 所示,试分析该药 的治疗效果。数据来源于李志辉和杜志成(2018),数据文档见"例 4-8.xls"。

性别	安慰剂 dose=1	中等剂量新药 dose=2	大剂量新药 dose=3
-+	38	33	23
文庄	35	32	26
sex=2	33	26	21
田州	33	34	34
为性	31	36	31
5CA-1	28	34	32

表 4-19 3 组患者接受治疗后的抑郁水平

1) 适用条件判断

"性别"和"剂量"为两个分类因子,因变量"抑郁水平"为定量数据,数据类型符合方差分 析的要求。本例结合试验设计、样本量,以及方差分析自身的稳健性,认为数据基本满足 正态分布和方差齐性的要求。

2) 是否考察交互作用

本例从既往研究成果和专业经验的角度,认为应当考察性别与剂量水平及二者的交互 作用对抑郁水平的影响,所以接下来的方差分析中包括交互作用的分析。

3) 效应显著性检验

∅ 进阶	方法 🗸	_	开始分析	Ō	二阶效应 ☑ ② 简单效应 ☑ ③ Bonferonni校 ~ 效应量 □
聚类	因子		抑郁水平		¥(定留)
主成分	分层回归				- Concerned
逐步回归	二元Logit		剂量		
多分类Logit	有序Logit		性别		
事后多重比较	偏相关				マクン 米 1745つろい
典型相关	双因素方差				AUE关,1X月277
三因素方差	多因素方差	-			Marken atoms arrests a
协方差	判别分析				忉受量(定量)【可选】

图 4-12 双因素方差分析操作界面

在仪表盘中依次单击【进阶方法】→【双因素方差】模块,将【剂量】与【性别】依 58/187 次拖曳至右侧【X(定类, 仅为 2 个)】分析框中, 将因变量【抑郁水平】拖曳至【Y(定量)】 分析框中。

勾选【二阶效应】和【简单效应】复选框,二阶效应即交互作用,在数据和分析目的 允许的情况下,双因素方差分析一般会默认考察两个因素的交互作用。简单效应即考察交 互作用的多重比较,可通俗理解为控制 A 因素的水平,来研究 B 因素各水平对 Y 的不同影 响。

选择多重比较方法下拉列表中的【Bonferroni 校正】,本例无协变量,因此忽略协变量 设定。双因素方差分析操作界面如图 4-12 所示,最后单击【开始分析】按钮。

双因素方差分析结果如表 4-20 所示, 平方和、自由度 df、均方属于中间计算值。通 常来说,我们可直接读取最后两列的 F 统计量及显著性检验 p 值。

差异源	平方和	自由度df	均方	F	p
Intercept	17422.222	1	17422.222	2825.225	0.000**
剂量	97.444	2	48.722	7.901	0.006**
性别	37.556	1	37.556	6.090	0.030*
剂量×性别	144.778	2	72.389	11.739	0.001**
Residual	74.000	12	6.167		

表 4-20 双因素方差分析结果

注: 1.R²: 0.791。

2 . * p<0.05, ** p<0.01.

方差分析中有交互作用时,要先看交互作用有无统计学意义。本例结果显示, F=11.739,p<0.01,剂量×性别对抑郁水平的影响具有统计学意义,即交互作用显著。性 别和剂量对抑郁水平的影响也是显著的 (p均小于 0.05)。

剂量×性别对抑郁水平的交互作用图如图 4-13 所示。

两条折线(线段)不是平行关系,有交叉现象,认为存在交互作用,结论和方差分析 表交互作用显著性检验一致。性别与剂量的交互作用显著,接下来应该进一步做简单效应 分析。



图 4-13 剂量×性别对抑郁水平的交互作用图

4.2.5 简单效应分析

按前面所述的分析思路,如果方差分析交互作用没有显著性(p>0.05),则可以继续针 对主效应因素的组间差异进行多重比较,操作和解释与单因素方差分析多重比较一致。

如果一个因素的效应大小在另一个因素不同水平下明显不同,则称两因素间存在交互 作用。当存在交互作用(p<0.05)时,单纯研究某个因素的作用没有意义,必须分别按另 一个因素的不同水平研究该因素的作用大小。为了和因素主效应的多重比较进行区分,本 节把针对交互作用的多重比较称为简单效应分析。

鉴于简单效应分析在实际科研数据分析中的重要性,本节单独对"例 4-8"双因素方差分 析的简单效应分析结果进行解读并介绍。

操作时, 在仪表盘中依次单击【进阶方法】→【双因素方差】模块, 重点是同时勾选 【简单效应】【二阶效应】复选框。

简单效应分析(剂量为条件)如表 4-21 所示,此表为在控制剂量各水平条件不变的 情况下考察男性与女性在抑郁水平上的差异。结果表明,接受安慰剂的第 1 组患者,男性 抑郁水平低于女性(均值差值为-4.67),t=-2.30,p=0.04<0.05,差异有统计学意义;接 受中等剂量新药的第 2 组患者,女性抑郁水平低于男性(均值差值为 4.33),t=2.14, p=0.05,差异无统计学意义;接受大剂量新药的第 3 组患者,女性抑郁水平明显低于男性 (均值差值为 9.00),t=4.44,p<0.01,差异有统计学意义。

剂量	性别	均值差值	标准误 SE	t	р
安慰剂	男性−女性	-4.67	2.03	-2.30	0.04
中等剂量	男性−女性	4.33	2.03	2.14	0.05
大剂量	男性−女性	9.00	2.03	4.44	0.00

表 4-21 简单效应分析 (剂量为条件)

简单效应分析(性别为条件)如表 4-22 所示,此表为在性别保持不变的情况下考察 不同剂量对抑郁缓解作用的差异比较。

性别	剂量	均值差值	标准误 SE	t	р
男性	安慰剂-中等剂量	-4.00	2.03	-1.97	0.22
男性	安慰剂-大剂量	-1.67	2.03	-0.82	1.00
男性	中等剂量-大剂量	2.33	2.03	1.15	0.82
女性	安慰剂-中等剂量	5.00	2.03	2.47	0.09
女性	安慰剂-大剂量	12.00	2.03	5.92	0.00
女性	中等剂量-大剂量	7.00	2.03	3.45	0.01

表 4-22 简单效应分析(性别为条件)

结果表明,对于男性患者来说,3 个剂量水平两两之间的抑郁差异比较,p 值依次为 0.22、1.00、0.82,均大于 0.05,无显著性,即对于男性患者来说,给予安慰剂、中等剂量 新药、大剂量新药其抑郁水平差异没有变化,差异无统计学意义,说明对于男性给药的不 同剂量疗效无差别。

对于女性患者,大剂量新药与安慰剂相比,t=5.92,p<0.05;大剂量新药与中等剂量 新药相比,t=3.45,p=0.01<0.05,表明给予大剂量新药的疗效要显著高于安慰剂和中等剂 量新药。中等剂量新药与安慰剂相比,t=2.47,p=0.09>0.05,表明中等剂量新药和安慰剂 的疗效没有差别。

4.3 卡方检验

卡方检验是以卡方分布为基础,针对定类数据资料的常用假设检验方法。其理论思想 是判断实际观测到的频数与有关总体的理论频数是否一致。

卡方统计量是实际频数与理论频数吻合程度的指标。卡方值越小,表明实际频数与理 论频数越接近,卡方值越大表明两者相差越大。卡方检验原假设实际频数与理论频数相同, 依据卡方分布计算显著性 p 值进行统计结论的推断。取显著性水平 α=0.05, 当 p 值小于 0.05 时说明实际频数与理论频数的差异有统计学意义;反之,若 p 值大于 0.05 则说明实际 频数与理论频数无差异。

4.3.1 方法概述

1.应用方向

在实际研究分析中,卡方检验分为独立性/差异性检验和拟合优度检验两个重要的应用 方向。

1) 独立性/差异性检验

独立性检验可理解为判断两组或多组计数资料是相互关联还是彼此独立,如是否患病 与是否吸烟的关联关系。根据研究侧重点的不同,也可用于研究两组或多组样本的总体率 (或构成比)之间的差别是否具有统计学意义,如某疾病联合化疗与单纯化疗两种治疗方 式的存活率有无差异,吸烟与不吸烟两组人群的患病率有无差异。

2) 拟合优度检验

拟合优度检验是检验实际观察的类别频数分布比例与已知类别频数分布比例是否符合的分析方法,通俗讲即检验总体是否服从某个指定分布。例如,在研究牛的相对性状分离现象时,用黑色无角牛和红色有角牛杂交,子二代出现黑色无角牛 192 头,黑色有角牛 78 头,红色无角牛 72 头,红色有角牛 18 头,观察这两对性状是否符合孟德尔遗传定律中 9:3:3:1 的遗传比例。

2.数据要求

用于卡方检验的数据录入格式有两种。第一种是列联表频数资料,即加权数据格式, 在分析时需要提前用频数进行加权。第二种是原始数据记录,即普通数据格式,在分析时 软件工具会自动汇总频数结果。

在实践中,列联表频数数据较常见。列联表也叫作交叉表,是两个分类变量各水平两 两交叉组合后的频数汇总表。

如果两个分类变量均为二水平,那么两者构成 2×2 交叉表,交叉组合频数所在的位置 也叫作单元格,因此 2×2 交叉表共有 4 个单元格,也就是常说的四格表。除四格表外 的列联表,我们可以统称为 R×C 列联表,其中 R 代表行个数,C 代表列个数。

3.卡方检验的类型

按列联表的形式,卡方检验可以简单划分为 2×2 四格表卡方检验、R×C 列联表卡方检 验;按研究设计不同,卡方检验可分为成组设计的独立性卡方检验和配对卡方检验;按研 究目的不同,卡方检验可分为卡方拟合优度检验和独立性卡方检验。此外根据是否有分层 变量,卡方检验还可分为分层卡方检验。常用卡方检验的研究目的和方法选择如表 4-23 所示。

一般进行卡方检验时,默认计算的是 Pearson 卡方统计量,而实际分析时,在 2×2 四 格表卡方检验、R×C 列联表卡方检验中通常会计算多个卡方统计量,包括 Pearson 卡方统 61/187 计量、Yates 校正卡方(连续校正卡方)统计量、趋势卡方统计量,以及 fisher 卡方 (fisher 确切概率)统计量等,应注意区分不同卡方统计量及其检验的适用条件,根据适 用条件选择恰当的结果进行解释和分析。

使用 SPSSAU 平台进行卡方检验时,应注意【通用方法】下的【交叉(卡方)】模块只能 提供较大样本时常用的 Pearson 卡方检验,而【实验/医学研究】下的【卡方检验】模块可 提供 Pearson 卡方检验、Yates 校正卡方检验,以及 fisher 卡方检验,适合小样本数据。

研究目的	卡方检验方法选择	SPSSAU 平台分析路径	
研究类别定类数据的实际 比例与预期比例是否吻合	拟合优度检验	【实验/医学研究】→【拟合优度检验】	
成组设计两组或多组率的 差异比较	 2×2 四格表卡方检验 2×2 四格表fisher 确切检验 <i>R</i>×2 列联表卡方检验 		
成组设计两组或多组的构 成比差异比较	2×C列联表卡方检验 R×C列联表卡方检验	【週用万法】→【父又下万】 【实验/医学研究】→【卡方检验】	
多组率的多重比较	卡方检验多重比较	【头验/医学研九】→【fisher下方】	
单向或双向有序资料的趋	Cookran Armitaga 趋热卡方检验		
势线检验	Coeman-Annitage 起另下力並並		
配对设计定类数据间的差 异关系	配对卡方检验	【实验/医学研究】→【配对卡方】	
控制混杂因素的卡方检验	分层卡方检验	【实验/医学研究】→【分层卡方】	

表 4-23 常用卡方检验的研究目的和方法选择

4.3.2 2×2 四格表卡方检验

四格表是最基础的列联表之一,四格表卡方检验也是较常见的一种检验方法。本节介绍的是成组设计的四格表卡方检验,具体分析时需区分 Pearson 卡方检验、Yates 校正卡方检验,以及 fisher 卡方检验。四格表卡方检验的适用条件如表 4-24 所示。

样本量与期望频数条件	方法选择
<i>n</i> ≥40 <u>∃</u> . <i>E</i> ≥5	Pearson 卡方检验
n≥40 但任意单元格出现 1≤E<5	Yates 校正卡方检验
任意单元格出现 E<1 或 n<40	fisher 卡方检验

表 4-24 四格表卡方检验的适用条件

表中 n 为样本量, E 为理论频数。在四格表资料中, Pearson 卡方检验适用于 n≥40 且全部单元格 E≥5 的情况。如果四格表 n≥40 但任意单元格出现 1≤E<5, 则一般考虑采 用 Yates 校正卡方检验, 它是对 Pearson 卡方检验的校正结果; 如果四格表中任意单元格 出现 E<1 或 n<40, 则应当采用 fisher 卡方检验。

一般通过在仪表盘中依次单击【实验/医学研究】→【卡方检验】模块来实现成组设计 的四格表卡方检验。

【例 4-9】某医院欲比较异梨醇口服液(试验组)和氢氯噻嗪+地塞米松(对照组)降 低颅内压的疗效。将 200 例颅内压增高症患者随机分为两组,两组降低颅内压有效率的比 较如表 4-25 所示。问两组降低颅内压的有效率有无差别?数据来源于孙振球和徐勇勇 (2014),数据文档见"例 4-9.xls"。

表 4-25 两组降低颅内压有效率的比较

组别	有效	无效	合计	有效率

试验组	99(95.19%)	5(4.81%)	104	95.19%
对照组	75(78.13%)	21(21.87%)	96	78.13%
合计	174	26	200	87.00%

1) 数据与案例分析

本例为频数资料,使用 SPSSAU 平台进行分析时,应按照加权数据格式录入数据。"组别""疗效"均有两个水平的分类变量,"频数"为定量数据。在具体分析时需要将"频数"作为权 重进行加权。本例加权数据格式如表 4-26 所示。

组别	疗效	频数
试验组	有效	99
试验组	无效	5
对照组	有效	75
对昭组	无效	21

表 4-26 例 4-9 加权数据格式

本例目的在于比较试验组与对照组的治疗有效率,是典型的两组率的差异检验。数据 资料为四格表频数资料,采用四格表卡方检验进行分析。

2) 卡方检验

读入数据后,在仪表盘中依次单击【实验/医学研究】→【卡方检验】模块,将【组别】 拖曳至【X(定类)】分析框中,将【疗效】拖曳至【Y(定类)】分析框中。本例需加权处理, 故要把【频数】拖曳至【加权项(可选)】分析框中。选择下拉列表中的【百分数(按列)】, 卡方检验操作界面如图 4-14 所示,最后单击【开始分析】按钮。

♀ 实验/医	♀ 实验/医学研究 ∨		开始分析	面 百分数(按列) ~ Q
卡方检验	Карра		组别	X(定类)
配对卡方	二元Probit		疗效	V(定类)
Poisson回归	Cox回归			
ICC组内相关 系数	单样本 Wilcoxon			
配对样本 Wilcoxon	多样本 Friedman	Ð		
游程检验	Kendall协调 系数		频数	加权项(可选)

图 4-14 卡方检验操作界面

3) 结果解读

为方便理解不同卡方统计量及其检验的适用条件,可先对卡方统计量、期望频数 E、 样本容量 n 进行统计汇总,如表 4-27 所示,最后按表 4-28 所示的卡方检验分析结果进行 报告。

表 4-27 卡方统计量、期望频数 E 及样本容量 n 统计汇总

项	名称	值
	Pearson 卡方	12.857(p=0.000**)
	Yates 校正卡方	11.392(<i>p</i> =0.001**)
	fisher 卡方	12.793(<i>p</i> =0.000**)
	<i>E</i> ≥5	4(100.00%)
组别×疗效(2×2)	1≤ <i>E</i> <5	0(0.00%)
	<i>E</i> <1	0(0.00%)
	Cnt	4
	n	200

自由度df值	1

注: ** p<0.01。

表 4-27 中包括了 Pearson 卡方、Yates 校正卡方和 fisher 卡方的统计量,显著性 p 值, 以及期望频数等基本信息的汇总结果。E 为列联表单元格的期望频数, n 为样本量。本例 中,样本量为 200 例,所有单元格的期望频数均大于 5。根据卡方检验适用条件,应选择 读取 Pearson 卡方的检验结果,即卡方值为 12.857, p<0.01,认为试验组与对照组两组人 群的治疗有效率差异且有统计学意义。

55日 夕む		组别		百计	.2	
赵曰	白小	试验组	对照组	心口	χ-	ρ
它为	有效	99(95.19%)	75(78.13%)	174(87.00%)		
91 XX	无效	5(4.81%)	21(21.88%)	26(13.00%)	12.857	0.000**
总	.H	104	96	200		

表 4-28 卡方检验分析结果

注: ** p<0.01。

表 4-28 所示最后两列给出了卡方统计量和 p 值, χ2=12.857, p<0.01, 其结果为平台 自动匹配选择输出的 Pearson 卡方检验结果。

结合以上分析,可知试验组和对照组在降低颅内压的疗效上并不相同。试验组的有效 率为 95.19%,对照组的有效率为 78.13%。经 Pearson 卡方检验,按 α=0.05 水平,认为两 组降低颅内压的总体有效率不相等,具体表现为异梨醇口服液的有效率高于和氢氯噻嗪+ 地塞米松的有效率。

本例样本量和期望频数的条件满足 Pearson 卡方检验的适用性,如果在四格表中有单元格出现期望频数在 1~5 的情况,则表 4-28 会自动选择输出 Yates 校正卡方的结果。一般建议读者直接读取和使用表 4-28 的分析结果进行报告。

4.3.3 R×C 列联表卡方检验与多重比较

R×C 列联表卡方检验包括常见的 2×C 两组构成比的差异比较, R×2 两组或多组率的差 异比较, 还有多行多列即 R×C 列联表卡方检验。

本节内容介绍列联表中行与列两个变量均为无序多分类变量的情况,常见 R×C 列联表 说明如表 4-29 所示。

常见 RxC 列联表卡方检验	方法的应用场景
2×C: 两组构成比的差异比较	脑梗死组与对照组在血型分布(A/B/AB/O型)上有无差异
R×2: 两组或多组率的差异比较	超短波、温热磁和蜡疗3种疗法的治愈率(无效、有效)有无差别
R×C: 两个分类变量的独立或关联关系	A、B、O 血型和 MN 血型之间是否有关联

表 4-29 常见 RxC 列联表说明

R×C 列联表卡方检验采用的仍然是 Pearson 卡方检验,其适用条件要求列联表中所有 的单元格理论频数不小于 1,并且 1≤E<5 的单元格数量不宜超过总数的 1/5。如果出现不 符合该条件的情况,研究者要考虑采用 fisher 卡方检验。

R×C 列联表卡方检验的结论是,总体上差异是否显著,如果还想进一步判断两两总体率的差异,应继续做卡方结果的多重比较。在进行卡方检验时,【实验/医学研究】下的 【卡方检验】模块会根据分组情况自动实现卡方分割法的多重比较。 以 R×2 列联表为例, 卡方分割的基本思想是将 R×2 拆分为多个 2×2 四格表资料从而 实现两两比较。例如, 3×2 列联表能检验 3 组总体率的差异, 假设 3 组名称分别为 A、B、 C, 可拆分出 AB (A、B 两组率的比较)、AC、BC 共 3 个 2×2 四格表卡方。

多重比较时,检验次数增多会增加 | 类错误的概率,此过程应当对卡方检验 p 值进行校正,建议使用校正显著性水平(Bonferroni 校正)。假设显著性水平为 0.05,当两两比较次数为 3 次时,Bonferroni 校正显著性水平为 0.05/3=0.0167,即 p 值需要与 0.0167 进行对比,从而做出统计推断,而不是与 0.05 比较。

【例 4-10】某医师研究物理疗法、药物疗法和外用膏药 3 种疗法治疗周围性面神经麻 痹的疗效, 3 种疗法对周围性面神经麻痹有效率的比较如表 4-30 所示。问 3 种疗法的有效 率有无差别?数据来源于孙振球和徐勇勇,本例数据文档见"例 4-10.xls"。

疗法	有效	无效	合计	有效率
物理疗法组	199	7	206	96.60%
药物疗法组	164	18	182	90.11%
外用膏药组	118	26	144	81.94%
合计	481	51	532	90.41%

表 4-30 3 种疗法对周围性面神经麻痹有效率的比较

1) 数据与案例分析

数据包括"组别""疗效""频数"3 个变量,为加权数据格式。在表 4-30 最后一列计算了 各组疗法的有效率,依次为 96.6%、90.11%、81.94%,本例目的是比较 3 种不同疗法的疗效 差异,为多组率的差异比较,接下来进行 R×2 卡方检验。

2) 卡方检验

读入数据后, 在仪表盘中依次单击【实验/医学研究】→【卡方检验】模块, 将【组别】 拖曳至【X(定类)】分析框中, 将【疗效】拖曳至【Y(定类)】分析框中。本例需要加权处理, 故要把【频数】拖曳至【加权项(可选)】分析框中。选择下拉列表中的【百分数(按列)】, 最后单击【开始分析】按钮。

和【例 4-09】一样,表 4-31 所示的卡方检验与期望频数统计汇总用于根据期望频数、 样本量等适用条件选择正确的统计量和检验结果。

项	名称	值
	Pearson 卡方	21.04(<i>p</i> =0.00**)
	Yates 校正卡方	21.04(<i>p</i> =0.00**)
	fisher 卡方	—
组别×疗效(3×2)	<i>E</i> ≥5	6(100.00%)
	1≤ <i>E</i> <5	0(0.00%)
	<i>E</i> <1	0(0.00%)
	Cnt	6
	n	532
EL/11/11 XX (3×2)	自由度df 值	2

表 4-31 卡方检验与期望频数统计汇总

注: ** p<0.01。

本例所有单元格 E≥5, 样本量为 532, 依据卡方检验的适用条件, 选择 Pearson 卡方 检验, 即卡方等于 21.04, p<0.01。

表 4-32 3 种疗法周围性面神经麻痹治愈率卡方检验

馬口	夕む		组别	百分	.2	2	
赵曰	白你	外用膏药组	物理疗法组	药物疗法组	下す	χ-	ρ

疗动	无效	26(18.06%)	7(3.40%)	18(9.89%)	51(9.59%)		
71 XX	有效	118(81.94%)	199(96.60%)	164(90.11%)	481(90.41%)	21.04	0.00**
总	计	144	206	182	532		

注: ** p<0.01。

表 4-32 所示的 3 种疗法周围性面神经麻痹治愈率卡方检验用于科研写作的结果报告, x2=21.04, p<0.01, 说明 3 种疗法治疗周围性面神经麻痹的有效率有差别。

卡方检验时常使用相关性指标作为效应量的估计,如果研究时需要报告组别的效应量,则可结合数据类型及交叉表格类型综合选择。卡方检验效应量如表 4-33 所示。本例选择 Cramer v 效应量,此处 Cramer v 为 0.20,说明疗法与治疗效果之间存在弱相关性。

表 4-33 卡方检验效应量

•	分析项	Phi	列联系数	校正列联系数	Cramer v	Lambda
-	疗效	0.20	0.20	0.28	0.20	0.00

综合以上信息,比较 3 种疗法的疗效。结果显示, χ²=21.04, p<0.01, 提示不同疗法 其疗效不同, 3 种疗法的有效率差异具有统计学意义。Cramer v=0.20, 提示疗法与治疗效 果之间存在弱相关性。

3) 卡方分割多重比较

当卡方检验整体有显著性时, 在仪表盘中依次单击【实验/医学研究】→【卡方检验】 模块, 平台会根据分组情况自动实现卡方分割法多重比较, 组别×疗效的卡方分割法多重 比较如表 4-34 所示。

比较	名称	外用膏药组	物理疗法组	药物疗法组	χ^2	р
第1次 无效 有效	无效	26	7	-	21.22	0.00**
	有效	118	199	-	21.32	0.00**
第2次	无效	26	-	18	4.50	0.02*
	有效	118	-	164	4.39	0.05*
笛 2 次	无效	-	7	18	676	0.01**
弗 5 八	有效	-	199	164	0.70	0.01**

表 4-34 组别×疗效的卡方分割法多重比较

注: 1.**p*<0.05 ***p*<0.01。

2.比较次数:3。

由表 4-34 可知,共进行了 3 次拆分。第 1 次是外用膏药组与物理疗法组之间的治疗 有效率差异比较;第 2 次是外用膏药组与药物疗法组之间的治疗有效率差异比较;第 3 次是物理疗法组与药物疗法组之间的治疗有效率差异比较。每一次拆分,都是一个独立 的四格表卡方检验,能输出对应的卡方统计量与显著性 p 值。

此处要注意, 表中最后一列的 p 值需要与校正后的显著性水平进行比较。本例进行了 3 次两两比较, 因此校正的显著性水平 α=0.05/3=0.0167, p 值要与 0.0167 进行比较, 从 而推断检验结果。

物理疗法组与外用膏药组的 p<0.0167,物理疗法组与药物疗法组的 p=0.01<0.0167,可以认为物理疗法与药物疗法组、外用膏药组的治疗有效率均有统计学差异,具体来说是物理疗法组的治疗有效率高于其他两种疗法,但我们尚不能认为药物治疗与外用膏药的治疗有效率有差异 (p=0.03>0.0167)。

4.3.4 fisher 卡方检验

fisher 卡方检验是由 R.A.Fisher (1934)提出的,其理论依据是超几何分布,是一种直接计算概率的假设检验方法。由于其在四格表中应用较广泛,所以也被称为四格表确切概率,此外它也可在 R×C 列联表卡方检验条件不满足时使用。在 SPSSAU 平台中,可通过 【实验/医学研究】下的【卡方检验】(fisher 卡方】两个模块实现 fisher 卡方检验。

1.2×2 四格表资料

当 2×2 四格表资料出现 E≤1 或 n<40 情况时,可使用 fisher 卡方检验进行假设检验。 此外,当四格表卡方检验后所得概率 p 接近检验水平 α 时,亦可使用 fisher 卡方检验。

【例 4-11】四格表 fisher 卡方检验。将 23 名抑郁症患者随机分为两组,分别用两种 药物治疗,效果如表 4-35 所示,问两种药物的治疗效果是否不同?数据来源于方积乾 (2012),数据文档见"例 4-11.xls"。

公组	治疗效果		△辻	方动变
刀组	有效	无效		有双平
甲药	7	5	12	58.3%
乙药	3	8	11	27.3%
合计	10	13	23	43.5%

表 4-35 两种药物治疗抑郁症的效果

1) 数据与案例分析

数据文档为加权数据格式,根据适用条件,本例四格表资料的样本量 n<40,Pearson 卡方检验、Yates 校正卡方检验不再适用, 宜采用 fisher 卡方检验进行分析。

2) 卡方检验

四格表资料的 fisher 卡方检验可直接通过【卡方检验】模块来实现,将【组别】拖曳 至【X(定类)】分析框中,将【疗效】拖曳至【Y(定类)】分析框中。本例要先进行加权处理, 再把【频数】拖曳至【加权项(可选)】分析框中。选择下拉列表中的【百分数(按列)】,最 后单击【开始分析】按钮。

本例 n=23<40, 且有单元格 1≤E<5。因此 fisher 卡方检验结果更可靠。卡方检验分 析结果如表 4-36 所示, 确切概率 p=0.214, 在 α=0.05 检验水平上尚不能认为两种药治疗 抑郁症的效果不同。

斯日	夕む	组别 n		台升	~2	n
	口仰	乙药	甲药	100 61	χ	ρ
应动	无效	8(72.73%)	5(41.67%)	13(56.52%)		
91 XX	有效	3(27.27%)	7(58.33%)	10(43.48%)	3.733	0.214
ė	针	11	12	23		

表 4-36 卡方检验分析结果

2.R×C列联表资料

R×C 列联表 Pearson 卡方检验要求列联表中所有单元格的理论频数不小于 1, 并且 1≤ E<5 的单元格数量不宜超过总数的 1/5。如果出现不符合该条件的情况,则研究者要采用 fisher 卡方检验进行差异检验。

【例 4-12】R×C 列联表 fisher 卡方检验。某研究回顾性分析 62 例 NME 乳腺癌患者 的 MRI 影像及临床病理资料,其中不同分子亚型在有无瘤周水肿的分布数据如表 4-37 所 示。试分析 4 种亚型在瘤周水肿的分布有无差别。数据来源于晋瑞等人(2019),数据文

档见"例 4-12.xls"。

瘤周水肿	Luminal A 型	Luminal B 型	HER-2 过表达型	三阴性型
有	3	23	12	7
无	7	10	0	0

表 4-37 不同分子亚型在有无瘤周水肿的分布数据

1) 数据与案例分析

本例数据是 4×2 列联表数据,有 8 个单元格频数,总样本量 n=62>40。通过【卡方 检验】模块进行分析,我们发现 1≤E<5 的单元格数比例为 37.5%,超过了 1/5 的单元格数, Pearson 卡方检验不再适用,此时宜采用 fisher 卡方检验进行检验。

2) fisher 卡方检验进行检验

【卡方检验】模块只提供 2×2 四格表资料的 fisher 卡方检验,针对 R×C 列联表我们使用【fisher 卡方】模块。该模块无须录入数据,而采取直接输入频数数字的方式进行假设检验。

在仪表盘中依次单击【实验/医学研究】→【fisher 卡方】模块,在弹出的电子表格区 域直接手动输入案例数据, fisher 卡方检验操作界面如图 4-15 所示,最后单击【开始分析】 按钮。

	А	B	С	D	F
1		- 有水肿	- 无水肿		-
2	亚型1	3	7		
3	亚型2	23	10		
4	亚型3	12	0		
5	亚型4	7	0		
开始分析 🖸					

图 4-15 fisher 卡方检验操作界面

3) 结果分析

特别强调, fisher 卡方检验可直接计算假设检验的 p 值, 因此中间并无统计量。分子 亚型×有无瘤周水肿 fisher 卡方检验如表 4-38 所示。

表 4-38 分子亚型×有无瘤周水肿 fisher 卡方检验

检验	卡方值	p
fisher 卡方	—	0.001
Pearson卡方	16.426	0.001
连续校正卡方	16.426	0.001

经计算,本例 fisher 卡方检验的 p=0.001<0.05,认为 4 种不同分子亚型在瘤周水肿上的分布差异有统计学意义。

4.3.5 配对卡方检验

前面介绍的是成组设计的卡方检验,如两组或多组率的比较,此处的组就是各自独立

的对象。与成组设计对应的是配对(或配伍)设计。

配对设计的特点是对同一样本分别用 A 和 B 两种方法处理,或对样本前后进行测量, 观察其水平分布的差异或一致性。若采用配对设计,当结局指标为计数资料时,则需要采 用配对设计的卡方检验方法。

如果是四格表的配对资料,则使用 McNemar 检验;如果是非四格表的配对资料,则 使用 McNemar-Bowker 检验,二者的区别如表 4-39 所示。

表 4-39 McNemar 检验与 McNemar-Bowker 检验的区别

配对数据类型	检验
2×2 四格表	McNemar 检验
R×C列联表(多行多列)	McNemar-Bowker 检验

【例 4-13】124 名学生参加 1000 米长跑,训练一个月前后两次测验的达标情况如表 4-40 所示,问一个月的训练是否有显著效果?数据来源于王孝玲(2007),数据文档见"例 4-13.xls"。

第一次测验	第二次测验			
	达标	未达标		
达标	61	19		
未达标	33	11		

表 4-40 学生训练一个月前后两次测验的达标情况

1) 数据与案例分析

数据文档为加权数据格式,包含3个变量。标题分别为"第一次测验"、"第二次测验"和 "频数",两次测验的结果均为"达标"或"未达标"。每个学生均有前后两次测验,并且测验的 结果均为"达标"或"未达标",这是配对设计的定类数据资料。综合这两点,本例不适用前面 学习的普通成组设计的卡方检验,而应采用专门用于配对设计的卡方检验方法。

2) 配对卡方检验

数据读入平台后, 在仪表盘中依次单击【实验/医学研究】→【配对卡方】模块, 首先 将【第一次测验】拖曳至【配对 1(定类)仅 1 项】分析框中, 将【第二次测验】拖曳至【配 对 2(定类)仅 1 项】分析框中, 然后将【频数】拖曳至【加权项(可选)】分析框中, 操作界 面如图 4-16 所示, 最后单击【开始分析】按钮。

☞ 实验/医	学研究 🗸		开始分析	0 m
卡方检验	Kappa		第一次测验	配对1 (定类)【仅1项】
配对卡方	二元Probit			
Poisson回归	Cox回归			
ICC组内相关 系数	单样本 Wilcoxon		第二次测验	配对2 (定类)【仅1项】
配对样本 Wilcoxon	多样本 Friedman	Ð		
游程检验	Kendall协调 系数			
Cochran's Q 检验	概要 t 检验		频数	加权项(可选)

图 4-16 配对卡方检验操作界面

3) 结果解读

四表格资料的配对设计使用 McNemar 检验;如果配对数据的组别大于 2,即配对多

分类时,则使用 McNemar-Bowker 检验。

本例使用的是 McNemar 检验, 配对卡方检验分析结果如表 4-41 所示。x2=19.000, p=0.070>0.05, 表明训练前第一次测验的结果与训练后第二次测验的结果无差异, 或差异 无统计学意义。

빠기 카카	而对 夕む		第二次测验		.2	2
自し入り	有你	未达标	达标		χ-	ρ
	未达标	11	33	44		
另 (八侧型	达标	19	61	80	19.000	0.070
总	.it	30	94	124		

表 4-41 配对卡方检验分析结果

4.3.6 分层卡方检验

卡方检验研究的是分类变量 X 对分类变量 Y 的关系,这样的关系可能会受第三个混杂 因素的影响。例如,研究是否吸烟(X)与是否患某病(Y)的关系,专业上认为年龄起到了干扰 作用,将其纳入分析范畴作为分层项,这种分析称为分层卡方检验。

1.基本概念

当给交叉表卡方检验加入分层变量后,卡方检验被拆分为不同的层次水平,每层均可 单独完成交叉表卡方检验进行分析,但是研究者关注的是扣除分层因素的干扰影响后,列 联表中行变量对列变量的影响是否显著。分层卡方可以用 Cochran-Mantel-Haenszel 检验 (简称 CMH 检验)进行分析,能很好地解决"辛普森悖论"问题,常用于病例对照试验及研 究。

OR 值是分层卡方中的一个重要概念。OR (Odds Ratio) 值又称比值比、优势比, 等于病例组中暴露人数与非暴露人数的比值除以对照组中暴露人数与非暴露人数的比值, 是流行病学研究中的一个常用指标, 反映的是疾病和暴露的关联强度。例如, 在研究高血压对心肌梗死的影响时, OR 值为 3.5, 通俗理解即患高血压的人发生心肌梗死的风险是未患高血压的人发生心肌梗死风险的 3.5 倍。在本书后续的 Logistic 回归、医学研究内容中还会详细介绍其应用。

目前应用较多的是 2×2×K 结构数据资料的分层卡方检验(X 和 Y 均为两个分类,分层 项为 K 层)。该方法一般只能考察或控制一个混杂因素,当混杂因素超过 1 个时,可考虑 Logistic 回归分析。

2.CMH 检验的分析思路

CMH 检验的基本思想是对各层的 OR 值进行合并,并进行合并后的独立性卡方检验,分析思路如图 4-17 所示。



70 / 187

图 4-17 CMH 检验的分析思路

1) 各层 2×2 卡方检验 p 值、OR 值

CMH 检验分别就各层频数数据进行普通的 2×2 四格表卡方检验, 解释和分析各层中 行变量与列变量间的关系, 计算并报告 OR 值。

2) OR 值齐性检验

对各层 OR 值进行齐性检验,如果各层 OR 值同质或一致,则说明当前纳入的分层因 素没有混杂干扰作用或干扰很弱,各层的 OR 值可进行合并,相当于消除分层因素影响后 用一个统一的 OR 值评价行与列变量的影响关系;如果各层的 OR 值不同质或不一致,则 说明分层因素存在混杂干扰作用,此时应分别报告各层的 OR 值。

一般用 Breslow-Day 检验和 Tarone's 检验完成 OR 值的齐性检验。当 p<0.05 时,说明各层的 OR 值是不一致的;当 p>0.05 时,说明各层的 OR 值是一致的。

3) 合并并报告调整后的 OR 值

当上一步检验出 OR 值具有一致性时,该结果提示可对各层的 OR 值进行合并计算, 合并后会得到一个调整后的 OR 值,它的大小反映的是排除分层因素影响后,行变量对列 变量的影响程度。在病例对照研究中,即排除分层因素影响后,暴露因素对结局的影响。

4) 条件独立性检验

CMH 检验在排除分层因素干扰后,可继续进行行变量与列变量的独立性或差异性关系 检验,最终得到分层卡方检验的结果。

3.实例分析

本节结合具体案例进一步介绍分层卡方检验在科研中的应用。

【例 4-14】某研究调查口服避孕药(OC)与心肌梗死的关系,考虑年龄是一个可能的混杂因素,将其纳入调查,案例数据如表 4-42 所示,试分析在年龄影响下心肌梗死与服用避孕药有无关系。数据来源于张文彤(2002),数据文档见"例 4-14.xls"。

避孕药	心肌梗死	年龄	频数
2	2	1	21
1	2	1	26
2	1	1	17
1	1	1	59
2	2	2	18
1	2	2	88
2	1	2	7
1	1	2	95

表 4-42 例 4-14 案例数据

1) 数据与案例分析

定类数据编码情况,分组变量"避孕药":数字1代表不服用避孕药,数字2代表服

用避孕药;结局变量"心肌梗死":数字1代表对照组,数字2代表病例组;分层项变量"年龄":数字1代表小于40岁,数字2代表大于或等于40岁。

考察两个分类变量间的关系,显然卡方检验是可选方法之一。现在还需要考察第三个 分类变量的干扰,可将其作为分层项,执行分层卡方检验。

2) 分层卡方检验

读入数据后, 在仪表盘中依次单击【实验/医学研究】→【分层卡方】模块, 该模块支 持普通数据格式与加权数据格式两种形式, 如果有加权项则在【Weight(加权项)】下拉列 表中选择具体的变量标题,如果没有加权项则不用选择。本例为加权数据格式,选择【频 数】变量进行加权。分层卡方操作界面如图 4-18 所示。

Group(X)	避孕药	~
Outcome(Y)	心肌梗死	~
Factor(分层项)	年龄	~
Weight(加权项)	频数	~
开始分析 Q	请选择【可选】 避孕药 心肌梗死 年龄 频数	

图 4-18 分层卡方操作界面

最后单击【开始分析】按钮,参照前面介绍的分析思路,按以下顺序对结果进行解释 和分析。

3) 各层 2×2 卡方检验

本例共有两层,分层卡方检验汇总表格如表 4-43 所示,不论年龄小于 40 岁还是大于 或等于 40 岁,心肌梗死与是否服用避孕药均有关联(卡方检验 p 均小于 0.05)。年龄小于 40 岁、大于或等于 40 岁两层的 OR 值依次为 2.80、2.78,均大于 1,提示不论年龄小于 40 岁还是大于或等于 40 岁,服用避孕药都是发生心肌梗死的危险因素。

	心肌梗死	对题	照	病	例				
年龄	波马达	不服用	服用	不服用	服用	χ^2	р	OR 值	OR 值 95% CI
	处于于到	避孕药	避孕药	避孕药	避孕药				
<	<40 岁	59	17	26	21	6.77	0.02	2.80	1.27~6.17
2	≥40 岁	95	7	88	18	5.03	0.03	2.78	1.11~6.97
	汇总	154	24	114	39	7.70	0.01	2.20	1.25~3.86

表 4-43 分层卡方检验汇总表格

注: 1.2×2表格时使用 fisher 卡方检验。

2. OR 值计算时使用较小值/较大值。

4) OR 值齐性检验

若要研究,结果包括各层 OR 值还是合并 OR 值,其前提条件是各层 OR 值是否满足 齐性 (同质或一致)。Breslow-Day-Tarone 检验结果如表 4-44 所示。

表 4-44	Breslow-Day-Tarone	检验结果
--------	--------------------	------

χ^2	df	p
0.00	1	0.99

本例 p 为 0.99>0.05, 表明两个年龄层间 OR 值同质或具有一致性。因此接下来我们需 要解读合并 OR 值及条件的独立性检验结果。如果 OR 值异质,则返回上一步骤分别报告 各层的 OR 值。
5) 合并并报告调整后的 OR 值

表 4-45 同时给出了合并计算的 OR 值,以及分层卡方检验结果(条件独立性检验)。

表 4-45 Cochran-Mantel-Haenszel 条件独立性检验

OR 值(Mantel-Haenszel Common)	95% CI	χ^2	df	p
2.79	$1.53 \sim 5.08$	10.73	1	0.00

合并 OR 值 (Mantel-Haenszel Common),即扣除分层变量影响后分类变量 X 对分类 变量 Y 的调整后 OR 值,在本例中即排除年龄混杂影响后,是否服用避孕药对心肌梗死的 优势比。本例调整后 OR=2.79,95% CI 为[1.53,5.08], CI 不包括 1,结论是排除年龄干扰影 响后,服用避孕药患心肌梗死的危险度是未服用避孕药的 2.79 倍。

6) 条件独立性检验

条件独立性检验(分层卡方检验)的卡方值等于 10.73, p<0.05, 表明排除年龄干扰 影响后, 心肌梗死与服用避孕药相关。

本例的最终结论:排除年龄干扰影响后,心肌梗死与服用避孕药相关 (p<0.05),服用 避孕药患心肌梗死的危险度是未服用避孕药的 2.79 倍。

4.3.7 卡方拟合优度检验

拟合优度检验 (Goodness of fit test), 是判断实际观察的类别频数分布比例与已知类 别频数分布比例是否符合的分析方法。

原假设样本实际频数分布与理论频数分布一致或相同,备择假设为不一致或不相同, 构造并计算卡方统计量及概率 p,利用 p 进行统计推断。当 p<0.05 时拒绝原假设,认为 实际频数分布与理论频数分布不相符;反之,当 p>0.05 时,认为实际频数分布与理论频 数分布一致。

常见的有等比例假设与不等比例(指定比例)假设。等比例假设,如某高校招生男研 究生 100 名、女研究生 120 名,分析研究生的性别比例有无差别(默认假设男女比例相 同);不等比例假设,如某高校研究生招生,原计划招生的男女性别比例为 7:3,实际招 男研究生 50 名,女研究生 35 名,分析实际研究生性别比例是否与原计划相符。

【例 4-15】用上述例子, 某高校研究生招生, 原计划招生的男女性别比例为 7:3, 实际招男研究生 50 名, 女研究生 35 名, 分析实际研究生性别比例是否与原计划相符(不 等比例假设的例子)。

1) 数据与案例分析

本例目的在于检验实际招生频数或比例是否与原计划频数或比例一致,首先将实际招 生频数录入 Excel 表格,原计划频数无须录入。实际招生情况 Excel 数据录入格式如图 4-19 所示,数据文档为"例 4-15.xls"。

然后将数据读入 SPSSAU 平台,可以在【数据处理】下的【数据标签】模块中设定本例的数据标签,数字1代表性别女,数字2代表性别男。

	А	В	С
1	性别	人数	
2	男	50	
3	女	35	
4			



73 / 187

图 4-19 实际招生情况 Excel 数据录入格式 图 4-20 【期望值设置】数值框

2) 拟合优度检验

在仪表盘中依次单击【实验/医学研究】→【卡方拟合优度】模块,将左侧的标题框中 的【性别】拖曳至【分析项(定类)】分析框中,将【人数】拖曳至【加权项(可选)】分析框 中。在分析框上方勾选【期望值设置】复选框,会弹出图 4-20 所示的【期望值设置】数 值框。在数字标签 1 数值框输入期望值【3】,在数字标签 2 数值框输入期望值【7】,期望 比

例可以设置成 3 和 7 或 0.3 和 0.7,也可以设置成 30 和 70,能准确表达相对比例即可, 平台会自动进行"归一化"计算。最后单击【开始分析】按钮。

3) 结果分析

本例卡方拟合优度检验结果如表 4-46 所示。

表 4-46 例 4-15 卡方拟合优度检验结果

名称	项	实际频数	期望频数	残差	实际比例	期望比例	χ^2	р
林見山	女	35.000	25.500	9.500	41.18%	30.00%	5.056	0.025*
任别	男	50.000	59.500	-9.500	58.82%	70.00%	5.050	0.023

注: **p<*0.05。

结果显示, 男女的实际比例是 58.82%、41.18%, 期望比例是 70.00%和 30.00%, \chi²=5.056, p=0.025<0.05, 说明男研究生和女研究生的实际招生比例与原计划比例的差异 具有统计学意义, 即实际招生的性别比例和原计划相比发生了变化。

4.4 非参数检验

t 检验与方差分析等参数检验方法,是事先假定总体服从正态分布或近似服从正态分 布,已知总体分布。实践中,我们并不总是已知总体的分布类型。这种在总体参数未知的 情况下,利用样本数据对总体分布形态等进行推断的方法,称为非参数检验。

在分析两组或多组均值的差异性时,当数据严重偏态,经数据转换也不能使其满足参数检验的正态性条件时,可考虑采用相对应的非参数秩和检验。

4.4.1 方法介绍

1.非参数检验

非参数检验与参数检验的区别,如表 4-47 所示。

表 4-47 非参数检验与参数检验的区别

分析方法	参数检验	非参数检验
适用范围	正态分布	分布未知
检验效能	盲	低
描述指标	平均值	中位数
图形展示	折线图	箱线图

非参数检验对数据分布没有要求,适用范围比参数检验广泛;但是由于其检验效能较

低,因此在不是严重偏离正态分布的情况下,应当优先选择参数检验方法。

在数据的描述统计方面,非参数检验因为分布未知,适合采用中位数进行描述,相对 应地,在统计展示时箱线图更合适。

2.秩和检验

本节介绍的秩和检验,是一类常用的非参数检验。秩和检验是先将数据从小到大,或 等级从弱到强转换成秩次,再计算秩和统计量,从而实现假设的检验方法。

秩和检验不受总体分布限制,适用于等级资料,适用面广。与参数检验类方法相比, 其统计检验功效较低。

3.类型与方法选择

当 t 检验、方差分析等参数检验不满足适用条件时,可用秩和检验代替。常见的非参数秩和检验包括 Wilcoxon 检验、Mann-Whitney 检验、Kruskal-Wallis 检验等,这些检验方法与 t 检验、方差分析等参数检验方法的对应关系,如表 4-48 所示。

表 4-48 非参数检验方法与参数检验方法的对应关系

研究目的	参数检验方法	非参数检验方法	分析路径
单一样本均值与已知总体	苗样木,松 砂	的样本Wilcovon 检心	【实验/医学研究】→【单样本
均值的差异比较	平作平 1 包包	平作本Wilcoxoli 应应	Wilcoxon检验】
两个独立样本均值的差异	油 立 样 木 水 於 心	Monn Whitney 松心	
比较	1出工作中 1 1亚亚	Wallin- w lituley 100.300	【通用方法】→【非参数检验】
多组数据均值的差异比较	单因素方差分析	Kruskal-Wallis 检验	
配对样本均值的差导比较	而对样本,检应	而对Wilcovon检心	【实验/医学研究】→【配对
癿朳杆平均值的左开比权	日山小小十一年 1 1 亚马亚	自L >J WIICOXOII 位 动	Wilcoxon检验】
多个相关样本分布的差异	随机区组支差分析	Friedman 检验	【实验/医学研究】→【多样本
比较	随机运知力左刀机	1 Houman 小巫司巫	Friedman]

需要说明的是, SPSSAU 平台【通用方法】下的【非参数检验】模块, 如果 X 的组别为两组, 则提供 Mann-Whitney 检验, 如果组别超过两组, 则提供 Kruskal-Wallis 检验。

此处要注意,在条件允许的情况下应优先使用参数检验方法;当条件不满足时,如非 正态分布或方差不齐时,才考虑使用非参数秩和检验。

4.4.2 单样本 Wilcoxon 检验

Wilcoxon 符号秩检验,可用于单个样本中位数和总体中位数的差异比较,本节将该过 程称为单样本 Wilcoxon 检验。此时,研究假设是判断某随机样本位置参数(如中位数)是 否和已知总体位置参数相等。

当单样本 t 检验的数据不满足正态分布要求时,可采用单样本 Wilcoxon 检验进行替代 分析,如研究某矿泉水容量是否明显不等于 500 毫升。

【例 4-16】已知某地正常人尿氟含量的中位数为 2.15 毫摩尔/升。现在该地某厂随机 抽取了 12 名工人,测得尿氟含量(毫摩尔/升)如下: 2.15、2.10、2.20、2.12、2.42、 2.52、 2.62、2.72、3.00、3.18、3.87、5.67,试分析该厂工人的尿氟含量是否与正常人一 致。案例数据来源于李晓松 (2017),数据文档见"例 4-16.xls"。

1) 数据与案例分析

本例研究目的是比较随机抽取的 12 名工人尿氟含量的平均水平与已知中位数的差异, 通过在仪表盘中依次单击【通用方法】→【正态性检验】模块得出,该单组样本数据不服 从正态分布 (p<0.05)。考虑使用单一样本位置参数与已知总体位置参数的差异进行比较, 选择单样本 Wilcoxon 检验。 2) 秩和检验

在仪表盘中依次单击【实验/医学研究】→【单样本 Wilcoxon】模块,将【niaofu】变 量拖曳至【分析项(定量)】分析框中,在最上面的数值框中输出数字【2.15】。单样本 Wilcoxon 检验操作界面如图 4-21 所示,最后单击【开始分析】按钮。

♀ 实验/图	三学研究 🗸	开始分析	2.15	Q
配对卡方	二元Probit	niaofu		A defree adarma
Poisson回归	Cox回归			分析顷(定量)
ICC组内相关 系数	单样本 Wilcoxon			

图 4-21 单样本 Wilcoxon 检验操作界面

3) 结果分析

单样本 Wilcoxon 检验结果如表 4-49 所示。随机抽取的 12 名工人测得的尿氟含量中 位数为 2.570 毫摩尔/ 升,与正常人尿氟含量的中位数 2.15 毫摩尔/ 升相比, z=2.667, p=0.008<0.01,二者的差异具有统计学意义。

表 4-49 单样本 Wilcoxon 检验结果

名称	样本量	25 分位数	中位数	75 分位数	统计量 z 值	р
niaofu	12	2.188	2.570	3.045	2.667	0.008**

注: ** p<0.01。

4.4.3 两组独立样本 Mann-Whitney 检验

Mann-Whitney 检验全称为 Mann-Whitney U test, 中文翻译为曼-惠特尼 U 检验, 用于判断定量数据资料或有序分类变量资料的两个总体分布是否有差别。当 t 检验不满足正态分布条件时, Mann-Whitney 检验可用作独立样本 t 检验的非参数替代方法, 也可用作单向有序的 2×C 列联表差异检验。

在仪表盘中依次单击【通用方法】→【非参数检验】模块,可进行两组独立样本 Mann- Whitney 检验。

【例 4-17】对 10 例肺癌病人和 12 例矽肺病人,用 X 光片测量肺门横径右侧距 RD 值 (厘米),数据如表 4-50 所示,试分析两组人群的 RD 值有无差别。数据来源于陈平雁 (2005),数据文档见"例 4-17.xls"。

		1	
group	RD/厘米	group	RD/厘米
1	3.23	2	2.78
1	3.50	2	3.23
1	4.04	2	4.20
1	4.15	2	4.87
1	4.28	2	5.12
1	4.34	2	6.21
1	4.47	2	7.18
1	4.64	2	8.05
1	4.75	2	8.56
1	4.82	2	9.60
1	4.95		

表 4-50 例 4-17 案例数据

1	5.10	

1) 数据与案例分析

自变量"组别"的编码情况:数字 1 代表矽肺组,数字 2 代表肺癌组。因变量"RD 值"为 定量数据资料,本例的目的是比较两组 RD 值数据均数的差异,优先考虑用独立样本 t 检 验,通过在仪表盘中依次单击【通用方法】→【正态性检验】模块得出,两组 RD 值数据 不服从正态分布(假设不服从),因此考虑采用独立样本 Mann-Whitney 检验。

2) 秩和检验

独立样本 Mann-Whitney 检验操作界面如图 4-22 所示。在仪表盘中依次单击【通用 方法】→【非参数检验】模块,将【group】拖曳至【X(定类)】分析框中,将【RD】拖曳 至【Y(定量)】分析框中,本例为两组数据的差异比较,无须做多重比较,最后单击【开始 分析】按钮。

口 通用	防法 🗸	开始分析	Ē	不进行(默认)	~
相关	线性回归	group			X(定类)
方差	t检验				AULX)
单样本t检验	配对t检验	RD			Y(定量)
正态性检验	非参数检验				

图 4-22 独立样本 Mann-Whitney 检验操作界面

3) 结果分析

在仪表盘中依次单击【通用方法】→【非参数检验】模块,如果【X(定类)】为两个水 平的分类数据,则平台会自动执行 Mann-Whitney 检验。

两样本 Mann-Whitney 检验分析结果如表 4-51 所示。矽肺组 RD 值的中位数为 4.405, 肺癌组 RD 值的中位数为 5.665, z 统计量为-1.748, p=0.080>0.05, 矽肺组与肺癌组的 RD 值水平差异无统计学意义。

表 4-51 两样本 Mann-Whitney 格	🕸 金验分析结果
---------------------------	----------

ாக	group 中位数 M(P2	5, P75)	MannWhitney 检验	MannWhitney 检验	
坝	矽肺(<i>n</i> =12)	肺癌(<i>n</i> =10)	统计量 U值	统计量 z 值	ρ
RD	4.405(4.1,4.8)	5.665(4.0,8.2)	33.500	-1.748	0.080

4.4.4 多组独立样本 Kruskal-Wallis 检验

Kruskal-Wallis 检验全称为 Kruskal-Wallis H 检验,简称"K-W 检验"或"H 检验",中文一 般翻译为克鲁斯卡尔-沃利斯检验,是完全随机设计多个独立样本比较的秩和检验方法,它 用于推断定量数据或等级资料的多个独立样本来自的多个总体分布是否有差别。单因素方 差分析在不满足正态分布、方差齐性条件时,Kruskal-Wallis 检验可用作其非参数检验的替 代方法,也可用作单向有序的 R×C 列联表差异检验。

Kruskal-Wallis 检验总体达到显著后,可通过多重比较进一步比较各组的差异。常见的 非参数类多重比较方法有 Nemenyi 法、Dunn´s t 检验和 Dunn´s t 检验(校正 p 值)等,其 中 Nemenyi 法是完全随机设计多样本秩和检验多重比较的常用方法。

在仪表盘中依次单击【通用方法】→【非参数检验】模块,可进行多个独立样本的

Kruskal-Wallis 检验及 Nemenyi 法多重比较。

【例 4-18】随机抽取 3 组不同人群各 10 人,测定其血浆总皮质醇值(102 微摩尔/升, 非正态),如表 4-52 所示。请问 3 组不同人群的血浆总皮质醇测定值有无差别?案例数据 来源于李晓松(2017),数据文档见"例 4-18.xls"。

表 4-52 3 组人群的血浆总皮质醇测定值 (单位: 10² 微摩尔/升)

健康人	单纯性肥胖	皮质醇增多症
0.11	0.17	2.70
0.52	0.33	2.81
0.61	0.55	2.92
0.69	0.66	3.59
0.77	0.86	3.86
0.86	1.13	4.08
1.02	1.38	4.30
1.08	1.63	4.30
1.27	2.04	5.96
1.92	3.75	9.62

1) 数据与案例分析

"组别"为自变量,"皮质醇"为因变量。本例目的是比较多组样本的均值差异,优先考虑 用单因素方差分析。经正态分布检验发现数据严重不满足正态分布,考虑使用多组独立样 本 Kruskal-Wallis 检验。

2) 非参数检验

数据读入平台后, 在仪表盘中依次单击【通用方法】→【非参数检验】模块, 将【组 别】拖曳至【X(定类)】分析框中, 将【皮质醇】拖曳至【Y(定量)】分析框中, 本例为 3 组 数据的差异比较, 当总体结果的显著性 p<0.05 时, 应继续做多重比较。本节指定了一种 多重比较方法, 在下拉列表中选择【Nemenyi 法】进行多重比较, 最后单击【开始分析】 按钮。Kruskal-Wallis 检验操作界面如图 4-23 所示。



图 4-23 Kruskal-Wallis 检验操作界面

3) 整体显著性检验——Kruskal-Wallis 检验

Kruskal-Wallis 非参数检验的分析结果如表 4-53 所示。在 Kruskal-Wallis 检验中, H=18.130, p<0.001, 可认为 3 组人群的血浆总皮质醇测定值的差异有统计学意义。由 3 组的中位数大小可知, 皮质醇增多症组的中位数明显高于其他两组, 而另外两组的中位数 水平基本相当。各组两两之间的差异是否有统计学意义, 尚需多重比较结果来判断, 也可 通过直观地观察 3 组人群总皮质醇值的箱线图进行比较, 此处略。

项 健康人 (<i>n</i> =10)	<u><u></u></u>	且别中位数 <i>M</i> (P ₂₅ , F	Kruckal Wallie		
	健康人	单纯性肥胖	皮质醇增多症	松哈纮计量 日值	p
	(<i>n</i> =10)	(<i>n</i> =10)	(<i>n</i> =10)	位视频的 里 77 国	
皮质醇	0.815(0.6,1.1)	0.995(0.5,1.7)	3.970(2.9,4.7)	18.130	0.000**

表 4-53 Kruskal-Wallis 非参数检验的分析结果

注: ** p<0.01。

4) Nemenyi 法多重比较

如果 Kruskal-Wallis 检验的结论是多组分布差异显著,则需要进一步对各组间的分布 位置差异做两两比较。在【非参数检验】模块中必须要选择一个多重比较方法,本例选择 【Nemenyi 法】, Nemenyi 法两两比较的结果如表 4-54 所示。

项	(1)名称	(J)名称	(1)中位数	(J)中位数	差值(1-J)	p
	健康人	单纯性肥胖	0.815	0.995	-0.180	0.867
皮质醇	健康人	皮质醇增多症	0.815	3.970	-3.155	0.000**
	单纯性肥胖	皮质醇增多症	0.995	3.970	-2.975	0.003**

表 4-54 Nemenyi 法两两比较的结果

注: ** p<0.01。

本例有健康人组、单纯性肥胖组、皮质醇增多症组3个组别,两两进行比较共需要对比3次,表4-54中的3行结果即组间两两比较的 Nemenyi 法检验结果。

皮质醇增多症组与健康人组、单纯性肥胖组相比总皮质醇测定值均有显著差异(p< 0.01),尚不能认为单纯性肥胖组与健康人组的总皮质醇测定值存在差异(p=0.867>0.05)。

4.4.5 配对样本 Wilcoxon 检验

Wilcoxon 检验可用于配对样本差值的中位数和数字 0 的差异比较。目的是判断配对的 两个相关样本所来自的两个总体中位数是否有差别。

例如,对 11 份工业污水水样同时采用 A 法与 B 法进行测定,如果数据不满足正态分布,则可以通过配对样本 Wilcoxon 检验判断两法的测定结果有无差异。又如,对专项减脂训练前与训练后肝功能的理化指标数据进行测定,如果严重违反正态分布条件,则考虑采用配对样本 Wilcoxon 检验。

从功能上讲, 配对样本 Wilcoxon 检验与配对样本 t 检验完全一致, 二者的区别在于数据(配对数据的差值)是否满足正态分布。如果数据满足正态分布,则使用配对样本 t 检验;反之,则使用配对样本 Wilcoxon 检验。

【例 4-19】研究长跑运动对增强普通高校学生的心功能效果,对某校 15 名男生进行 测试,经过 5 个月的长跑锻炼后看其晨脉是否减少。案例数据如表 4-55 所示,试分析锻 炼前后的晨脉有无显著差异。数据来源于卢纹岱(2006),数据文档见"例 4-19.xls"。

锻炼前	锻炼后	锻炼前	锻炼后	锻炼前	锻炼后
70	48	56	55	75	56
76	54	58	54	66	48
56	60	60	45	56	64
63	64	65	51	59	50
63	48	65	48	70	54

表 4-55 例 4-19 案例数据

1) 数据与案例分析

配对资料不需要分组变量,锻炼前、锻炼后分别作为一个变量即可。本例目的在于比较两配对样本数据均值的差异,可选择配对样本 t 检验。假设长跑锻炼前后晨脉数据的差值不服从正态分布,此时配对样本 t 检验不再合适,可考虑非参数的配对样本 Wilcoxon 检

验代替配对样本 t 检验。

2) 非参数检验

读入数据后, 在仪表盘中依次单击【实验/医学研究】→【配对样本 Wilcoxon】模块, 操作界面如图 4-24 所示, 将【锻炼前】【锻炼后】变量分别拖曳至【配对 1(定量)】分析框、 【配对 2(定量)】分析框中, 最后单击【开始分析】按钮。

♀ 实验/医	学研究 🗸		开始分析	ê Q	
卡方检验	Kappa		锻炼前		配对1(定量)
配对卡方	二元Probit				
Poisson回归	Cox回归				
ICC组内相关	单样本 Wilcoxon				
配对样本 Wilcoxon	多样本 Friedman	€	锻炼后		配对2(定量)
游程检验	Kendall协调 系数				

图 4-24 配对样本 Wilcoxon 检验操作界面

3) 结果解读

配对样本 Wilcoxon 检验结果如表 4-56 所示。配对样本 Wilcoxon 检验的 z=2.842, p= 0.004<0.01, 锻炼前后的晨脉差异显著, 或锻炼前后的晨脉差异有统计学意义。

表 4-56 配对样本 Wilcoxon 检验结果

夕护	配对中位数	<i>M</i> (P ₂₅ , P ₇₅)	中位数 M的差值	纮壮景 → 佑	2
名称	配对 1	配对 2	(配对 1-配对 2)	501里 Z 但	ρ
锻炼前配对锻炼后	63.0(58.5,68.0)	54.0(48.0,55.5)	9.0	2.842	0.004**

注: ** p<0.01。

4.4.6 多样本 Friedman 检验

多样本 Friedman 检验也属于非参数检验方法,用于检验多个相关样本是否具有显著 差异的统计检验方法。在正态分布条件不满足时,也用来作为单个组内因素的重复测量方 差分析的替代方法。

例如,8 名受试对象在相同实验条件下分别接受4 种不同频率声音的刺激,4 种反应 率数据均来自同一个被试,因此有相关性。如果数据不满足正态分布,则可利用多样本 Friedman 检验来判断4 种刺激的反应率是否有差别。

【例 4-20】在某项随机区组设计的动物实验中,不同种系雌性家兔注射不同剂量雌激 素后子宫质量(克)的数据如表 4-57 所示。试比较 4 种剂量组雌性家兔子宫质量的差别 有无统计学意义。数据来源于颜虹和徐勇勇(2010),数据文档见"例 4-20.xls"。

YA	YB	YC	YD
63	54	138	188
90	144	220	238
54	92	83	300
45	100	213	140
54	36	150	175
72	90	163	300

表 4-57 例 4-20 案例数据

64	87	185	207

1) 数据与案例分析

相关性数据资料的数据格式为宽型数据, 4 种剂量组的测定数据分别作为单个变量, 依次为"YA""YB""YC""YD", 这 4 个变量间是有相关性的。

研究目的是,比较 4 种剂量雌激素注射后家兔子宫质量的总体分布是否相同。通过在 仪表盘中依次单击【通用方法】→【正态性检验】模块得出,4 组数据均服从正态分布。 数据先由宽型数据转换成长型数据(或重新录入为长型数据),再通过在仪表盘中依次单击

【通用方法】→【方差】模块进行方差齐性检验, F=3.47, p=0.03, 按 α=0.05 水平, 4 种 剂量组家兔子宫质量的总体方差不齐。所以本例不宜采用方差分析, 应采用多样本 Friedman 检验。

2) Friedman 检验

读入数据后, 在仪表盘中依次单击【实验/医学研究】→【多样本 Friedman】模块, 操作界面如图 4-25 所示, 将【YA】【YB】【YC】【YD】4 个变量拖曳至【分析项(定量)】分 析框中, 勾选【Nemenyi 两两比较】复选框, 最后单击【开始分析】按钮。



图 4-25 多样本 Friedman 检验操作界面

3) 整体显著性检验

多样本 Friedman 检验研究多个配对定量数据是否存在显著差异,结果如表 4-58 所示。 χ^2 =16.71, p<0.01,按 α =0.01 检验水平,认为注射不同剂量雌激素后,家兔子宫质量的差别有统计学意义。

名称	样本量	25 分位数	中位数	75 分位数	统计量 χ^2 值	р
YA	7	54.00	63.00	68.00		
YB	7	70.50	90.00	96.00	1671	0.00**
YC	7	144.00	163.00	199.00	10.71	0.00***
YD	7	181.50	207.00	269.00		

表 4-58 多样本 Friedman 检验结果

注: ** p<0.01。

4) 多重比较

按前面总结的方差分析流程(多组样本的非参数秩和检验类似流程),如果总体检验 p<0.05,则要对组间差异进行多重比较。【多样本 Friedman】模块默认提供的是 Nemenyi 法, Nemenyi 法多重比较结果如表 4-59 所示。

表 4-59 Nemenyi 法多重比较结果

(1) 名称	(J)名称	(ル)中位数	(J)中位数	中位数差值(<i>I-J</i>)	р
YA	YB	63.00	90.00	-27.00	0.82
YA	YC	63.00	163.00	-100.00	0.06
YA	YD	63.00	207.00	-144.00	0.00**

YB	YC	90.00	163.00	-73.00	0.35
YB	YD	90.00	207.00	-117.00	0.02*
YC	YD	163.00	207.00	-44.00	0.59

注: * p<0.05 ** p<0.01。

4 组数据任意两组进行配对比较,总共需要比较 6 次。YD 与 YA、YB 相比, p 均小于 0.05,差异有统计学意义,其他各组间无差异。结合中位数差值及输出的箱线图、统计图,结果表明,YD 组的家兔子宫质量高于其他 3 组,与 YA 组和 YB 组差异显著,但是 YD 组 与 YC 组无差异。

第5章 问卷数据相关影响关系研究

两个变量之间独立是指两个变量完全没有关系。两个变量之间相关是指两个变量存在 不确定性关系。此处相关关系不同于确定性关系,确定性关系一般指函数关系,若自变量 给定,则函数值确定。而不确定性相关关系,通俗理解即一个为自变量,另一个为随机变 量或两个都为随机变量,当一个变量发生变化时,另一个变量的取值有一定的随机性变化。

例如,研究者收集了同一家企业 1000 名职工的工资和受教育数据,每个人的受教育 年限不同,工资也不同,此时这两个变量都是随机变量。很多研究文献证实工资和受教育 年限存在相关性,或者说一名职工的工资受他接受教育时长的影响。当受教育年限改变时, 工资水平也随之改变,即不同的受教育程度使职工获得的工资水平不完全相同,有一定的 随机性。

职工工资受其受教育年限的影响,但工资与受教育年限并非严格的因果关系。我们可 以说工资与受教育年限存在相关性,也可以说受教育年限会影响职工工资。

本章介绍研究分析中的相关与影响关系,主要包括相关分析、线性回归、Logistic 回归 及非线性回归,相关与回归分析方法的使用场景如表 5-1 所示,接下来在各节内容中会结 合具体案例展开阐述。

方法	数据基本要求	使用场景举例
相关分析	定量数据或有序定类数据	销售收入与广告支出存在正相关关系,但销售收入并不由广告支出 唯一确定
线性回归	因变量必须是定量数据	分析广告支出、产品价格、产品质量、服务质量对销售收入的影响
Logistic 回归	因变量为定类数据	患病结局中 1 表示患病, 0 表示不患病, 研究是否患病的影响因素 分析
曲线与非线 性回归	主要针对定量数据	人口随时间(年)的变化而呈现 S 曲线增长过程,线性回归不再 适用

表 5-1 相关与回归分析方法的使用场景

5.1 相关分析

相关分析是用来研究两个或多个变量间的相关关系的统计学方法。例如,分析销售收 入与广告支出间的相关关系,或销售收入、广告支出、产品价格、产品质量、服务质量之 间的相关关系。变量可以是同等地位,也可以区分为一个或多个自变量对因变量的影响关 系。

5.1.1 相关关系概述

1.线性、非线性与无相关

相关关系按不同的划分方式可分为多种类型,结合直线趋势与关系密切程度将相关关系分为正线性相关、负线性相关、非线性相关、无相关4种,如图5-1所示。

两个变量在散点图中的点总体上呈一条直线,称为线性相关。图 5-1 (a) 表示一个变 量增加或减少,另一个变量相应地会线性增加或减少,称为正线性相关(简称正相关);图 5-1 (b)表示一个变量增加或减少,另一个变量反而会线性地减少或增加,称为负线性相关(简称负相关)。

83 / 187

两个变量在散点图中的点不是一条直线时,称为非线性相关或曲线相关,如图 5-1 (c) 所示。两个变量彼此互不影响,其数量变化各自独立,称为无相关,如图 5-1 (d)所示。



图 5-1 相关关系的类型

2.认识线性相关系数

线性相关包括正相关和负相关, Pearson 相关系数是衡量变量之间线性相关关系密切 程度和方向的统计指标。

Pearson 相关系数又称皮尔逊积差相关系数,通常被简称为相关系数,一般用 r 表示。 r 的取值为-1≤r≤1,绝对值越大表示相关密切程度越高。根据贾俊平(2014),对于一个 具体的 r 绝对值, r≥0.8 视为高度相关,0.5≤r<0.8 视为中度相关,0.3≤r<0.5 视为低度相 关,r<0.3 视为两个变量之间是弱相关。相关系数的强弱程度如表 5-2 所示。r 为正数时, 表示两个变量的变化方向一致,称为正相关;r 为负数时,表示两个变量的变化方向相反,称为负相关。

r 值(绝对值)	相关程度强弱
≥0.8	高度相关
$0.5 \le r < 0.8$	中度相关
0.3≤r<0.5	低度相关
<0.3	弱相关

表 5-2 相关系数的强弱程度

例如, 某省份国民收入与居民储蓄存款余额 r=0.9, 表明二者是高度正相关关系。若国民收入增加,则居民储蓄存款余额也随之增加;反之,若国民收入减少,则居民储蓄存款余额也随之减少。

在实际分析中,我们是基于样本数据计算相关系数的,用于估计总体相关系数时必定存在抽样误差,因此相关系数必须通过显著性检验才能说明其成立与否。一般先假设总体相关系数为 0,若相关系数的显著性概率 p 小于 0.05,则表明实际相关系数明显不等于 0,可说明相关关系存在。

3.如何选择线性相关系数

除 Pearson 相关系数之外,常用的相关系数还有 Spearman 相关系数、Kendall 相关系数。3 个相关系数可依据变量数据类型及正态分布条件进行选择,如图 5-2 所示。

Pearson 相关系数适用于两个变量均为定量数据的情况,要求数据服从二元正态分布, 通常我们简化为两个变量分别服从正态分布,并且无明显异常值,可以借助图形法或更为 严格的正态分布检验方法判断该条件,一般来说,不严重违反正态分布仍然可以继续使用 Pearson 相关系数,多数情况下结果较稳健。



图 5-2 相关系数的选择

Spearman 相关系数又称秩相关系数或等级相关系数,适用于定量数据或等级(有序 分类)数据,用两个变量的秩次大小做相关分析。其对数据分布没有明确要求,属于非参 数方法。在进行相关分析时,当 Pearson 相关系数不满足正态分布条件时, Spearman 相 关系数可用作 Pearson 相关系数的非参数替代。

Kendall 相关系数同样是用秩次进行相关分析的,也属于非参数方法,适用于连续性数 据或等级(有序分类)数据,主要用于两个有序分类变量的相关性,也称作和谐系数,可 用作一致性分析。

4.线性相关系数的报告

计算相关系数后,应对其准确解读和报告。一般来说解读和报告时按顺序包括以下3 个方面。

(1) 首先解读相关系数显著性检验的概率 p, 当 p<0.05 时认为存在相关关系。

(2) 然后解读相关系数的正负方向,若相关系数为正数则为正相关,若相关系数为负 数则为负相关。

(3) 最后解读相关系数绝对值的大小,明确相关关系的密切程度。

5.1.2 相关分析步骤

变量的数据类型,以及是否服从正态分布对相关分析过程有一定的影响,相关分析的 一般步骤如图 5-3 所示。



图 5-3 相关分析的一般步骤

1) 数据类型

本节主要介绍 Pearson 相关系数、Spearman 相关系数及 Kendall 相关系数针对定量数 据和有序分类数据的相关分析,如果遇到两个无序分类变量间的相关关系,则通过卡方检 验输出列联系数或 Phi and Cramer's v 系数来实现。

2) 散点图

散点图能直观地观察变量间的变化关系,开始相关分析前可先绘制两个变量的散点图,观 察和判断两个变量线性相关还是非线性相关,本节介绍的内容适用线性相关的情况。

3) 正态检验

对于定量数据可先通过直方图、正态 PP 图、QQ 图等图形法,或正态分布检验法判断 数据的分析状况。

4) 相关系数

根据正态分布条件及变量数据类型,对于两个定量数据且满足正态分布的情形,一般 选择 Pearson 相关系数,如果严重偏态则考虑使用 Spearman 相关系数;对于两个有序分 类数据,一般选择 Kendall 相关系数。

5) 结果解读

对相关系数的显著性、相关方向、相关程度进行正确解读。对多个数据进行两两相关 分析时,可报告和分析相关系数矩阵。

在呈现相关系数时常用"*"符号在相关系数 r 右侧标注其显著性水平, 当显著性检验的 p<0.05 时, 标注"*"; 当显著性检验的 p<0.01 时, 标注"**"; 当显著性检验的 p<0.001 时, 标注"**"; 当显著性检验的 p<0.001 时, 标注"***"。

5.1.3 两个变量相关实例分析

研究及分析两个变量间的相关关系较常见,如考察销售额与广告费的相关关系,计算 两个变量的相关系数即可;也可以同时考察多个变量两两之间的相关关系,如考察销售额、 广告费、产品单价、产品质量满意度、服务质量满意度两两之间的相关关系,共需要计算 10个相关系数,一般采用相关系数矩阵进行高效输出。

1.两个变量 Pearson 相关

【例 5-1】某研究收集到 758 名美国年轻男子的数据,变量包括"工资""教育年限""工龄""现雇佣年""年龄""智商""世界观""母亲受教育年限""婚否""是否住美国南方",以及"是否住大城市",试对"智商"与"世界观"进行相关分析。案例数据来源于陈强(2015),本例对数据进行了编辑及修改,数据文档见"例 5-1.xls"。

1) 数据与案例分析

"智商"与"世界观"均为测验评分定量数据,评分越高表示智商水平越高,对世界运行的 看法越客观深刻,按一般性常识,欲考察是不是智商水平越高的人他的世界观评分也相应 越高,此处采用两个变量间的相关系数来进行分析。

2) 散点图与正态检验

数据读入平台后, 在仪表盘中依次单击【可视化】→【散点图】模块, 以"智商"为自 变量, "世界观"为因变量, 将【智商】与【世界观】两个变量分别拖曳至【X(定量)】与 【Y(定量)】分析框中, 单击【开始分析】按钮绘制散点图。"智商"与"世界观"评分间的散 点图如图 5-4 所示。



图 5-4 "智商"与"世界观"评分间的散点图

由图 5-4 可知,"智商"与"世界观"之间无明显的非线性相关趋势,存在线性相关关系,可通过线性相关系数衡量关系的密切程度。

在仪表盘中依次单击【可视化】→【直方图】模块,绘制"智商"与"世界观"两个变量的 直方图,如图 5-5 所示,不难看出数据近似满足正态分布。



(a) "智商"的直方图 (b) "世界观"的直方图图 5-5 "智商"与"世界观"两个变量的直方图

综上,可使用 Pearson 相关系数考察"智商"与"世界观"两个变量间的线性相关关系。

3) 选择相关系数

读入数据后, 在仪表盘中依次单击【通用方法】→【相关】模块。具体分析时, 应根 据数据类型与正态分布情况选择恰当的相关系数。区分自变量与因变量时, 平台只输出一 个或多个自变量与因变量的相关系数; 不区分自变量与因变量时, 平台会以相关系数矩阵 的形式输出所有变量两两之间的相关系数。

本例"智商"为自变量,"世界观"为因变量,因此将【智商】和【世界观】分别拖曳至对 应的【分析项 X(定量)】和【分析项 Y(定量)】分析框中,同时选择下拉列表中的【Pearson 相关系数】,最后单击【开始分析】按钮。区分自变量与因变量的相关分析操作界面如图 5-6 所示。

□ 通用	方法 🗸		开始分析	Pears Pears	on相关系数 ~	2
频数	分类汇总		世界观	Spear	man相关系数	
描述	交叉(卡方)			Include		-
相关	线性回归				分忻墺Y(定量)
方差	t检验		智商			
单样本t检验	配对检验	Ð			11 JP 77 1 1 1 1 1 1 1 1	
正态性检验	非参数检验				分忻坝X(定量)	【刊选】

图 5-6 区分自变量与因变量的相关分析操作界面

4) 结果分析

"智商"与"世界观"的 Pearson 相关系数如表 5-3 所示。

表 5-2 "智商"与"世界观"的 Pearson 相关系数

项	世界观
智商	0. 340**

"智商"与"世界观"间的 Pearson 相关系数标记为两颗*, 等价于相关系数显著性检验 p<0.01, 说明相关系数有统计学意义, 二者存在线性相关关系。具体来说 r=0.340, 参考 表 5-2 所示的相关系数的强弱程度, 即"智商"与"世界观"间存在正向的低度相关关系, 可通 俗理解为年轻男子的"智商"越高其"世界观"评分也越高。

2.两个变量 Spearman 相关

【例 5-2】沿用【例 5-1】的数据文档,试分析"工龄"与"工资"的相关性。 1)数据与 案例分析"工龄"为受访者参加工作的年数,"工资"为工作薪资,均为定量数据。考察"工资" 与"工龄"是否存在线性相关关系,如是否工龄越长的人其工资水平也相应越高。

2) 散点图与正态检验

在仪表盘中依次单击【可视化】→【散点图】模块,绘制两个变量的散点图,如图 5-7 所示。"工龄"与"工资"无明显曲线关系,可能存在一定的线性关系。



图 5-7 "工龄"与"工资"间的散点图 本例通过在仪表盘中依次单击【可视化】→【P-P/Q-Q 图】模块,绘制"工龄"与"工资"

的正态 PP 图, 如图 5-8 所示。

图中的点偏离对角线较多,认为"工龄"与"工资"两个变量不服从正态分布,二者的线性 相关宜采用 Spearman 相关系数。

3) 相关分析与结果解读

在仪表盘中依次单击【通用方法】→【相关】模块,将【工龄】和【工资】分别拖曳 至【分析项 X(定量)】和【分析项 Y(定量)】分析框中,选择相关系数下拉列表中的 【Spearman 相关系数】,最后单击【开始分析】按钮。Spearman 相关系数如表 5-4 所示。



(a) "工龄"的正态 PP 图 (b) "工资"的正态 PP 图 图 5-8 "工龄"与"工资"的正态 PP 图

表 5-4	Spearman	相关系数
-------	----------	------

项	工资
工龄	0.074*

Spearman 相关系数标记为一颗*, 说明相关系数显著性检验 p<0.05, 相关系数有统计 学意义, 即二者存在线性相关关系。具体来说 r=0.074, 说明"工龄"与"工资"间存在正向的 弱相关关系。

3.多个变量两两间的相关系数矩阵

【例 5-3】继续沿用【例 5-1】的数据文件, 在【例 5-1】的基础上, 研究者想了解包 括 "母亲受教育年限""智商""世界观""年龄""教育年限""工龄""现雇佣年""工资"等 8 个变量两 两之间的相关关系,此时从研究目的上变量间没有明显的因变量、自变量区分,可输出相 关系数矩阵进行分析。



图 5-9 多个变量的相关分析界面

1) 数据与案例分析

8 个变量均为定量数据,通过绘制两两变量的散点图,变量间均存在线性相关关系。 通过图形法发现多数变量近似正态分布,8 个变量间的相关性本例采用 Pearson 相关系数 进行分析。

2) 相关分析

在仪表盘中依次单击【通用方法】→【相关】模块,将【母亲受教育年限】【智商】 【世界观】【年龄】【教育年限】【工龄】【现雇佣年】【工资】8 个变量统一拖曳至【分析项 Y(定量)】分析框中,选择下拉列表中的【Pearson 相关系数】。多个变量的相关分析界面如 图 5-9 所示。最后单击【开始分析】按钮。

3) 结果分析

本例输出的是相关系数矩阵,标准格式如表 5-5 所示。

变量	平均值	标准差	母亲受 教育 年限	智商	世界观	年龄	教育 年限	工龄	现雇 佣年	工资
母亲受 教育 年限	10.910	2.741	1.000							
智商	103.856	13.619	0.263**	1.000						
世界观	36.574	7.302	0.195**	0.340**	1.000					
年龄	21.835	2.982	0.103**	0.178**	0.444**	1.000				
教育 年限	13.405	2.232	0.340**	0.513**	0.381**	0.448**	1.000			
工龄	1.735	2.106	-0.159**	-0.166**	0.043	0.388**	-0.242**	1.000		
现雇 佣年	1.831	1.674	-0.048	0.019	0.182**	0.339**	-0.050	0.231**	1.000	
工资	322.820	145.685	0.193**	0.344**	0.306**	0.533**	0.500**	0.069	0.166**	1.000

表 5-5 相关系数矩阵标准格式

注: ** p<0.01。

该表格第 2 列和第 3 列给出了各变量的平均值与标准差指标,两两变量间的相关系数, 以相关系数矩阵形式呈现。由表 5-5 可知,"现雇佣年"与"母亲受教育年限"、"现雇佣年"与 "智商"、"工龄"与"世界观"这 3 组无线性相关关系(未标记*符号,均 p>0.05),其他变量间 存在不同程度的线性相关关系(均 p<0.05)。

5.1.4 偏相关实例分析

偏相关是指消除第3个变量C的作用后研究A和B这两个变量间的线性相关关系。用 来衡量偏相关的相关系数一般称作偏相关系数,特指在相关分析中,消除其他变量影响的 条件下,计算某两个变量之间的相关系数。例如,只分析身高与肺活量间存在线性相关可 能得到错误结论,而先控制体重的影响再研究身高和肺活量的关系更为科学、稳妥。

1.偏相关分析注意事项

第3个变量C同样是定量数据类型,一般要求该变量与A和B均存在相关关系。

偏相关分析并不是新的相关分析方法,其仍然是线性相关分析,只是多了第 3 个需要 控制或排除的干扰因素。

2.一般分析步骤

在偏相关关系实际研究分析中,偏相关分析步骤如图 5-10 所示。

(1) 正态分布条件判断。对 A、B、C 3 个变量数据进行正态分布检验, 若服从正态

分布则以 Pearson 相关系数衡量相关性,否则采用 Spearman 相关系数。

(2) 控制条件判断。对 A、B、C 3 个变量进行简单相关分析, 作为条件判断第 3 个 变量 C 是否与 A 和 B 均存在相关关系。



图 5-10 偏相关分析步骤

(3) 偏相关分析。确认 C 需要控制后,进行偏相关分析,并对偏相关结果进行解释和分析。

3.偏相关实例分析

【例 5-4】研究者收集了 29 名学生的身高、体重、肺活量数据,专业上认为身高、体 重都和肺活量有一定的相关关系,试分析身高与肺活量之间的相关关系。数据来源于卢纹 岱 (2006),数据文档见"例 5-4.xls"。

1) 数据与案例分析

数据文档中的身高、体重、肺活量均为定量数据,通过图形法观察认为 3 个变量近似 服从正态分布,接下来用 Pearson 相关系数来衡量相关关系。

专业上认为肺活量和身高、体重均相关,要研究身高与体重的关系则应当控制体重的 影响。偏相关分析由独立的模块完成,读入数据后,在仪表盘中依次单击【进阶方法】→ 【偏相关】模块。

将【身高】【肺活量】拖曳至【分析项(定量)(至少 2 个)】分析框中,将【体重】拖曳 至【控制变量(至少 1 个)】分析框中,最后单击【开始分析】按钮。偏相关分析操作界面 如图 5-11 所示。

∅ 进阶	访法 🗸		开始分析	<u>n</u> Q
聚类	因子		身高	
主成分	分层回归		肺活量	分析面(完景)(至少2个)
逐步回归	二元Logit			73 III X (ALE) (LL) = 17
多分类Logit	有序Logit			
事后多重比较	偏相关	e	体重	
典型相关	双因素方差	1		控制变量(至少1个)

图 5-11 偏相关分析操作界面

2) 控制条件判断

偏相关分析要求第 3 个变量与待研究的两个变量间存在线性相关性,是一个基本条件, 所以首先要分析控制变量与其他两个变量的相关性。控制变量相关性条件判断如表 5-6 所 示。

项	身高	肺活量
体重	0.741**	0.751**

表 5-6 控制变量相关性条件判断

注: **p<0.01。

体重与肺活量、体重与身高的相关系数分别为 0.751**、0.741**, 均有统计学意义, 说明体重同时与待研究的重点变量间存在强正相关关系, 研究身高与肺活量间的关系, 需 要消除体重的干扰影响。

3) 偏相关结果分析

偏相关分析结果如表 5-7 所示,身高与肺活量间的偏相关系数 r=0.098,未标记*符号, 说明 p>0.05,无相关性,即当控制体重的影响后,实际上身高与肺活量之间没有相关性, 该结论和未控制体重的结果完全相反。

项	平均值	标准差	身高	肺活量
身高	152.593	8.356	1.000	
肺活量	2.190	0.451	0.098	1.000

表 5-7 偏相关分析结果

分析总结:如果不消除体重的影响,那么身高与肺活量呈强正相关关系(r=0.6**,单 独做相关分析),消除体重影响后发现身高与肺活量间实际无相关性(r=0.098, p>0.05)。

5.2 线性回归

当因变量与自变量间存在线性相关关系时,可以使用线性回归分析方法确定它们之间 相互依赖的定量关系。此处所说的定量关系,并非严格的因果关系,而是自变量 X 对因变 量 Y 的影响或预测作用。

例如,研究者分析广告费、产品单价、产品满意度、服务满意度对销售收入的影响, 如果各自变量与因变量间存在线性相关关系,则可建立以销售收入作为因变量,其他 4 个 变量作为自变量的线性回归模型,用于分析它们之间的线性相关关系,或者说研究销售收 入的影响因素,并对销售收入进行预测。

本节主要介绍线性回归模型的相关概念,并通过具体实例对多重线性回归、哑变量线 性回归做具体研究及分析,通过在仪表盘中依次单击【通用方法】→【线性回归】模块或 依次单击【进阶方法】→【逐步回归】模块来实现。

5.2.1 线性回归模型与检验

1. 归模型与种类

线性回归可通过回归函数定量化地解释自变量与因变量的关系,这种回归函数被称作 线性回归模型,用样本数据估计所得的回归方程表达式如下:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon$$

上式中, Ŷ为因变量的估计值, β_0 为常数项, 也叫截距。 X_1 、 X_2 、 X_p 为自变量, β_1 、 β_2 、 β_p 称为偏回归系数, 它表示其他自变量不变时指定的某自变量 X 每变动一个单位时因 变量 Y 的平均变化量。 ε 称为残差, 是因变量真实取值与估计值之间的差值, 是一个随机 变量。

一般用普通最小二乘法(记为 OLS),通过样本数据估计未知的 , 拟合一条直线使 各样本点与直线的纵向距离最小。

根据线性回归中自变量的个数多少进行分类,当线性回归中仅有一个自变量时,称作

一元线性回归,如研究产品质量评分与顾客满意度间的相关关系;当线性回归中有多个自 变量时,称作多重线性回归(通常叫作多元线性回归),其回归系数称为偏回归系数。例如, 以销售收入为因变量,同时研究广告费、产品单价、产品满意度、服务满意度与因变量销 售收入的相关关系。

多重线性回归根据自变量引入模型和筛选方式,可以分为强制引入法多重线性回归和 逐步多重线性回归。

2.模型检验和评价

拟合线性回归模型后,要对模型总体拟合状况进行检验和评价,通过检验后方可用于 影响因素分析或回归预测。线性回归模型的检验如表 5-8 所示。

检验项目	检验方法或指标
回归方程总体显著性检验	F 检验(ANOVA)
回归系数显著性检验	<i>t</i> 检验
回归方程拟合优度评价	一元线性回归 R ² 、多重线性回归调整后的 R ²

回表 5-8 线性 归模型的检验

1) 回归方程总体显著性检验

研究者采用方差分析 F 检验,对回归方程总体是否显著(有统计学意义)进行检验。 该检验原假设回归方程中至少有一个自变量的回归系数不为 0,当回归模型 F 检验的概率 p 小于 0.05 时说明模型显著,即至少有一个自变量对因变量的影响有统计学意义;反之, 若 p 大于 0.05 则说明模型不成立。

2) 回归系数显著性检验

回归方程总体显著,如果想进一步判断哪些自变量的回归系数是显著的,则需要进行 t 检验。原假设自变量回归系数等于 0,如果回归系数 t 检验概率 p<0.05,则说明该变量 回归系数不为 0,回归系数有统计学意义,其对因变量有显著影响;反之,若 p>0.05,则 说明该自变量的回归系数为 0,自变量的影响无统计学意义。

3) 回归方程拟合优度评价

拟合优度是指样本数据各点围绕回归直线的密集程度,用来评价回归模型的拟合质量。 一般是用决定系数 R2 作为评价指标, R2 接近 1 说明回归方程拟合优度良好, R2 接近 0 说明回归方程拟合优度差。 R2 一般解释为回归方程对因变量 Y 总变异的解释力度,如 R2 为 0.8,即回归方程可解释因变量 Y 总变异原因的 80%。

一元线性回归时仅输出 R2, 多重线性回归时可同时输出 R2 和调整后的 R2 。 R2 会随自变量的个数或样本量的增加而增大,为了消除这种影响,引进了调整后的 R2, 因此 多重线性回归时决定系数用调整后的 R2。

5.2.2 线性回归适用条件

线性回归对数据资料是有要求的,因变量必须是定量数据,自变量可以是定量数据也可以是定类数据,当遇到分类数据为自变量时,要根据实际情况考虑以哑变量形式进行线性回归。线性回归的正确使用,还应满足以下的主要适用条件,如表 5-9 所示。

适用条件	说明	检验方法/指标
线性关系	自变量 X 与因变量 Y 间存在线性关系	前验: 散点图、相关分析
残差正态性	回归残差服从正态分布	后验:残差直方图、残差 QQ 图/PP 图
残差等方差性	回归残差的方差齐	后验: 残差散点图 (残差诊断图)

回表 5-9 线性 归的主要适用条件

残差独立性	回归残差满足独立性特征	后验: D-W 检验
无多重共线性	多重线性回归自变量无明显共线性的影响	后验: VIF(方差膨胀因子)
无明显异常值	无明显异常值对回归结果不利	后验:残差散点图/箱线图、标准差等

前验是指在线性回归开始之前进行的检验,后验是指在线性回归后利用回归结果(如 残差)进行的检验。残差即因变量的观测值与利用回归模型求出的预测值之间的差值,能 反映利用回归模型进行预测引起的误差。

1) 线性关系

自变量与因变量间存在线性关系,这是线性回归最基本的条件之一。一般在开始线性 回归之前,通过绘制自变量与因变量的散点图或进行两者的相关分析可加以判断。如果自 变量与因变量是非线性关系,那么需要先进行数据转换,再进行线性回归或曲线回归。

2) 残差正态性

线性回归模型要求其残差服从均值为 0、方差为 2 的正态分布,回归拟合后对其残 差进行正态分布检验,常用方法有残差直方图、残差 PP 图/QQ 图等,或采用显著性检验 方法。如果残差不服从正态分布,则考虑对因变量进行正态转换使其满足条件,如先对因 变量取对数函数再重新进行回归分析。

3) 残差等方差性

残差等方差性即残差齐次,理论上是指自变量取值不同时,因变量 Y 的方差相等,可 通俗理解为不同 Y 预测值情况下,残差的方差相等。研究者可利用残差数据绘制残差散点 图,用以观察残差随因变量取值或随预测值的变化趋势。如果残差随机分布,无明显规律 可循,则说明残差等方差;如果残差的分布有迹可循,如自变量 X 值越大,残差项越大或 越小,常见的残差呈"喇叭状",则说明残差分布不均,模型具有异方差性,模型质量较差。 如果模型有明显的异方差性,则建议处理后重新进行回归分析,如先对 Y 取对数再构建模 型等。

4) 残差独立性

针对回归残差的独立性条件,研究者通常采用 Durbin 和 Watson 提出的 D-W 检验方法。如果 D-W 值在 2 附近(1.7~2.3),则说明残差独立,即没有自相关性;反之,若 D-W 值明显偏离 2,则说明存在自相关性。如果有明显的自相关性,则考虑对因变量进行差分处理或更换分析方法。

5) 无多重共线性

线性回归中的多重共线性,是指线性回归模型中的自变量之间由于存在强相关关系而 使模型估计失真或难以准确估计。对共线性的判断,常见方法有分析变量间的相关性及排 查方差膨胀因子(VIF)值。如果自变量中出现 VIF 大于 10(严格一些也可以将 VIF 大于 5 作为标准),则说明模型中存在严重的多重共线性问题,模型结果不可靠;反之,若 VIF 小 于 10(或小于 5),则说明模型存在的共线性问题不严重。如果模型呈现共线性问题,则 考虑使用逐步回归分析、主成分回归,或者利用专业经验及变量间的相关性删除个别自变 量。

6) 无明显异常值

由于异常值的存在对于回归直线方程的拟合、判定系数及显著性检验的结果都有很大 的影响,因此对线性回归中异常值的分析不容忽视,可通过残差散点图进行观察,如果发 现有明显的离群点应当重视,必要时予以删除或替换处理。

5.2.3 线性回归的一般步骤

在实际科研分析时,线性回归分析的一般步骤如图 5-12 所示。



图 5-12 线性回归分析的一般步骤

1) 准备数据

按普通数据格式录入数据,即每行是一个个案,每列是一个变量。线性回归的因变量 必须是定量数据,如果因变量是定类数据,则应该进行 Logistic 回归。线性回归的自变量 允许是定量数据或定类数据,分类数据可先通过在仪表盘中依次单击【数据处理】→【生 成变量】模块进行哑变量处理,再参与回归。

2) 线性条件判断

自变量与因变量间存在线性相关关系,这是基础条件。如果自变量与因变量为非线性 关系,则考虑先对数据进行转换或进行曲线回归。对于多重线性回归一般需要逐个观察各 自变量与因变量的散点图是否有线性趋势,或进行线性相关分析,利用线性相关系数进行 判断。

3) 建立线性回归模型

如果只有一个自变量,则可进行一元线性回归,如果有多个自变量,则可进行多重线 性回归。当进行多重线性回归时,可根据自变量的个数及自变量的筛选方式,选择逐步回 归方式对自变量进行筛选。

4) 模型检验与评价

构建模型后,我们需要对模型整体进行显著性检验,判断模型是否有效,并进一步对 各自变量的回归系数进行显著性检验,利用 R2 或调整后的 R2 对模型拟合优度进行评价, R2 的大小没有固定标准,应结合行业经验或文献资料进行评判。

5) 残差及共线性诊断

这一部分分析内容可以称为后验分析,利用回归残差进行残差正态性、残差等方差性、 残差独立性诊断,判断残差是否满足相应要求。同时对回归模型有无多重共线性问题、有 无明显异常值影响进行判断和处理。

在后验分析中,如发现非正态、异方差、自相关、共线性、异常值等问题,要进行有 针对性的处理,常见的有对因变量做对数函数转换、剔除某些引发共线性问题的自变量, 或剔除个别异常值数据后,返回第3步重新拟合线性回归。

6) 结果报告

综合显著性检验、模型拟合优度、残差诊断、异常值判断等结果,对回归结果进行解 读、分析和报告。侧重点在回归模型的有效性、回归系数的解释、确认影响因素,有时也 会写出线性回归方程予以报告。

5.2.4 多重线性回归的实例分析

在实际研究中,因变量的变化往往受多个因素的影响,因此多重线性回归更能反映变 量间的客观关系。

1.一元线性 归与多重线性 归的区别

根据自变量的个数,线性回归可分为一元线性回归与多重线性回归。一元线性回归与 多重线性回归对数据的要求、适用条件、模型检验、预测等基本是一致的,二者在应用时 的主要区别如表 5-10 所示。

线性回归	一元线性回归	多重线性回归
模型举例	<i>y</i> =6.6+0.8 <i>x</i>	$y=6.6+0.5x_1+0.1x_2+0.2x_3$
自变量	仅一个自变量	多个自变量
自变量选择	强迫法	强迫法、逐步法
共线性	无	自变量间要求无多重共线性影响
回归系数	回归系数 X 每改变一个单位,Y随着改变多少个单位	偏回归系数 控制其他自变量保持不变的情况下,X每改 变一个单位对因变量Y均值的影响程度
模型评价	R ²	调整后的 R ²
应用范围	一般作为中间分析过程	更广泛,用作变量关系研究的依据

表 5-10 一元线性回归与多重线性回归的主要区别

(1)一元线性回归:只包括一个自变量,不论其有无显著性意义都必须纳入模型并建 立回归方程,自变量的选择主要受专业知识和研究目的的影响,不存在多重共线性、逐步 回归筛选的问题。回归系数表示自变量每改变一个单位,因变量随之发生的改变量,回归 模型的评价指标可采用决定系数 R2,通常作为多重线性回归或其他统计方法前的初步分 析。

(2)多重线性回归:可以分析多个自变量对因变量的影响关系,也可以靠专业知识选择少数重要的自变量进行强迫法回归分析,还可以通过逐步回归法从众多自变量中筛选对因变量有显著影响的自变量进行分析,并要求多个自变量间不能有多重共线性。回归系数表示控制其他自变量保持不变的情况下,自变量每改变一个单位对因变量均值的影响程度。在采用 R2 对模型进行拟合评价时,应对 R2 进行校正,因此一般采用调整后的 R2。

因变量的变化是系统性的,与一元线性回归相比较,多重线性回归更符合实际研究情况,使用范围更广泛,当因变量为定量数据时,常作为变量影响关系的依据。

2.线性回归实例

【例 5-5】本例以实例分析的形式,结合图 5-12 所示的步骤对线性回归分析进行进一步阐述。继续沿用【例 5-1】的数据文档,拟建立以"年龄""教育年限""工龄""现雇佣年"为自变量,"工资"为因变量的多重线性回归模型。案例数据来源于陈强(2015),本例对数据进行了编辑及修改,数据文档见"例 5-5.xls"。

一元线性回归的操作步骤、结果分析和多重线性回归基本一致,此处不再单独进行实 例介绍。



1) 准备数据

从研究目的上分析"年龄""教育年限""工龄""现雇佣年"4 个自变量对"工资"的相关性, 或 对"工资"的影响, 由于"工资"变量为定量数据, 因此选用线性回归分析。如果因变量为分类 数据, 则选用 Logistic 回归分析。

2) 线性条件判断

在 5.1 节已经介绍过用散点图和相关系数来判断两个变量间是否存在线性相关关系。 通过在仪表盘中依次单击【可视化】→【散点图】模块来绘制散点图,因为 4 个自变量在 量纲单位上比较接近,所以本例在一个直角坐标系中绘制了 4 个散点图(量纲单位差距较 大时,可采用矩阵排列)。4 个自变量与因变量的散点图如图 5-13 所示。



图 5-13 4 个自变量与因变量的散点图(续)

通过【直方图】模块初步判断正态性后, 在仪表盘中依次单击【通用方法】→【相关】 模块, 以"工资"为因变量进行相关分析, 结果发现"年龄""教育年限""工龄""现雇佣年"与因 变量"工资"的 Spearman 相关系数依次为 0.51、0.49、0.1、0.07, p 均小于 0.05(至少标记 一个*符号), 且均为正相关关系。综合散点图和 Spearman 相关系数, "工资"与"年龄"和 "教育年限"存在强线性相关关系, 而与"工龄"和"现雇佣年"存在弱线性相关关系。

3) 建立线性回归模型

本例考察的 4 个自变量均为定量数据,可直接进行线性回归分析。在仪表盘中依次单击【通用方法】→【线性回归】模块,将【工资】拖曳至【Y(定量)】分析框中,将其他 4 个变量拖曳至【X(定量/定类)】分析框中,勾选【保存残差和预测值】复选框,命令平台计算残差和预测值用于后验分析,多重线性回归操作界面如图 5-14 所示,最后单击【开始分析】按钮。

口 通用	防法 🗸		开始分析	保存预	线差和预测值 🔽	ΘQ
频数	分类汇总		工资			Y(定量)
描述	交叉(卡方)		年龄			
相关	线性回归		教育年限			
方差	t检验		工龄		X(定	量/定类)
单样本t检验	配对检验	Ð	现雇佣年			
正态性检验	非参数检验					

图 5-14 多重线性回归操作界面

线性回归结果表格较多,包括线性回归分析结果、ANOVA 表格、coefPlot 图等,我们可以按步骤进行解读和分析。

4) 回归模型检验与评价

(1)回归方程总体显著性检验。回归模型是否有效采用方差分析 F 检验,若该检验的 概率 p<0.05,则说明模型显著,即当前回归模型中至少有一个自变量的回归系数不为 0, 对因变量 Y 有显著影响;反之,模型无统计学意义。

线性回归模型总体显著性检验如表 5-11 所示, F=111.78, p<0.05, 表明该线性回归 模型总体上有统计学意义。

项	平方和	df	均方	F	р
回归	5982995.92	4	1495748.98	111.78	0.00
残差	10062378.19	752	13380.82		
总计	16045374.11	756			

表 5-11 线性回归模型总体显著性检验

(2) 偏回归系数显著性检验。某个自变量的 t 检验概率 p 如果小于 0.05,则说明该 变量系数不为 0,偏回归系数有统计学意义,其对因变量有显著影响;反之,则说明偏回 归系数无统计学意义。

偏回归系数(n=757)如表 5-12 所示, 第 2 列、第 4 列分别给出了非标准化系数 B 与标准化系数 Beta, 第 5、6 列为各自变量偏回归系数的 t 检验结果, 第 7 列为偏回归系数的 95% CI, 第 8 列为共线性指标 VIF (下文有相应解释及分析)。

TE	市 非标准	准化系数 标准化系数		4		05% 01		
坝	В	标准误	Beta	Ĺ	ρ	95% CI		
常数	-370.70	33.98	—	-10.91	0.00**	-437.29~-304.11	—	
年龄	17.59	2.05	0.36	8.59	0.00**	13.58~21.61	2.10	
教育年限	22.37	2.50	0.34	8.94	0.00**	17.47~27.28	1.76	
工龄	-0.03	2.51	-0.00	-0.01	0.99	-4.95~4.89	1.58	
现雇佣年	5.35	2.75	0.06	1.95	0.05	$-0.04 \sim 10.74$	1.20	

回表 5-12 偏 归系数 (n=757)

注: 1.因变量: 工资。

2 . ***p*<0.01_°

t 检验结果显示, "年龄""教育年限"的 p 值均小于 0.01, "工龄"的 p 值为 0.99 大于 0.05, "现雇佣年"的 p 值等于 0.05,均不显著,无统计学意义。常数项的显著性一般无实际意义, 可以不做解读。本例所考察的 4 个自变量,"年龄""教育年限"对"工资"有显著影响,而另外 两个变量对"工资"无影响。

"年龄"变量的非标准化系数为 17.59, 说明"年龄"和"工资"存在正相关关系, 在其他变量不变的情况下, 年龄每增长 1 岁, 工资水平会相应提高 17.59 个单位。"教育年限"非标 准化系数为 22.37, 表明在其他变量不变的情况下, 多接受一年教育, 工资水平会相应提 高 22.37 个单位。

利用非标准化偏线性回归系数可写出回归方程的表达式,即

Y=-370.70+17.59×年龄+22.37×教育年限-0.03×工龄+5.35×现雇佣年

标准化偏线性回归系数可用其绝对值对各自变量进行重要性排序。例如,本例 4 个自 变量的重要性依次为"年龄""教育年限""现雇佣年""工龄"。 (3)回归方程拟合度评价。模型汇总表能列出诸多和模型拟合有关的指标。线性回归 模型拟合评价如表 5-13 所示。

R	R ²	调整后的 R ²	模型误差 RMSE	D-W 值	AIC 值	BIC 值
0.61	0.37	0.37	115.29	1.85	9345.95	9369.10

表 5-13 线性回归模型拟合评价

R 即复相关系数,也叫作多元相关系数,反映的是所有自变量与因变量的总体相关关系。AIC 为赤池信息量准则,BIC 为贝叶斯信息准则,这两个指标一般用于多个回归模型的优劣比较。RMSE 为均方根误差,和原始数据的单位一致,解释结果时更容易理解,一般用于多个模型的比较。R、AIC 值、BIC 值、RMSE 这 4 个指标不常用,本例不做解读。D-W 值用于残差独立性检验,稍后在第5步进行解读和分析。

一元线性回归时以 R2 作为拟合优度评价指标,多重线性回归时则采用调整后的 R2 作 为拟合优度评价指标。本例调整后的 R2 为 0.37,表示建立的回归模型可解释"工资"总变异 信息的 37%。R2 越接近 1 越能说明模型的解释能力,不同行业及领域对 R2 的接受不尽相 同,没有绝对的标准,可参考行业文献进行判断。

5) 残差及共线性等条件诊断

在 SPSSAU 平台分析界面打开【我的数据】模块,在当前数据集名称右侧,单击【查 看】按钮可查看数据,即可观察平台计算并另存为新变量的残差值和预测值数据。查看数 据集中新生成的残差和预测值如图 5-15 所示,左侧第 1 列"Regression_Prediction_XXXX"即 回归方程对现有数据的预测值(以下简称 Prediction),第 2 列"Regression_Residual_XXXX" 即本次回归的非标准化残差值(以下简称 Residual)。后验分析中对残差的分析需要使用这 两列新生成的变量,它也会自动出现在待分析模块的变量列表中。

第1列	第2列	第3列
Regression_Prediction_9936	Regression_Residual_9936	南方否
231.98866653897704	133.04884346102295	0
402.5638503917339	-172.58204039173393	0
299.6786714163149	2.1924185836850825	0
219.75436970031416	20.332280299685834	0
303.77946079629237	71.2483092037076	0
117.45452330880805	4.542896691191956	0
591.8164508571579	81.35499914284219	0
374.7968413584697	-41.84426135846968	0
295.41384801696785	14.718751983032178	0

图 5-15 查看数据集中新生成的残差和预测值

(1)残差正态性。本例采用操作简单的直方图,在仪表盘中依次单击【可视化】→ 【直方图】模块,将【Residual】拖曳至【分析项(定量)】分析框中,单击【开始分析】按钮。

线性回归残差直方图如图 5-16 所示,直方图呈现左右对称的形态,比较接近正态分 布,认为其近似正态分布,残差满足正态分布的要求。



图 5-16 线性回归残差直方图

(2)残差等方差性。一般利用残差数据绘制残差散点图,观察样本点是否随机分布。 在仪表盘中依次单击【可视化】→【散点图】模块,将【Residual】拖曳至【Y(定量)】分 析框中,将【Prediction】拖曳至【X(定量)】分析框中,此处注意散点图 X 轴及 Y 轴上分 配的变量,一般来说残差诊断的散点图,会将其预测值作为 X 轴,残差值作为 Y 轴,最后 单击【开始分析】按钮。

线性回归预测值与回归残差值散点图如图 5-17 所示。点的分布并不是随机的,随着 回归预测值的增大,残差有"逐渐放大"的趋势,呈现开口向右的"喇叭状"形态,提示本次建 立的线性回归模型残差不满足等方差性,存在残差异方差的问题,这对线性回归过程是不 利的,影响结果的准确性,应当重视并想办法予以处理。



图 5-17 线性回归预测值与回归残差值散点图

常见的处理方式可先对回归分析的因变量进行对数函数的变换,再重新建立线性回归 模型,此分析过程稍后进行。

(3) 残差独立性。残差独立性通常用 D-W 检验方法,如果 D-W 值在 2 附近(1.7~2.3),则说明残差独立。D-W 检验结果在线性回归中能自动计算并输出。

本例的 D-W 值为 1.85,在 1.7~2.3 范围内,认为残差独立(见表 5-13)。

(4)无多重共线性。统计实践中常通过 VIF 指标来快速判断有无共线性问题,一般认为 VIF 大于 10 (较严格的标准是大于 5),说明线性回归模型有较严重的共线性问题。表 5-12 中的最后一列即 VIF,本例分析了 4 个自变量的 VIF,均介于 1.2~2.1,小于 10,因此本例不考虑共线性问题。

(5) 无明显异常值。回归分析的异常值可通过残差散点图进行直观的观察及判断(见

图 5-17), 散点图顶部和底部有极少数点距离偏远, 值得关注。一般来说对异常值的判断 应谨慎, 需要结合专业认知对这些点进行判断, 本例暂不处理。

6) 回归结果分析

在以上步骤中,对以"年龄""教育年限""工龄""现雇佣年"为自变量,"工资"为因变量的 多重线性回归模型进行了显著性检验、后验分析、共线性检验、异常值等初步分析,目前 的结论是,模型总体上有统计学意义,其中"年龄""教育年限"对"工资"有显著影响,但是残 差存在异方差性,结合因变量"工资"不服从正态分布的特征,本例考虑先对因变量进行对 数函数转换,再重新构建多重线性回归模型。

7) 重新建立回归模型

首先通过在仪表盘中依次单击【数据处理】→【生成变量】模块,再选中【工资】变 量,然后在【生成变量】下方的下拉列表中选择【自然对数(Ln)】,单击底部的【确认处理】

第1列	第2列
Ln_工资	Regression_Prediction_9936
5.9000001154256	231.98866653897704
5.438000218839135	402.5638503917339
5.710000071942002	299.6786714163149
5.480999899848799	219.75436970031416
5.9270000765619315	303.77946079629237
4.80399989696866	117.45452330880805
6.51200005202935	591.8164508571579
5.808000077437857	374.7968413584697
5.7369999479591725	295.41384801696785

按钮,平台会自动计算并将新的对数数据保存到数据文件中。对"工资"进行自然对数转换 如图 5-18 所示。

生成变量 💿 🧕 高级公式	D
自然对数(Ln)	~
常用	
平均值	
求和	
虚拟(哑)变量	
Z标准化(S)	
中心化(C)	
乘积(交互项)	_
自然对数(Ln)	
10为底对数(Log10)	
量纲处理	
归—化(MMS)	

图 5-18 对"工资"进行自然对数转换

新变量"Ln_工资"即"工资"数据的自然对数函数转换值。接下来以"年龄""教育年限""工 龄""现雇佣年"为自变量,以"Ln_工资"为因变量,重新拟合多重线性回归模型,并计算新模 型的残差值和预测值。基于残差与预测数据,重新绘制参数散点图并进行诊断。相关操作 和前面第一次线性回归一致,本节不再赘述,可直接进行主要结果的解释分析和报告。为 和上一次回归结果相比较,两次模型分别简称为模型1和模型2。

图 5-19 所示为模型 2 的残差散点图,显然因变量进行对数转换后,模型 2 的残差分 布得到了明显改善。点的分布随机化、均匀化,无明显趋势规律,可认为回归残差等方差。 其他关于残差正态性、独立性的条件同样满足。



图 5-19 模型 2 的残差散点图

模型 2 线性回归分析汇总表 (n=757) 如表 5-14 所示, F(4,752)=111.46, p<0.05, 回 归模型 2 有统计学意义。其中"年龄"和"教育年限"对"Ln_工资"有显著影响(均 p<0.05), "工龄"和"现雇佣年"对"Ln_工资"无影响(均 p>0.05)。

调整后的 R2 =0.37, 模型 2 可解释"Ln_工资"的 37%变异来源, 模型表达式为 Ln_工资=3.661+0.050×年龄+0.067×教育年限+0.004×工龄+0.015×现雇佣年

	非标准化系数		标准化系数	+	5		
坝	В	标准误	Beta	L L	ρ	VIE	
常数	3.661	0.10	—	36.80	0.00**	—	
年龄	0.050	0.01	0.35	8.32	0.00**	2.10	
教育年限	0.067	0.01	0.35	9.17	0.00**	1.76	
工龄	0.004	0.01	0.02	0.53	0.60	1.58	
现雇佣年	0.015	0.01	0.06	1.87	0.06	1.20	
R^2			0.37				
调整后的 R ²	0.37						
F	F(4,752)=111.46,p=0.00						
D-W 值			1.76				

表 5-14 模型 2 线性回归分析汇总表(n=757)

注: 1 . 因变量: Ln_工资。

2 . ** $p < 0.01_{\circ}$

因变量对数变换后,关于自变量回归系数的解释分析和模型1不同。例如,本例的"教育年限"偏回归线性系数为0.067,每增加一年教育,将使工资增加6.7%。同理,年龄每增加一年,工资会增加5%。

5.2.5 逐步线性回归的实例分析

当要研究分析的自变量个数较少,且这些自变量得到专业及文献资料的支持时,自变 量一般未提前筛选,采用全部纳入模型的方式,即将专业上欲考察的自变量全部纳入线性 回归模型中。 在实际分析过程中,以探索为目的的相关影响关系研究,往往是从众多的自变量中搜 寻对因变量有影响的因素,研究者并不清楚哪些自变量与因变量是相关的,此时为提高分 析效率,回归过程只引入作用显著的自变量而剔除无意义的自变量,可采用逐步回归的形 式进行多重线性回归分析。

1.逐步归法

逐步回归包括 3 种自变量筛选方式,其方法如表 5-15 所示。

表 5-15 逐步回归的 3 种方法

逐步回归的方法	说明
向前法	自变量逐个引入回归方程,只进不出
向后法	自变量先全部引入方程,然后将无意义的自变量逐个剔除
逐步法	向前法与向后法的综合,可进可出

(1)向前法:先对每个自变量做线性回归,然后按重要性逐个引入有显著性的自变量 建立多重线性回归方程,不对已引入的自变量做显著性检验,只进不出,直到没有自变量 被引入为止。

(2)向后法:先将所有自变量引入建立多重线性回归方程,然后按重要性逐个剔除无显著性意义的自变量,每剔除一次要针对剩余的自变量重新建立回归方程,直到回归方程 中的自变量不能被剔除为止。

(3)逐步法:先按向前法引入自变量,每引入一个自变量当前模型的所有自变量都要做一次显著性检验,已进入模型的自变量如果变得无显著性则剔除,能确保每次引入新的自变量前回归方程中都是有显著意义的自变量,属于双向筛选过程,直到没有自变量可以被引入,也没有自变量可以被剔除为止。

逐步法可克服向前法与向后法的一些缺陷,得到普遍使用。3种逐步回归方法的选择, 并无严格的规定或标准,实践中必须同时重视专业上对自变量的筛选和评价,不能完全依赖统计方法进行自变量筛选。

2. 逐步 归实例分析

【例 5-6】在【例 5-5】的基础上,研究者考虑加入更多的自变量以研究它们对"工资" 的影响。拟建立以"是否住美国南方""婚否""是否住大城市""母亲受教育年限""智商""世界 观""年龄""教育年限""工龄""现雇佣年"为自变量,"Ln_工资"为因变量,采用逐步回归法构建 多重线性回归模型。

1) 准备数据

因变量"Ln_工资"为定量数据, 自变量"年龄""教育年限""现雇佣年""工龄""智商""世界 观""母亲受教育年限"为定量数据, "婚否"、"是否住美国南方"及"是否住大城市"为二水平分 类变量, 如果进行哑变量处理后仍为自身, 则二水平分类变量也可以直接进行线性回归, 数据类型符合线性回归的基本要求。

2) 线性相关条件判断

针对待分析的 10 个自变量,分别绘制其与因变量的散点图或矩阵式散点图。经观察 发现,其与因变量之间基本呈线性相关关系,没有发现明显的非线性关系,可以进行下一 步的多重线性回归。具体操作及散点图解释与 5.2.4 节内容类似,此处略。

3) 建立线性回归模型

当研究者对自变量的筛选无明确专业或理论依据,在自变量较多时可考虑采用逐步回 归的方式进行筛选。在仪表盘中依次单击【进阶方法】→【逐步回归】模块,操作界面如 图 5-20 所示。将变量【Ln_工资】拖曳至【Y(定量)】分析框中,其他变量拖曳至【X(定量/ 定类)】分析框中,在下拉列表的向前法、向后法、逐步法中,选择【逐步法】,勾选保存 残差和预测值】复选框,最后单击【开始分析】按钮。

∅ 进阶	方法 🗸	开始分析	Ē	保存残差和预测的	∎ 🔽 O
聚类	因子	Ln_工资		方法	···· Y(定量)
主成分	分层回归	日不住美国主主		<u>逐步法</u> 向前法	
逐步回归	二元Logit	定台'吐天国南 刀 紙本		向后法	J
多分类Logit	有序Logit	星否住大城市			
事后多重比较	偏相关	母亲受教育年限			
典型相关	双因素方差	智商			
三因素方差	多因素方差	世界观			
协方差	判别分析	年龄			
岭回归	分层聚类	教育年限			X(定量/定类)
曲线回归	PLS回归	工龄			
Lasso回归	RFM模型	现雇佣年			

图 5-20 例 5-6 逐步回归操作界面

4) 模型检验与评价

逐步回归分析结果汇总(n=757)如表 5-16 所示。

(1) 回归方程总体显著性检验, F(6,750)=93.43, p<0.01, 按 α=0.01 水平, 认为本 次拟合所得的回归方程具有统计学意义。

(2) 偏回归系数检验,待分析的 10 个自变量,最终保留在模型的包括"是否住美国南方""婚否""是否住大城市""智商""年龄""教育年限"6 个自变量, t 检验 p 值全部小于 0.05,说明这 6 个自变量对"Ln_工资"的影响有统计学意义。由标准化系数看出,"年龄""教育年限""是否住大城市"对"Ln_工资"的影响位列前三位。

项	非标准	化系数	标准化系数	+	n			
	В	标准误	Beta	L	ρ	VIE		
常数	3.450	0.12	—	28.72	0.00**	_		
是否住美国南方	-0.090	0.03	-0.09	-3.35	0.00**	1.04		
婚否	0.100	0.03	0.11	3.64	0.00**	1.28		
是否住大城市	0.130	0.03	0.14	5.07	0.00**	1.04		
智商	0.004	0.00	0.13	4.08	0.00**	1.39		
年龄	0.050	0.00	0.34	9.78	0.00**	1.58		
教育年限	0.050	0.01	0.25	6.88	0.00**	1.67		
R^2		0.43						
调整后的 R ²	0.42							
F	F(6,750)=93.43,p=0.00							
D-W 值			1.81					

表 5-16 逐步回归分析结果汇总(n=757)

注: 1.因变量: Ln_工资。

2 . ** $p{<}0.01_{\circ}$

(3) 经过逐步回归的筛选,最终回归方程为

Ln_工资=3.450-0.090×是否住美国南方+0.100×婚否+0.130×是否住大城市+0.004×智 商+0.050×年龄+0.050×教育年限。

(4) 模型拟合评价,回归方程调整后 R2 =0.42,表示"Ln_工资"变异的 42%能被上述 多重线性回归方程所解释。

5) 模型残差及共线性诊断

(1) 残差诊断,首先通过勾选【保存残差和预测值】复选框,平台会自动计算本次回 归方程的残差和预测值,两个新变量名称跟前缀字符简称为"Residual"和"Prediction",然后 通过单击【可视化】下的【散点图】【直方图】或【P-P 图/Q-Q 图】模块绘制残差散点图、 残差直方图、残差正态 PP 图,并进行残差的各条件诊断。经诊断,残差符合正态分布、 等方差性。具体操作解释及分析与 5.2.4 节相同,此处略。残差独立性,由表 5-16 中的 D-W 值可知,本例回归方程的 D-W=1.81,接近 2,可粗略认为本例残差独立。

(2) 共线性诊断,由表 5-16 中的 VIF 指标可知,纳入回归方程的 6 个自变量,其各 自的 VIF 指标均小于 5,认为不存在多重线性问题。

6) 结果分析报告

综上所述,通过本次逐步回归发现,"是否住美国南方""婚否""是否住大城市""智商""年龄""教育年限"这 6 个自变量对工资有显著影响,可解释工资变异的 42%,整个回归方程有统计学意义。

5.2.6 有哑变量的线性回归

前面介绍了线性回归的因变量必须是定量数据,自变量允许是定量数据或定类数据。 在定类数据中,如果是二水平分类变量一般视为定量数据,可直接进行线性回归。而对于 多个分类水平的分类自变量,应酌情考虑将其转换为哑变量,以哑变量的形式参与线性回 归分析。

1.哑变量转换

例如, A、B、AB、O 4 种血型数据, 依次用数字 1、2、3、4 表示, 如果直接以血型 作为自变量, 则回归系数表示血型每增加/减少一个单位, 因变量随之增加/减少的改变量, 这与实际情况不符, 因为 4 种血型是平等关系, 并不存在递增或递减效应。遇到此类自变 量的线性回归分析, 应考虑将分类变量转换为数个哑变量, 每个哑变量只代表与参考水平 相比的差异, 这样做所得的回归系数才有实际意义。

哑变量又称虚拟变量,它是人为虚设的变量,所以有些地方也称之为虚设变量。哑变 量较常见的表示方式是"指示符法",即用 0-1 数据进行组织。

一个有 k 个水平的多分类变量转换为哑变量时, 可生成 k 个哑变量, 每个哑变量均为 0-1 数据, 1 表示原分类水平的一个分类, 0 表示非此类。

哑变量转换示例如表 5-17 所示。哑变量"血型_1"的编码 1 对应的是"A 型", 编码 0 则 表示"非 A 型", 该哑变量代表的就是"A 型"血型; 哑变量"血型_2"对应的是"B 型"; 哑变量 "血型_3"对应的是"AB 型"; 哑变量"血型_4"对应的是"O 型"。同一个分类变量转换所得的多 个哑变量, 一般简称为"一组"或"一簇"哑变量。

	血型及水	k平编码	血型_1	血型_2	血型_3	血型_4	
_	A 型	1	1	0	0	0	
	B 型	2	0	1	0	0	
	AB 型	3	0	0	1	0	
_	O 型	4	0	0	0	1	

表 5-17 哑变量转换示例

在 SPSSAU 平台中,可通过在仪表盘中依次单击【数据处理】→【生成变量】模块对 多分类的自变量进行哑变量处理,相关介绍见本书 3.4 节内容。

2.参照水平

多分类变量转换为哑变量参与线性回归时,应选择一个恰当的分类作为参照水平,即 哑变量回归时,纳入回归模型的哑变量为 k-1 个,减掉的这一个作为参照。例如,我们可 选择"O 型"作为参照,此时参与回归的仅包括"血型_1""血型_2""血型_3"这 3 个哑变量,而 哑变量"血型_4"作为参照不纳入回归模型。

哑变量回归时,应注意遵守"同进同出"原则,即若任意一个哑变量对因变量 Y 有显著 性,则其同组哑变量均一并纳入回归模型;若一组哑变量对因变量 Y 无显著性,则该组哑 变量全部被剔除模型。

值得注意的是,参照水平的选择不是随意的,主要根据专业和研究目的进行选择。怎 么理解呢?例如,研究目的在于考察"吸烟"对患某疾病的影响,则以"不吸烟"作为参照;再 例如,研究病情严重程度对预后质量的影响,根据专业知识将病情严重程度划分为4个等级,可考虑将等级最低的水平作为参考,有利于临床意义的解释。

3.实例分析

【例 5-7】对【例 5-1】的案例背景和数据进行重新整理,数据文档为例"例 5-7.xls"。 某研究收集到 758 名美国年轻男子的数据,行业经验认为"年龄"、"教育年限"及"智商等级" 对"Ln_工资"的对数数据有预测作用,试拟合多重线性回归进行分析。

1) 哑变量转换

本例"智商等级"为有 4 个分类水平的分类变量, 4 个等级数字编码依次为 1、2、3、 4。 其作为线性回归自变量时, 考虑对其进行哑变量转换生成 3 个哑变量, 以 3 个哑变量形式 参与线性回归。

在仪表盘中依次单击【数据处理】→【生成变量】模块,先在左侧的变量列表中选中 【智商等级】变量,然后在右侧下拉列表中选择常用的【虚拟(哑)变量】,最后单击底部的 【确认处理】按钮。哑变量转换操作与结果如图 5-21 所示。



图 5-21 哑变量转换操作与结果

转换后,在原始数据中新增"智商等级_1~智商等级_4"4 个哑变量,分别对应的是"智 商等级"的 4 个水平。此处应注意,选择其中一个水平作为参照,本例选择"智商等级_1"即 第一个水平作为参照,其余 3 个哑变量参与接下来的线性回归。

2) 线性回归

同一组哑变量应同步进入模型或同步退出模型,为此哑变量不适合按照逐步回归方式 进行筛选,本例选择使用【线性回归】完成线性回归分析。在仪表盘中依次单击【通用方 法】→【线性回归】模块,将变量【Ln_工资】拖曳至【Y(定量)】分析框中,将【年龄】 与【教育年限】拖曳至【X(定量/定类)】分析框中。哑变量回归操作界面如图 5-22 所示。

□ 通用	II 通用方法 V		开始分析	Ō	保存残差和预测值 🗌 异常点信息保存 🗌	
频数	分类汇总		Ln_工资		∀(造量)	
描述	交叉(卡方)					
相关	线性回归		年龄 教育年限			
方差	t检验				a contra per contra Mila	
			習商寺级_4		X(定量/定奕)	
里样本t检验	配对t检验	θ	智商等级_3			
正态性检验	非参数检验		智商等级_2			

图 5-22 哑变量回归操作界面

本例以"智商等级_1"为参照,先将【智商等级_2】~【智商等级_4】这 3 个哑变量拖 曳至【X(定量/定类)】分析框中,然后勾选【保存残差和预测值】复选框,最后单击【开始 分析】按钮。

3) 结果分析

线性回归分析结果汇总(n=757)如表 5-18 所示。

(1)回归方程总体显著性检验, F(5,751)=95.35, p<0.01, 按 α=0.01 水平, 认为本 次拟合所得的回归方程具有统计学意义。

(2) 偏回归系数检验, "年龄"和"教育年限"两个自变量, 以及"智商等级"的 3 个哑变 量, t 检验 p 值全部小于 0.01, 说明这 5 个自变量对"Ln_工资"的影响有统计学意义。相对 于"智商等级_1"来说, 由"智商等级_1"变换到"智商等级_2", 工资水平增加了 10%; 由"智商 等级_1"变换到"智商等级_3", 工资水平增加了 11%; 由"智商等级_1"变换到"智商等级_4", 工资水平增加了 20%, 可见智商对工资的影响。

TE	非标准化系数		标准化系数	4	p	VIF		
坝	В	标准误	佳误 Beta					
常数	3.707	0.10	—	36.51	0.00**	_		
年龄	0.057	0.00	0.40	12.49	0.00**	1.26		
教育年限	0.046	0.01	0.24	6.65	0.00**	1.64		
智商等级_4	0.196	0.04	0.20	4.85	0.00**	2.08		
智商等级_3	0.112	0.04	0.12	3.08	0.00**	1.79		
智商等级_2	0.102	0.04	0.10	2.82	0.00**	1.61		
R^2		0.39						
调整后的 R ²	0.38							
F	F(5,751)=95.35,p=0.00							
D-W 值	1.77							

表 5-18 线性回归分析结果汇总(n=757)

注: 1.因变量: Ln_工资。

2 . ** $p{<}0.01_{\circ}$

(3) 最终回归方程为: Ln_工资=3.707+0.057×年龄+0.046×教育年限+0.196×智商等

级_4+0.112×智商等级_3+0.102×智商等级_2。

(4) 模型拟合评价,回归方程调整后 R2 =0.38,表示"Ln_工资"变异的 38%能被上述 多重线性回归方程所解释。

5.3 Logistic 回归

线性回归的因变量是定量数据,如果遇到分类变量,则不适合拟合线性回归。例如, 贷款违约的相关性研究,研究目标只有两种结局,数字1表示违约,数字0表示未违约, 如果用线性回归方程解决该问题,方程等号左侧只有两种结局0或1,而方程等号右侧的 值域是任意解,显然此类问题不能用线性回归方程来解决。

Logistic 回归有多种类型,包括二元 Logistic 回归、多分类 Logistic 回归、有序 Logistic 回归,以及条件 Logistic 回归,在 SPSSAU 平台中对应有【二元 Logit】、【有序 Logit】、【多分类 Logit】及【条件 Logit 回归】模块。本节主要介绍这 4 种类型的 Logistic 回归原理及应用。

5.3.1 方法概述

Logistic 回归是一种广义的线性回归分析模型,也是研究分类型因变量与某些影响因素 之间关系的一种回归分析方法。

1.Logistic 归的类型

Logistic 回归的类型如图 5-23 所示。根据数据资料的情况, Logistic 回归可分为成组 资料的非条件 Logistic 回归与配伍资料的条件 Logistic 回归。其中, 非条件 Logistic 回归根 据因变量的分类水平个数,可分为二元 Logistic 回归、多分类 Logistic 回归和有序 Logistic 回归。

(1) 二元 Logistic 回归:也称二项 Logistic 回归、二分类 Logistic 回归,因变量只有 两种结局,且结局是互斥的,如死亡与未死亡、癌症淋巴结转移与未转移。



图 5-23 Logistic 回归的类型

(2) 多分类 Logistic 回归:也称多项 Logistic 回归、多元 Logistic 回归,因变量是无序多分类变量,如某研究想了解不同社区与性别之间成年居民获取健康知识的途径是否不同,获取健康知识的途径包括 3 种,分别是传统大众媒介=1,网络=2,社区宣传=3,该因变量即为无序多分类变量,该问题适合采用多分类 Logistic 回归进行分析。

(3) 有序 Logistic 回归: 因变量为有序分类变量(等级数据), 如医学研究中关于某病的治疗效果, 无效=1, 有效=2, 痊愈=3, 如果要研究疗效的影响因素, 则采用有序
Logistic 回归。

(4) 条件 Logistic 回归:又称配对 Logistic 回归,其主要用于配对资料或分层资料的 多因素分析,包括1:1和1:M 配对资料的研究及分析。

2.Logistic 归的适用条件

Logistic 回归的因变量必须是二分类变量、无序多分类变量、有序分类变量, 自变量可 以是定量数据也可以是定类数据, 多水平的分类自变量应先转换为哑变量, 主要包括以下 适用条件。

(1) 定量数据的自变量与因变量的 Logistic 转换值之间存在线性关系,这是由 Logistic 回归的原理决定的,一般情况下该条件是满足的。

(2) 自变量之间无多重共线性,和线性回归类似,在考察多个自变量的影响时,如果存在共线性问题会影响 Logistic 回归的拟合结果。

(3) 样本量的经验要求:因变量结局中较少的那一类样本量是自变量个数的至少 10 倍,以上经验方法估计的只是样本量的温饱水平,相对而言样本量越大越好。例如,结果 为阳性与阴性,普遍来说阳性人群比例较低,100个阳性样本最多只能支持 10 个自变量, 或者说,研究者需要考察 8 个自变量要求阳性样本至少有 80 例。孙振球和徐勇勇(2014) 指出,对于配对资料,样本的匹配组数应为纳入方程中的自变量个数 p 的 20 倍以上。

3.Logistic 归的一般步骤

线性回归一般采用最小二乘法进行参数估计,而 Logistic 回归采用的是最大似然估计法,虽然原理上有所不同,但是整体的分析思路是类似的。非条件 Logistic 回归分析一般步骤如图 5-24 所示,适合二元 Logistic 回归与多分类 Logistic 回归。



图 5-24 非条件 Logistic 回归分析一般步骤

1) 基本条件判断

首先检查因变量是否为二分类变量、多分类变量、有序分类变量中的一种,若因变量 为其中一种,则采用对应的二元 Logistic 回归、多分类 Logistic 回归、有序 Logistic 回归; 若因变量为定量数据则采用线性回归。

建议在回归开始前先检查异常值情况,以及通过【线性回归】模块以结果变量为因变 量,输出各自变量的 VIF 来判断有无多重共线性影响,如发现共线性、异常值等问题,考 虑进行有针对性的处理措施,常见的有剔除或替换个别异常值数据,然后采取逐步回归等 操作。

2) 建立 Logistic 回归模型

建立 Logistic 回归模型的过程,较常见的是"先单后多",即先通过单因素分析筛选自变量,然后仅保留有显著影响的自变量进行多因素回归。这种场景在探索性研究目的、自变量较多或样本量不足的情况下应用较多。许汝福(2016)研究指出,当危险因素较多时可采用单因素分析进行初筛,但应注意适当调整检验水准并结合专业选择及纳入多因素分析变量,不要随意舍弃单因素分析无统计意义的自变量,避免漏掉重要的危险因素。

单因素分析的常见方法有卡方检验、t 检验、方差分析和秩和检验, 差异的显著性水 平可以由 0.05 适当放宽至 0.1、0.15, 甚至 0.2。有一点必须明确, 在进行多因素 Logistic 回归前进行单因素分析并不是绝对的, 在样本量充足、研究目标明确、有足够专业理论支 持的情况下,可将所有自变量一起进行多因素 Logistic 回归。

和线性回归类似,在进行多因素 Logistic 回归时,也可采用逐步回归法对变量进行筛选,如向前法、向后法或逐步法,尤其是逐步法 Logistic 回归在科研中使用较多。

3) Logistic 回归模型的检验与评价

Logistic 回归模型和线性回归模型检验类似,要先对模型总体显著性进行检验。具体判断时,可以直接读取似然比卡方检验的概率 p,如果 p<0.05 则认为模型总体有统计学意义;反之,如果 p>0.05 则认为模型无效。

Logistic 回归常用 Hosmer-Lemeshow 检验(简称 HL 检验)进行拟合优度评价,适用 于含有定量自变量的模型拟合优度评价。原假设模型拟合值和观测值的吻合程度一致,如 果 p>0.05 则说明通过 HL 检验,可认为模型拟合优度良好;反之,如果 p<0.05 则说明模 型没有通过 HL 检验,模型拟合优度一般或较差。

决定系数 R2 作为线性回归模型拟合优度的重要指标,其结果得到重视和应用。 Logistic 回归也提供 R2,常见的包括 McFadden R2、Cox & Snell R2、Nagelkerke R2, 它们被称为伪 R2,其含义和线性回归的决定系数 R2 类似,但是在实际应用中,它们在 Logistic 回归中较少使用。

此外,模型预测准确率也可用作模型拟合优度的评价,没有严格的标准,具体由专业 经验决定。

4) 偏回归系数与 OR 值解释与分析

对 Logistic 回归方程中各自变量的偏回归系数、OR 值及其 CI 进行解读,分析哪些因素对研究结局有影响,哪些因素无影响,以及通过 OR 值对影响程度进行分析。

5) 结论报告

综合回归模型显著性检验、模型拟合评价,以及偏回归系数和 OR 值的情况,总结并 呈现最终的分析结果和结论。

5.3.2 二元 Logistic 回归

二元 Logistic 回归中因变量只有两种结局,且两种结局是互斥的。举个例子进行说明,现在假设 y=1 为死亡, y=0 为未死亡, Logistic 回归最终可以做到的是将病例判定为死亡 或未死亡,以及出现该结局的概率。若死亡的概率为 P,则未死亡的概率为 1-P,令 ln(P/1- P)=logit(P),这一过程被称为 logit 对数变换。

1.二元 Logistic 归模型

当有多个因素时, Logistic 回归的一般形式为

$$\operatorname{logit}(P) = \ln\left[\frac{P}{1-P}\right] = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \dots + \beta_m \mathbf{x}_m$$

整个模型以最大似然法进行参数估计,以医学、流行病学为例,模型中有以下主要概 念。

(1) P/1-P: 称为比值或优势 (Odds), ln(P/1-P)=logit(P)称为优势的对数, 大量实践 证明 logit(P)与定量的自变量呈线性关系。

(2) OR (Odds Ratio) 值:又称比值比、优势比,主要指用病例组中的比值 P/1-P 除以对照组中的比值 P/1-P,是流行病学、医学研究中的一个常用指标。

(3) 偏回归系数 j (j=1,2,…,m): 表示在其他条件不变的情况下, 自变量每改变一 个单位时 Logit(P)的改变量。如果回归系数是正数, 则表示自变量与因变量正相关; 如果 回归系数是负数,则表示自变量与因变量负相关。

(4)回归系数与 OR 值的关系:回归系数主要解释、分析自变量的显著性及对因变量 影响的正负方向,OR 值用于衡量自变量对因变量的作用程度,OR 值等于回归系数的自然 对数值。若自变量 X 的偏回归系数为 0.6,则其 OR=exp(0.6)= e0.6 =1.822。回归系数、 OR 值及对因变量的意义如表 5-19 所示。

回归系数 OR 值		对因变量的意义	
$\beta_j < 0$	OR<1	该因素是保护或抑制因素	
$\beta_j = 0$	OR =1	该因素对结局的发生与否不起作用	
$\beta_j > 0$	OR>1	该因素是危险或促进因素	

表 5-19 回归系数、OR 值及对因变量的意义

2.逐步法筛选自变量

和线性回归类似,多因素 Logistic 回归也可采用逐步回归方法对变量进行筛选,如向前法、向后法或逐步法,尤其逐步法在多因素 Logistic 回归中受到科研工作的青睐。此处 注意,SPSSAU 平台会采用 Wald 检验进行对自变量的逐步筛选。

3.二元 Logistic 归实例分析

【例 5-8】研究者收集了银行贷款客户的个人负债信息,以及曾经是否有过还贷违约 记录,数据赋值说明如表 5-20 所示,试分析是否违约的相关因素。数据来源于 SPSS 统计 软件自带数据集"bankloan.sav",数据文档为"例 5-8.xls"。

因素	变量说明	赋值说明
年龄	岁、分类数据(三水平)	<30=1, 30~40=2, >40=3
教育水平	年、分类数据(三水平)	高中及以下等于 1, 高中等于 2, 大学及以上等于 3
当前雇佣时长	年、分类数据(三水平)	小于或等于 3 等于 1, 4~10=2, 大于 10 等于 3
当前居住时长	年、分类数据(三水平)	小于或等于 3 等于 1, 4~10=2, 大于 10 等于 3
家庭收入	千美元、定量数据	
负债收入比率	%、定量数据	
信用卡负债	千美元、定量数据	
其他负债	千美元、定量数据	
曾经违约	分类数据(二水平)	否=0, 是=1

表 5-20 例 5-8 案例数据赋值说明

1) 基本条件判断

研究贷款违约发生的相关因素,因变量"曾经违约"有两种结局,因此选择使用二元 Logistic 回归。通过在仪表盘中依次单击【通用方法】→【频率】模块,针对"曾经违约"进 行频率统计,曾经发生过违约行为的有 183 例,按样本量与自变量个数的 10 倍关系计算, 样本量基本满足本次分析需要。

通过在仪表盘中依次单击【通用方法】→【线性回归】模块,以"曾经违约"为因变量, 其他数据为自变量进行线性回归,在输出结果中发现,各自变量的 VIF 值均未超过 5,初 步认为自变量间不存在共线性问题。

2) 建立 Logistic 回归模型

本例要考察的自变量有 8 个, 在建立多因素回归模型之前, 可先通过卡方检验、t 检验分析、了解各自变量对因变量的影响。

本例通过在仪表盘中依次单击【实验/医学研究】→【卡方检验】模块,分析"年龄"、 "教育水平"、"当前雇佣时长"、"当前居住时长"与"曾经违约"间的关系。单因素分析1 如表 5-21 所示。

田麦	夕む	曾经证	违约	百十	·2	n
四条	口小小	否	是	105 11	χ	ρ
	30 岁以下	136(26.31%)	82(44.81%)	218(31.14%)		
年龄	30~40 岁	222(42.94%)	59(32.24%)	281(40.14%)	21.58	0.00**
	40 岁以上	159(30.75%)	42(22.95%)	201(28.71%)		
	高中以下	293(56.67%)	79(43.17%)	372(53.14%)		
教育水平	高中	139(26.89%)	59(32.24%)	198(28.29%)	10.84	0.00**
	大学及以上	85(16.44%)	45(24.59%)	130(18.57%)		
	3年以下	104(20.12%)	93(50.82%)	197(28.14%)		
当前雇佣时长	4~10年	211(40.81%)	62(33.88%)	273(39.00%)	70.18	0.00**
	10年以上	202(39.07%)	28(15.30%)	230(32.86%)		
	3年以下	136(26.31%)	78(42.62%)	214(30.57%)		
当前居住时长	4~10年	210(40.62%)	60(32.79%)	270(38.57%)	17.07	0.00**
	10年以上	171(33.08%)	45(24.59%)	216(30.86%)		

表 5-21 单因素分析 1

注: ** p<0.01。

在仪表盘中依次单击【通用方法】→【t 检验】模块, 分析"家庭收入"、"负债收入比率"、"信用卡负债"、"其他负债"与"曾经违约"间的关系。单因素分析 2 如表 5-22 所示。

田表	曾经违约(平均	+	n	
四条	否(<i>n</i> =517)	是(<i>n</i> =183)	Ĺ	ρ
家庭收入	47.15±34.22	41.21±43.12	1.88	0.06
负债收入比率	8.68 ± 5.62	14.73±7.90	-9.54	0.00**
信用卡负债	1.25 ± 1.42	2.42 ± 3.23	-4.77	0.00**
其他负债	2.77 ± 2.81	3.86±4.26	-3.22	0.00**

表 5-22 单因素分析 2

注: ** p<0.01。

当进行单因素分析时,假设检验的显著性水平可适当放宽到 0.1 甚至 0.2。由表 5-21、 表 5-22 可知,经卡方检验与 t 检验,按 α=0.1 的显著性水平,所有自变量的 p 值均小于 0.1,说明待分析的 8 个自变量分别和因变量"曾经违约"有相关性。

对单因素阶段有显著性的各自变量继续进行多因素二元 Logistic 回归分析,因素较多时可采用逐步回归的方法进行筛选。SPSSAU 平台基于 Wald 检验提供了 3 种逐步回归的方法,分别是逐步法、向前法、向后法,逐步法是向前法与向后法的综合应用,一般情况下使用逐步法较多。

3) 建立多因素 Logistic 回归模型

本例变量"年龄""教育水平""当前雇佣时长""当前居住时长"为多分类变量,通过在仪表 盘中依次单击【数据处理】→【生成变量】模块,对这 4 个变量进行哑变量转换。在仪表 盘中依次单击【进阶方法】→【二元 Logit】模块,将【曾经违约】拖曳至【Y(定类)】分 析框中,特别注意,因变量的两个水平数字编码必须是 0 和 1,可提前通过【数据处理】

筛选样本 700 @ 😋 🧕		开始分析	Ē	保存残差和预测值 ☑ 逐步法(基于wald ŗ▼) ♀		
年龄_30岁以下		曾经违约		方法 全进入(默认)		
当前居住时长_3年以下		年龄 40岁以上		逐步法(基于wald p) 向前法(基于wald p)		
当前雇佣时长_3年以下		年龄_30-40岁		向后法(基于wald p)		
教育水平_高中以下		- 当前居住时长_10年以上				
年龄	e	当前居住时长_4-10年				
教育水平		当前雇佣时长_10年以上				
当前雇佣时长		当前雇佣时长_4-10年				_
当前居住时长		教育水平_大学及以上			112 / 18	7
		教育水平_高中				
		家庭收入				
		负债收入比率				
		信用卡负债				
		其他负债				

下的【数据编码】模块查看或进行编辑转换。将【家庭收入】【负债收入比率】【信用卡负 债】【其他负债】,以及【年龄】【教育水平】【当前雇佣时长】【当前居住时长】这4个变量 生成的哑变量全部拖曳至【X(定量/定类)】分析框中,注意本例全部选择第一个水平作为参 照,4 个分类变量的一水平哑变量不移入【X(定量/定类)】分析框中,勾选【保存残差和预 测值】复选框,最后单击【开始分析】按钮。二元 Logistic 回归操作界面如图 5-25 所示。

图 5-25 二元 Logistic 回归操作界面

Logistic 回归输出包括基本汇总、模型似然比卡方检验、分析结果汇总、回归预测准确 率、Hosmer-Lemeshow 拟合度检验、coefPlot 图等结果,我们可以按步骤进行解释和分析。

4) Logistic 回归模型的检验与评价

模型似然比卡方检验用于对整体模型的有效性进行分析。二元 Logistic 回归模型的总体显著性检验如表 5-23 所示。χ2=229.287, p<0.01, 认为二元 Logistic 回归模型总体上有统计学意义,模型中引入的自变量至少有一个对因变量有影响,模型是有效的。

表 5-23 二元 Logistic 回归模型的总体显著性检验

模型	-2 倍对数似然值	χ^2	df	р	AIC 值	BIC 值
仅截距	804.364					
最终模型	575.077	229.287	6	0.000	589.077	620.934

AIC 值、BIC 值这两个统计量用于模型间的比较,取值越低模型拟合越好,本例只有一 个有效模型,没有可比较的对象,因此 AIC 值和 BIC 值并无实际用处。"-2 倍对数似然值" 即其他统计软件工具输出的"-2LL"统计量,也用于多个模型间的比较,取值越低模型拟合 越好。

表 5-24 Hosmer-Lemeshow 拟合度检验结果

χ^2	df	р
5.219	8	0.734

Hosmer-Lemeshow 拟合度检验结果如表 5-24 所示, χ2=5.219, p=0.734>0.05, 说明 模型拟合良好。

二元 Logistic 回归预测准确率如表 5-25 所示。本例二元 Logistic 回归模型对结局 0 即 未违约的预测准确率为 93.04% (481/517),对结局 1 即违约的预测准确率为 45.90%,总体 预测准确率为 80.71%。从银行贷款业务风险预警角度来看,本例更关注对违约结局的预测 能力,显然 45.90%是比较低的,该模型的实用价值有待进一步提高。

预测值 项 预测准确率 预测错误率 0 1 481 93.04% 0 36 6.96% 真实值 1 99 84 45.90% 54.10% 汇总 80.71% 19.29%

表 5-25 二元 Logistic 回归预测准确率

注意,有些研究并不看中模型的预测能力,而主要关注的是因变量的相关影响因素。 因此,预测准确率表的结果应综合研究目的来解释及分析。

在模型分析结果汇总表(偏回归系数解释时使用),即表 5-26 的底部, SPSSAU 平台 提供了 3 个伪 R2 指标,其含义类似线性回归中的决定系数 R2,取值越大越好,在实际分 析中应用较少,可以不做关注。

5) 偏回归系数与 OR 值解释及分析

二元 Logistic 回归分析结果汇总如表 5-26 所示。通过逐步法,模型能自动根据显著性 情况对自变量进行引入或剔除。在 4 个定量数据中,Wald 卡方检验显示"负债收入比率" "信用卡负债"的 p 值均小于 0.05,它们对"曾经违约"的影响有统计学意义。

	偏回归 系数	标准误	z值	Wald χ^2	р	OR 值	OR 值 95% CI
当前居住时长_10年以上	-0.947	0.281	-3.375	11.390	0.001	0.388	0.224~0.672
当前居住时长_4~10年	-0.789	0.242	-3.264	10.652	0.001	0.454	0.283~0.730
当前雇佣时长_10年以上	-3.211	0.391	-8.223	67.611	0.000	0.040	0.019~0.087
当前雇佣时长_4~10年	-1.292	0.231	-5.594	31.293	0.000	0.275	0.175~0.432
负债收入比率	0.103	0.018	5.690	32.381	0.000	1.108	1.070~1.148
信用卡负债	0.426	0.078	5.447	29.666	0.000	1.530	1.313~1.784
截距	-1.099	0.257	-4.276	18.282	0.000	0.333	0.201~0.551

表 5-26 二元 Logistic 回归分析结果汇总

注: 1.因变量: 曾经违约。

2. McFadden R^2 : 0.285.

3. Cox & Snell R^2 : 0.279°

4. Nagelkerke R^2 : 0.409°

Wald 卡方检验显示"当前居住时长""当前雇佣时长"的 4 个哑变量, p 值均小于 0.05, 它们对"曾经违约"的影响也有统计学意义。其他没有引入模型的项, 如"家庭收入""其他负 债""教育水平"的两个哑变量及"年龄"的两个哑变量均没有统计学意义。

标准误和 z 值这两个为中间计算过程的统计量,标准误不宜过大,z 值一般不用解读。 表中重点是各因素的偏回归系数、OR 值及其 95% CI。

两个定量数据"负债收入比率""信用卡负债"的偏回归系数为正数,认为其与"是否违约" 存在正向相关关系。相对应的 OR 值大于 1, OR 值 95% CI 不包括 1,说明"负债收入比 率""信用卡负债"越高越容易出现偿还贷款违约的情况。以"信用卡负债"为例, Wald X2=29.666,p<0.01,认为其对"是否违约"的影响有统计学意义,二者存在正相关关系。 OR=1.530>1,说明其为发生违约的危险因素或促进因素,"信用卡负债"每增加一个单位, 其发生违约的可能性是原来的 1.530 倍,或发生违约的可能性比原来增加 53%。

4 个哑变量的偏回归系数均为负数,说明其与"曾经违约"存在负相关关系,相对应的 OR 值均小于 1, OR 值 95% CI 不包括 1,说明"当前居住时长"4~10 年、10 年以上,"当 前雇佣时长"4~10 年、10 年以上对"是否违约"起抑制作用,"当前居住时长""当前工作时长" 越长(稳定)越不容易出现还贷违约的情况。以"当前雇佣时长_10 年以上"为例,Wald X2=67.611,p<0.01,相较于"当前雇佣时长_4 年以下"认为其对"是否违约"的影响有统计学 意义,二者存在负相关关系。OR=0.040<1,说明其为发生违约的保护因素 或抑制因素, "当前雇佣时长"每改变一个等级,其发生违约的可能性是原来的 0.040 倍,或发生违约的 可能性比原来降低 99.6%。

6) 结果报告

本例建立的贷款违约二元 Logistic 回归模型为

In(P/1-P)=-1.099-0.947×当前居住时长_10年以上-0.789×当前居住时长_4~10年-3.211×当前雇佣时长_10年以上-1.292×当前雇佣时长_4~10年+0.103×负债收入比率+ 0.426×信用卡负债

其中, P 代表"曾经违约"为1的概率, 1-P 代表"曾经违约"为0的概率。总体而言模型 有统计学意义。"负债收入比率"和"信用卡负债"正向影响违约的发生, 而"当前居住时长"和 "当前雇佣时长"则反向抑制违约的发生。 coefPlot 图形可直观地展示模型中引入的自变量,以及各自变量对因变量影响的 OR 值情况。Logistic 回归的 OR 值结果绘制的 coefPlot 图形如图 5-26 所示。



图 5-26 Logistic 回归的 OR 值结果绘制的 coefPlot 图形

图中垂直的虚线代表 OR 值等于 1,为无效线,图中的横线段为各自变量的 OR 值 Cl, 线段中间的圆点为具体的 OR 值。若各自变量的 OR 值 Cl 和虚线无交叉或重叠,则表示对 应的自变量有显著性,位于虚线右侧表示 OR 值大于 1,为危险因素;位于虚线左侧表示 OR 值小于 1,为保护因素。

5.3.3 多分类 Logistic 回归

因变量为无序多分类变量,如研究成人早餐选择的相关因素,早餐种类包括谷物类、 燕麦类、复合类,此时因变量有3种结局,而且3种早餐是平等的没有顺序或等级属性, 此类回归问题可以使用多分类 Logistic 回归进行分析。

1.模型原理

多分类 Logistic 回归有时也被称为多元 Logistic 回归,从因变量的多个类别中选一个水平 作为对照拟合其他类别水平,相较于该对照水平的 Logistic 回归模型,因此 k 个分类水平 的因变量,最终会得到 k-1 个 Logistic 回归模型。

2.重要概念

(1) 多分类 Logistic 回归模型的参数估计与二元 Logistic 回归模型类似,同样采用最大似然法。

(2) 在模型检验方面,多分类 Logistic 回归模型与二元 Logistic 回归模型有一些差别, 常用的拟合优度检验为 Pearson 卡方检验和偏差似然比卡方检验。其他概念和二元 Logistic 回归模型基本类似。

3.多分类 Logistic 归实例分析

【例 5-9】以 1992 年美国总统选举的部分数据为例,总统投票对象包括 Perot、Bush、 Clinton,数据赋值说明如表 5-27 所示,试分析选民的投票情况。案例数据来源于卢纹岱 (2006),数据文档见"例 5-9.xls"。

因素	变量说明	赋值说明
age	年龄	定量数据

表 5-27 例 5-9 案例数据赋值说明

sex	性别	male=1, female=2
pres	投票对象	Perot=1, Bush=2, Clinton=3

1) 基本条件判断

本例研究投票对象的相关影响因素,投票对象(变量)为"pres",有 3 个分类水平, 为无序多分类变量,总投票数为 1847,通过在仪表盘中依次单击【通用方法】→【频数】 模块,"pres"的 3 个投票对象 Perot、Bush、Clinton 依次获得 278 票、661 票、908 票,样 本量能满足 Logistic 回归的经验要求,本例仅包括年龄、性别两个自变量,暂不考虑多重 共线性问题。

本例拟以 Perot 作为参照水平,采用多分类 Logistic 回归进行分析。

2) 建立 Logistic 回归模型

数据读入平台后,在仪表盘中依次单击【进阶方法】→【多分类 Logit】模块,将 【pres】变量拖曳至【Y(定类)】分析框中,将【sex】和【age】变量拖曳至【X(定量/定类)】 分析框中。此处应注意,常见的参照水平主要包括第一个类别或最后一个类别,平台默认 以第一个数字编码或较小的数据作为参照组。

∅ 进阶	方法 🗸	开始分析	Ē	保存预测类别 🖌
聚类	因子	pres		¥(定类)
主成分	分层回归			- 0
逐步回归	二元Logit	sex		
多分类Logit	有序Logit	age		w/하리 /하⊁\
事后多重比较	偏相关			X(正重/正尖)
典型相关	双因素方差			

图 5-27 多分类 Logistic 回归操作界面

多分类 Logistic 回归的自变量可以是定量数据,也可以是定类数据。如果是多分类定 类数据可根据实际情况提前做哑变量转换,如果未做哑变量转换,其移入【X(定量/定类)】 分析框后,平台会按定量数据进行回归分析。勾选【保存预测类别】复选框,命令平台对 案例数据进行类别预测,最后单击【开始分析】按钮。多分类 Logistic 回归操作界面如图 5-27 所示。

多分类 Logistic 回归输出包括基本汇总、模型似然比卡方检验、回归分析结果汇总、 预测准确率等结果。在结果解释及分析时,可参考二元 Logistic 回归,先判断模型总体是 否有效,再评价模型拟合质量,最后检验各自变量因素的显著性及分析 OR 值结果。

3) Logistic 回归模型的检验与评价

多分类 Logistic 回归模型和二元 Logistic 回归模型一样,总体显著性检验仍然采用的 是似然比卡方检验。多分类 Logistic 回归模型总体显著性检验如表 5-28 所示。经检验, x2=89.743, p<0.05,认为模型总体上有统计学意义,模型有效。表 5-28 中的 AIC 值、 BIC 值及-2 倍对数似然值,和二元 Logistic 回归解读一致,均为取值越小越好,主要用于 多个模型间的比较,此处可解释及分析的意义不大。

表 5-28 多分类 Logistic 回归模型总体显著性检验

模型	-2 倍对数似然值	χ^2	df	р	AIC 值	BIC 值
仅截距	3700.829					

最终模型	3611.086	89.743	4	0.000	3623.086	3656.214

我们也可以用预测准确率来评价模型的拟合优度,如表 5-29 所示。本次拟合的多分 类 Logistic 回归模型,对 Perot、Bush 的投票预测准确率都很低,对 Clinton 的投票预测准 确率可达 99.34%。

项		预测值			预测作确密	环测性记录	
		Perot	Bush	Clinton	"贝州/庄1用平	贝侧相厌平	
	Perot(n=278)	0	4	274	0.00%	100.000%	
真实值	Bush(<i>n</i> =661)	0	9	652	1.36%	98.638%	
	Clinton(n=908)	0	6	902	99.34%	0.661%	
		汇总			49.32%	50.68%	

表 5-29 多分类 Logistic 预测准确率

夜 5-50 多力关 LOUISIIC 巴归力协组者	表 5-30	多分类 Logistic 回归分	·析结果
----------------------------	--------	------------------	------

Bush	偏回归系数	标准误	z 值	Wald χ^2	р	OR 值	OR 值 95% CI
sex	0.301	0.145	2.072	4.292	0.038	1.351	1.016~1.796
age	0.031	0.005	6.239	38.921	0.000	1.032	1.022~1.042
截距	-0.992	0.319	-3.108	9.662	0.002	0.371	0.198~0.693
Clinton	偏回归系数	标准误	z值	Wald χ^2	р	OR 值	OR 值 95% CI
Clinton	偏回归系数 0.734	标准误 0.141	z值 5.214	Wald χ ² 27.191	р 0.000	OR 值 2.084	OR 值 95% CI 1.581~2.747
Clinton sex age	偏回归系数 0.734 0.034	标准误 0.141 0.005	z值 5.214 6.939	Wald χ ² 27.191 48.151	р 0.000 0.000	OR 值 2.084 1.035	OR 值 95% CI 1.581~2.747 1.025~1.045

注: 1. McFadden R²: 0.024。

2. Cox & Snell R^2 : 0.047.

3. Nagelkerke R^2 : 0.055°

4) 回归系数与 OR 值解释与分析

对 k 个分类水平的因变量进行多分类 Logistic 回归, 将得到 k-1 个模型, 每个模型都 能独立计算各自变量对因变量的回归结果。多分类 Logistic 回归分析结果如表 5-30 所示。 应注意该表格分为上下两部分, 前 4 行为与 Petor 相比, 投票给 Bush 的影响因素分析; 而 后 4 行为与 Petor 相比, 投票给 Clinton 的影响因素分析。

(1) 与 Petor 相比, 投票给 Bush 的影响因素分析。

经 Wald 卡方检验, 性别 sex (χ2=4.292, p<0.05)、年龄 age (χ2=38.921, p<0.01), 认为性别和年龄对投票给 Bush 的影响有统计学意义。这两个因素的偏回归系数均为正数, 说明与投票给 Bush 有正相关关系。相对应的 OR 值均大于 1, OR 值 95% CI 不包括 1, 说 明性别、年龄对投票结果有影响。

以性别为例, OR=1.351, 表示与 Petor 相比, 女性投票给 Bush 的可能性是男性的 1.351 倍 (默认以低编码水平为参照)。

(2) 与 Petor 相比, 投票给 Clinton 的影响因素分析。

经 Wald 卡方检验, 性别 sex (χ2=27.191, p<0.01)、年龄 age (χ2=48.151, p<0.01), 认为性别和年龄对投票给 Clinton 的影响有统计学意义。这两个因素的偏回归系数均为正数, 说明与投票给 Clinton 有正相关关系。相对应的 OR 值均大于 1, OR 值 95% Cl 不包括 1, 说明性别、年龄对投票结果有影响。

性别的 OR=2.084,为促进因素,表示与 Petor 相比,女性投票给 Clinton 的可能性是 男性的 2.084 倍。年龄的 OR=1.035,同样也属于促进因素,表示与 Petor 相比,年龄越大 的群体越愿意投票给 Clinton。

3) 结果报告

根据表中的常数项和偏回归系数,两个模型的表达式为 In(Bush/Petor)=-0.992+0.301sex+0.031age In(Clinton/Petor)=-1.486+0.734sex+0.034age 性别、年龄对投票结果的影响均有统计学意义,是候选人选取成功的显著影响因素。

5.3.4 有序 Logistic 回归

有序 Logistic 回归和多分类 Logistic 回归的因变量均有多个分类水平,但是前者分类 水平是有顺序、等级属性的。例如,临床试验的疗效分为无效、好转、有效和治愈 4 个等 级,社会调查类研究满意度分为 1~5 层级。

1.模型原理

由于因变量是等级资料的特征,因此有序 Logistic 回归模型被称为累加 Logit 模型,其 原理是,对因变量水平分割后形成多个二元 Logistic 回归模型,假设多个模型中的自变 量回归系数不变,不同的仅是模型的常数项,可通俗理解为模型的回归曲线是平行的。 有序 Logistic 回归分割原理示意图如图 5-28 所示。以 Y 有 3 水平为例,依次编码为 1、2、 3,按序依次产生两个分割点,拆分出两个二元 Logistic 回归模型。第一个模型为(1 vs 2+3),第二个模型为(1+2 vs 3),一般参照水平均取较高等级。也就是说,k个水平将得 到 k-1 个二元 Logistic 回归模型。这 k-1 个二元 Logistic 回归模型要求回归自变量不变, 仅常数项改变。



图 5-28 有序 Logistic 回归分割原理示意图

2.重要概念

1) 连接函数

连接函数可以理解为是累计概率的转换形式,用于累计概率模型的估计。有序 Logistic 回归通常包括 5 种连接函数,使用说明如表 5-31 所示。

连接函数	模型公式	使用说明
Logit	$f(x) = \log[x/(1-x)]$	因变量各选项分布较均匀,或者选项个数较少时使用
Probit	$f(x)=\Phi^{-1}(x)$	因变量接近正态分布时使用
补充log-log	$f(x)=\log[-\log(1-x)]$	因变量水平高的选项出现概率高,且选项个数较多时使用
负 log-log	$f(x) = -\log(-\log x)$	因变量水平低的选项出现概率高,且选项个数较多时使用
Cauchit	$f(x) = \tan[\pi(x-0.5)]$	因变量存在极端值时使用

表 5-31 有序 Logistic 回归 5 种连接函数使用说明

平台默认使用 Logit 连接函数,若模型没有特殊要求,一般建议使用 Logit 连接函数, 其常用在因变量分类水平较少时。连接函数可能会影响平行性检验,如果平行性检验无法 通过,则考虑根据因变量分布情况选择更为合适的连接函数。

2) 平行性检验

累加 Logit 模型需要先对因变量分类水平进行分割,然后对分割后的数据进行 Logistic

回归,此时要求分割后的模型参数满足平行性,即 Logistic 回归模型中的各自变量偏回归 系数相等。

该检验假设数据满足平行性,作为一个条件需要在有序 Logistic 回归时进行检验和判断。若 p>0.05 则说明数据满足平行性;反之,若 p<0.05 则说明数据不满足平行性。当平行性条件未被满足时,说明数据不适合进行有序 Logistic 回归,应该采用多分类 Logistic 回 归分析模型。

3.有序 Logistic 归分析的步骤

有序 Logistic 回归分析的步骤和前面的二元 Logistic 回归、多分类 Logistic 回归略有不同,如图 5-29 所示。



图 5-29 有序 Logistic 回归分析的步骤

1) 基本条件判断

因变量必须是等级数据,如药物治疗的疗效(无效、有效、显著有效),或者顾客满意 度(非常不满意、不满意、一般、满意、非常满意)。共线性、异常值等要求与二元 Logistic 回归一样。

2) 建立有序 Logistic 回归模型

一般选因变量的最大水平作为参考水平,以 Logit 连接函数建立有序 Logistic 回归模型, 也可以根据专业经验调整因变量的编码水平,以便对后面结果的解释及分析。

3) 平行性检验

模型分割后要求各自变量偏回归系数相等,并通过平行性检验。如果不满足平行性检验条件,则考虑调整连接函数重新进行检验。若最终结果认定无法满足平行性检验条件,则考虑将有序因变量视为无序多分类数据类型,使用多分类 Logistic 回归进行分析。

步骤 4) 5) 6) 的分析和二元 Logistic 回归、多分类 Logistic 回归基本一致,本节不做 赘述。

4.有序 Logistic 归实例分析

【例 5-10】研究者研究性别和两种治疗方法对某病疗效的影响,疗效的评价分为 3 个 等级,无效、有效和显效,数据赋值说明如表 5-32 所示。试分析疗效与性别、治疗方法 之间的关系。案例数据来源于张文彤(2002),数据文档见"例 5-10.xls"。

因素	变量说明	赋值说明
sex 性别	二水平	男=0, 女=1
treat 治疗方法	二水平	旧疗法=0,新疗法=1
effect 疗效	三水平	无效=1,有效=2,显效=3

表 5-32 例 5-10 数据赋值说明

1) 基础条件判断

治疗效果分为 3 个层级, 依次是无效=1, 有效=2, 显效=3, 为有序的多分类变量, 主要考虑性别和新旧两种治疗方法对疗效的影响, 应采用有序 Logistic 回归。

2) 建立有序 Logistic 回归模型

数据读入平台后, 在仪表盘中依次单击【进阶方法】→【有序 Logit】模块, 将 【effect】拖曳至【Y(定类)】分析框中, 将【sex】和【treat】拖曳至【X(定量/定类)】分析 框中。在【平行性检验】下拉列表中选择【进行检验】,在【连接函数】下拉列表中选择 【默认 Logit】。有序 logistic 回归操作界面如图 5-30 所示。

∅ 进阶	方法 🗸		开始分析	Ô	连接函数(默认L	ogit ∽ Ω
聚类	因子		effect		平行性检验 不进行(默认)	Y(定类)
主成分	分层回归		sex		进行检验	6
逐步回归	二元Logit		treat			
多分类Logit	有序Logit					
事后多重比较	偏相关	0				A(正重/ 定尖)
典型相关	双因素方差					

图 5-30 有序 logistic 回归操作界面

3) 平行性检验

作为有序 logistic 回归的适用条件,我们要先判断数据是否满足平行性检验条件。有序 Logistic 回归模型平行性检验如表 5-33 所示。经检验, x2=1.469, p=0.480>0.05,模型通 过平行性检验,使用有序 Logistic 回归进行分析是合适的。

表 5-33 有序 Logistic 回归模型平行性检验

模型	−2 倍对数似然值	χ^2	df	p
原假设	150.029			
最终	148.560	1.469	2	0.480

4) 模型总体显著性检验

有序 Logistic 回归模型总体显著性检验如表 5-34 所示。经检验, χ2=19.887, p<0.01, 认为模型总体上有统计学意义,模型有效;反之,如果 p>0.05 则认为模型无效。

表 5-34 有序 Logistic 回归模型总体显著性检验

模型	-2 倍对数似然值	χ^2	df	р	AIC 值	BIC 值
仅截距	169.916					
最终模型	150.029	19.887	2	0.000	158.029	167.753
<u> </u>		- 12- 1341				

表 5-34 中的 AIC 值、BIC 值及-2 倍对数似然值,和二元 Logistic 回归解读一致,均为 取值越小越好,主要用于多个模型间的比较,此处意义不大。

5) 模型回归系数、OR 值

本例因变量有 3 个等级,将得到两个模型,这两个模型中自变量的偏回归系数不变而 常数项不同。有序 Logistic 回归模型分析结果如表 5-35 所示。表中"因变量阈值"是指两个 模型的常数项估计结果,"自变量"是指两个模型的自变量偏回归系数的估计结果。

项 项 偏回归系数 标准误 z值 Wald χ^2 OR 值 OR 值 95% CI р 无效 1.813 0.557 3.257 10.607 0.001 0.163 $0.055 \sim 0.486$ 因变量阈值 0.021~0.225 19.781 0.000 0.069 有效 2.667 0.600 4.448 1.319 0.013 0.529 2.492 6.210 3.739 $1.325 \sim 10.548$ sex 自变量 1.797 14.449 0.000 0.473 3 801 6.033 treat $2.388 \sim 15.241$

表 5-35 有序 Logistic 回归模型分析结果

注: 1 . McFadden R^2 : 0.117。

2. Cox 和 Snell R²: 0.211。

3. Nagelkerke R^2 : 0.243.

两个模型的常数项依次为 1.813、2.667, 经 Wald 卡方检验, 性别"sex"的 χ 2=6.210, p=0.013<0.05, 治疗方法"treat"的 χ 2=14.449, p<0.01, 认为性别和治疗方法对疗效的作用都有统计学意义。性别"sex"和治疗方法"treat"的 OR 值依次为 3.739、6.033, 表明对女性患者来说疗效优于男性, 而新治疗方法的疗效优于旧治疗方法。

6) 结果报告

根据表 5-35 中的常数项和偏回归系数,两个模型的表达式为 logit[P(effect≤无效)/(1-P(effect≤无效))]=1.813+1.319sex+1.797treat logit[P(effect≤有效)/(1-P(effect≤有效))]=2.667+1.319sex+1.797treat

性别和治疗方法对疗效的作用有显著影响,对女性患者来说疗效优于男性,而新治疗 方法的疗效优于旧治疗方法的疗效。

第6章 问卷专属研究分析方法

对问卷数据资料进行研究分析是量化研究的重要内容,根据问卷题型的不同,所获得 的数据类型也不同,这影响到不同题型或不同数据类型采取何种统计方法的问题。问卷从 广义上可分为普通问卷和量表问卷,针对普通问卷和量表问卷的分析方法和应用也各有不 同。

问卷类型与分析方法如表 6-1 所示。本章从实用的角度介绍问卷数据分析方法的应用, 内容包括普通问卷常见题型的数据分析,如单选题、多选题的分析方法;量表问卷常见题 型的数据分析,如信度分析、效度分析、项目分析等。

表 6-1 问卷类型与分析方法

问卷类型	常用研究分析方法
並活门半	单选题、多选题等数据资料
百旭円仓	频数统计、交叉表卡方检验、t 检验、方差分析、线性回归、Logistic 回归等
	本书中特指李克特量表数据资料
量表问卷	信度分析、效度分析、项目分析、探索性因子分析、验证性因子分析、观察变量路径分析、
	潜变量结构方程模型、中介效应分析、调节效应分析、有调节的中介分析等

具体统计分析方法包括频数统计、交叉表卡方检验、项目分析、信度分析、探索性因 子分析、验证性因子分析、观察变量路径分析、潜变量结构方程模型,以及中介效应分析、 调节效应分析、有调节的中介分析等。

6.1单选题与多选题分析

在一份调查问卷中,常利用单选题、多选题进行基本事实现状的调查和分析。例如, 将被访者个人基本背景信息中的性别、年龄、学历、收入等设置为单选题,将研究对象的 现状、行为等根据实际情况设置为多选题。多选题"您通过以下哪些渠道获取新闻信息"的 选项包括电视新闻、今日头条、微博、微信、新闻客户端等。

SPSSAU 平台在【问卷研究】模块中提供了多个用于单选题、多选题的子模块,本节 仅介绍常用的【多选题】【单选-多选】两个子模块。

6.1.1 分析思路

单选题和多选题的分析,常见使用场景是单个单选题、单个多选题、单选题与单选题 交叉、单选题与多选题交叉,个别时候也会出现多选题与多选题交叉的情况。普通问卷中 单选题、多选题的分析思路如表 6-2 所示。

1.单个单选题/多选题

在数据录入时单选题或多选题一般可定义为二分类变量或多分类变量,属于定类数据 资料,单个单选题主要统计各选项(分类水平)的频数、百分比,必要时可通过柱形图/条 形图、饼图进行展示,如分析受访者的性别、年龄段、职业等的频数分布。

单个多选题的分析一般称为多重响应分析,主要涉及每个选项的响应频数、响应率、 普及率的描述统计,图形方面可使用帕累托图进行频数响应的展示和分析。根据研究分析 需求,也可以对多选题的各选项响应频数进行等比例的拟合优度检验。原假设各选项被选 中的响应率相等,如果卡方检验 p 值小于 0.05 则拒绝原假设,认为各选项的响应率差异显 著;反之,如果 p 值大于 0.05 则说明各选项的响应率无差异。

2.单选题与单选题交叉

两个单选题进行交叉分析时,首先要进行频数、百分比的统计,基于频数数据再使用 交叉表卡方检验进行差异比较、关联关系的研究。例如,分析不同职业(单选题)受访者 在是否吸烟(单选题)行为上的差异,或研究职业(单选题)与是否吸烟(单选题)间的 关联关系。

使用场景	分析思路	SPSSAU 路径
单个单选题	描述统计:频数、百分比描述统计;柱形图/条形 图、饼图	【通用方法】→【频数】
单个多选题	多重响应分析:各选项响应频数、响应率、普及 率;帕累托图	【问卷研究】→【多选题】
单选题与单	差异比较或关系研究, 交叉表卡方检验	【通田方注】→【交叉(卡方)】
选题交叉	至开记我这八示明儿: 文人私下方恒温	
单选题与多	差异比较或关系研究, 交叉表卡方检验	【问类研究】→【单选_名选】
选题交叉	至开记我这八示明儿: 文人私下方恒温	
单选题作为	单选题为因变量时使用 Logistic 回归分析影响因	【进阶方法】→【二元 Logit】【多分类
因变量	素或进行预测研究	Logit】【有序 Logit】
单选题作为 自变量	其他题型的数据作为因变量,单选题作为自变量 /分组变量进行 t 检验、方差分析、回归分析等,做 分组特征描述统计或影响因素分析	【通用方法】→【t 检验】【方差】【线 性回归】等

表 6-2 普通问卷中单选题、多选题的分析思路

3.单选题与多选题交叉

单选题与多选题交叉有两种情况,即单-多交叉(单选题作为自变量)、多-单交叉(单 选题作为因变量),前者在实际分析中较常见,为本节主要介绍的内容。单-多交叉分析, 实际上也是使用卡方检验进行差异比较或关联关系研究的。例如,性别(单选题)与新闻 获取渠道(多选题)进行交叉,通过卡方检验分析不同性别受访者在新闻获取渠道的分布 差异,或者说研究性别与新闻获取渠道的关系。

4.单选题作为因变量

在影响因素的调查研究中,有时候会以单选题形式收集因变量信息,进而采用 Logistic 回归分析影响因素。例如,某研究的因变量来自问卷中的题目"总体而言,您对自己所过生 活的感觉是怎样的呢?"选项分别为很不幸福、不太幸福、一般、比较幸福、非常幸福,为 有序多分类变量,影响因素分析时应采用有序 Logistic 回归。

5.单选题作为自变量

单选题常作为分组变量和其他题型的数据结合进行 t 检验、方差分析、回归分析等, 主要用于差异比较、控制变量、影响因素研究。例如, 在量表问卷中, 以性别、学历等单 选题数据作为分组变量, 以某维度或量表总得分数据作为因变量, 进行 t 检验、方差分析, 研究不同组的差异或是否对目标变量有影响作用。

此类情况下的 t 检验、方差分析、回归分析在 SPSSAU 平台的实现, 读者可阅读本书 对应的章节内容, 本节不做赘述。

6.分析思路总结

具体分析时应依据研究目的提出合理的分析需求,在一份调查问卷中,并不需要将所 有的单选题、多选题全部进行频数分析或将所有的单选题、多选题进行交叉分析,尤其是 两个题型的交叉,并不是所有的交叉分析都有实际意义。

6.1.2 频数统计实例分析

接下来通过实例进一步介绍单选题、多选题的分析方法。某调查研究在线英语学习网站课程购买意愿的影响因素,初步拟定产品、促销、渠道、价格、个性化服务和隐私保护这 6 个因素对网站用户购买意愿的影响,除核心量表题项外,研究者还设计了受访者"性别""年龄"等一般的背景信息题项,以及受访者"为什么学习英语""让你决定购买该课程的因素是什么"等基本特征或态度题项。试对该问卷中部分单选题、多选题的数据进行分析。案例数据来源于周俊(2017),问卷数据文档见"例 11-1.xls",以下简称英语问卷。

1.单选题实例

【例 6-1】对英语问卷中"您的性别""您的年龄"两个单选题进行频数、百分比的描述统 计。

1) 数据与案例分析

性别、年龄段的数据从类型上属于定类数据,可进行频数、百分比的描述统计。

2) 频数统计

针对单选题的描述统计,可通过在仪表盘中依次单击【通用方法】→【频数】模块来 实现。单选题频数统计操作界面如图 6-1 所示。从标题框中将【性别】【年龄】变量拖曳 至【分析项(定类)】分析框中,单击【开始分析】按钮即可。



图 6-1 单选题频数统计操作界面

3) 结果分析

输出的结果包括频数分析表及饼图、柱形图等,重点解释分析频数表的结果。问卷单 选题频数、百分比统计如表 6-3 所示。表中同时给出了性别、年龄的频数、百分比,以及 累积百分比结果。例如,年龄在 19~22 岁的受访者有 119 人,占样本总量的 39.67%。

名称	选项	频数	百分比	累积百分比
林見山	男	87	29.00%	29.00%
王加	女	213	71.00%	100.00%
	18 岁及以下	18	6.00%	6.00%
	19~22 岁	119	39.67%	45.67%
	23~26岁	42	14.00%	59.67%
石山山	27~30 岁	28	9.33%	69.00%
平时	31~34 岁	52	17.33%	86.33%
	35~38 岁	17	5.67%	92.00%
	39~42 岁	10	3.33%	95.33%
	43 岁以上	14	4.67%	100.00%
	合计	300	100.00%	100.00%

表 6-3 问卷单选题频数、百分比统计

2.多选题实例

【例 11-2】英语问卷中的题项"让你决定购买该课程的因素是什么"为多选题,以下简 124 / 187 称"购课因素"多选题,试对题项进行频数、百分比的描述统计。

1) 数据与案例分析

"购课因素"多选题的选项有"课程内容""师资力量""教学质量""课程价格""优惠折扣"及 "其它"共6项。每个选项录入为一个变量,一般建议采用"0-1"编码形式录入 多选题数据,即数字0表示未选中该选项,数字1表示选中该选项。本例为"0-1"编码形式, 6个选项共6个变量。

2) 频数统计

在仪表盘中依次单击【问卷研究】→【多选题】模块,该模块专门用于分析多选题的 响应频数和百分比。多选题频数统计操作界面如图 6-2 所示。从标题框中将【课程内容】、

【师资力量】、【教学质量】、【课程价格】、【优惠折扣】及【其它】6 个变量全部或逐一拖 曳至右侧的【1 个多选题对应选项】分析框中。如果多选题为"0-1"编码,则在上方的【计 数值】下拉列表中选择【计数值,默认 1】,即默认数字 1 表示选中该选项,多选题为"0-1" 编码的问卷默认设定即可。

■ 问考	新研究 🗸
信度	效度
多选题	单选-多选
多选-单选	多选-多选
项目分析	权重
验证性因子分析	对应分析
路径分析	结构方程模型 SEM
调节作用	中介作用
调节中介	KANO模型

图 6-2 多选题频数统计操作界面

3) 结果分析

与单选题的频数、百分比不同,多选题针对频数数据有两个新的概念,分别为响应率 和普及率。响应率是指选项被选中的次数占所有选项总选中次数的比例,普及率是指选中 某选项的受访者占总受访者人数的比例,前者是从选项的角度计算百分比的,后者是从受 访者的角度计算百分比的。二者的区别在于计算时分母不同,响应率的加和为 100%,普及 率的加和可能大于 100%。

本次分析的问卷多选题响应率、普及率、卡方拟合优度检验的结果如表 6-4 所示。 表 6-4 问卷多选题响应率、普及率、卡方拟合优度检验的结果

表 6-4 问卷多选题响应率、普及率、卡方拟合优度检验的结果

TE	响应		並乃亥 (<i>~</i> 300)	
坝	n	响应率	自及平(11-300)	
课程内容	211	26.34%	70.33%	
师资力量	142	17.73%	47.33%	
教学质量	213	26.59%	71.00%	
课程价格	146	18.23%	48.67%	
优惠折扣	74	9.24%	24.67%	
其它	15	1.87%	5.00%	
汇总	801	100.00%	267.00%	

注:卡方拟合优度检验: χ²=225.749, p=0.000。

(1)选项的响应率。本例的问卷共有 300 个受访者,这 300 个受访者在该多选题的 6 个选项中共选择了 801 个(次)选项,平均 1 个人选 2.67(801/300)个选项。"教学质量"被选中 213 次,即 213 次响应,占响应量的 26.59%(213/801)。同理"课程价格"被选中 146 次,响应率为 18.23%(146/801)。

使用响应率对各选项进行排序,即分析各选项在决定购买课程中的重要性。例如,本 例决定购买课程最重要的因素是"教学质量",第二因素是"课程内容",第三因素是"课程价 格"。此处也可以直接查看输出响应率的柱形图,以直观地观察和总结各选项的重要性。

(2)选项的普及率。"教学质量"选项被选中 213 次,意味着有 213 个受访者选择了该选项,这 213 个受访者在总受访者中占比为 71.00%(213/300),普及率呈现的是多选题具体选项在受访者中的可接受、受众普及程度,从"人"的角度反映了具体选项被重视的程度。同理可以按普及率对各选项进行排序,并通过对应的普及率条形图进行直观的观察和分析。

显然,我们发现本例中 6 个选项在响应率和普及率的前后排序是一致的,依次为教学 质量>课程内容>课程价格>师资力量>优惠折扣>其它,由此可见在决定购买课程的因素方 面,大家最为看重的前三项是"教学质量"、"课程内容"及"课程价格"。

(3)选项响应是否存在差异。本例卡方拟合优度检验的结果为 X²=225.749, p<0.05, 认为"教学质量""课程内容""课程价格""师资力量""优惠折扣""其它"这 6 个选项的响应率间具 有显著差异,或差异具有统计学意义,即决定受访者购买课程的因素被重视的程度是不一 样的,是有差异的。

(4) 帕累托图重要性分析。基于响应率制作的帕累托图,从"二八原则"解读该多选题 各选项的重要程度。例 11-2 多选题分析结果的帕累托图如图 6-3 所示,"教学质量""课程 内容""课程价格"3 个选项的累计比率为 71.2%,加上师资力量后累计比率达 88.9%,即对于 购买课程的因素,"教学质量""课程内容""课程价格""师资力量"4 个选项是"重要的",其他选 项是"次要的"。



图 6-3 例 11-2 多选题分析结果的帕累托图

6.1.3 卡方检验实例分析

单选题与单选题的交叉分析,可将两个单选题变量通过在仪表盘中依次单击【通用方法】→【交叉(卡方)】模块进行卡方检验,相关操作及卡方检验的原理、适用条件等知识 点介绍详见 4.3 节内容。 下面继续沿用 11.1.2 节"在线英语学习网站课程购买意愿影响因素研究"的调查问卷数据集, 重点介绍单选题与多选题数据的交叉分析。

【例 11-3】本研究尝试对"性别"单选题与"购课因素"多选题进行交叉分析。

1) 数据与案例分析

将"性别"单选题与"购课因素"多选题交叉分析转换为统计分析问题,即研究不同性别受 访者的购课因素有无差异,或者说分析"性别"与"购课因素"间的关联关系。从数据类型上两 个变量均为定类数据,可使用交叉表卡方检验进行定类数据间的关系研究。

2) 交叉表卡方检验

我们重点介绍卡方检验在问卷数据资料中的应用,在仪表盘中依次单击【问卷研究】 →【单选-多选】模块,该模块专门用于分析常见的单选题和多选题的交叉问题。

单选题与多选题交叉卡方检验操作界面如图 6-4 所示。从标题框中将【性别】变量拖 曳至【X(定类)【可选】】分析框中,将购课因素多选题的【课程内容】【师资力量】【教学 质量】【课程价格】【优惠折扣】【其它】6 个变量拖曳至【1 个多选题对应选项】分析框中, 在【计数值】下拉列表中选择【计数值,默认 1】,最后单击【开始分析】按钮。

🗉 问卷	研究 🗸		开始分析	Ē	计数值, 默认1	~ Q
信度	效度		性别			X(定类)[可选]
多选题	单选-多选		课程内容			
多选-单选	多选-多选		师资力量			
项目分析	权重		教学质量			
验证性因子分析	对应分析	Ð	课程价格			
路径分析	结构方程模型 SEM		优惠折扣			
调节作用	中介作用		其它			
调节中介	KANO模型				1	人文洪斯时南洪语
NPS	PSM				1	137268月四九四

图 6-4 单选题与多选题交叉卡方检验操作界面

注意,如果【X(定类)【可选】】不指定具体的单选题变量,则只针对多选题进行统计分析。

3) 结果分析

单选题与多选题交叉频数及卡方检验结果,如表 6-5 所示。

表 6-5 单选题与多选题父义频数及卞万检验约

15	,	汇出(1-200)		
坝	男(n=87) 女(n=213)		////////_	
课程内容	54 (62.07%)	157 (73.71%)	211 (70.33%)	
师资力量	44 (50.57%)	98 (46.01%)	142 (47.33%)	
教学质量	61 (70.11%)	152 (71.36%)	213 (71.00%)	
课程价格	40 (45.98%)	106 (49.77%)	146 (48.67%)	
优惠折扣	18 (20.69%)	56 (26.29%)	74 (24.67%)	
其它	4 (4.60%)	11 (5.16%)	15 (5.00%)	

注:卡方检验: χ²=1.762, p=0.881。

表 6-5 中的交叉频数及卡方检验结果(见表格底部注释)可用于分析不同组别受访者 在选项上的选择差异。本例 χ²=1.762, p=0.881>0.05, 认为男性组和女性组在购课因素的 选择上无差异,或差异无统计学意义。

127 / 187

6.2填空题分析

填空题作为一种开放式题项,常用在非量表类问卷中。它可以一个独立的填空题出现,如"请您对售后服务提出改进建议或意见";也可以某个单选题或多选题在"其它"选项中进行 文字补充,如"春节假期旅游度假您的理想目的地是"有 6 个固定选项,第 7 个选项为"其 它",并要求填写具体的目的地。

6.2.1 分析思路

填空题获取的数据一般为非标准化或非结构化的信息,如"希望提供洗车服务"文本信息,以及"请填写您的年龄"。受访者可能会给出五花八门的答案,如"58岁""我 60 了""35"。 类似这些数据难以进行统计计算,而在实际应用中问卷的填空题缺失情况比较严重,很多 为无效作答。

1.词云图

限于填空题的数据类型,填空题的分析一般从作答中挖掘有价值的关键点在问题讨论 时使用。针对填空题收集的文本数据,可使用文本分析方法,如文本分词后绘制的词云图, 直观地展示高频分词的分布情况,突出展示关键信息。

词云图也称云图,是由美国西北大学新闻学教授 Rich Gordon 提出的,"词云"可对文 字中出现频率较高的"关键词"予以视觉上的突出,形成"关键词云层"或"关键词渲染",从而 使浏览者通过快速观察文本就可以领略文本的主旨。当然也可使用 SPSSAU 的【文本分析 模块】进行深度的文本分析。

2.词云分析数据格式要求

SPSSAU 平台的【词云】模块,可针对填空题进行词云制作和展示。【词云】模块允许 两种数据格式,第一种是 txt 文本格式,一般建议一个受访者的作答文字作为一行,所有 受访者的文字信息另存为一个 txt 文本文档;第二种是词频数据格式,利用其他工具对所 有文字信息进行分词并统计词频,形成两列数据,第一列为关键词,第二列为频次。

3.txt 文本分析步骤

1) 创建文本文档

对填空题获得的文本数据,要先进行适度整理、清洗、规范,然后另存为 txt 文本文档。一般建议一个受访者的文本信息占一行,假设收集了 1000 份问卷,则将 1000 行文本信息另存为 txt 文本文档。

2) 上传文本文档

将保存的 txt 文本文档上传到 SPSSAU 平台的【词云】模块中。

3) 词云图及分析

【词云】模块将自动对文本进行分词,并统计各分词出现的频数从而形成词频数据, 据此绘制词云图,并结合研究主题对词云图进行解释,分析方法主要用于观察和分析高频 关键词。

4.词频数据分析步骤

1) 创建频数数据

首先,研究者可通过其他软件工具提前将受访者填空题的所有文字信息进行分词,统 计并汇总各关键词的频次,建立两列数据,第一列为"关键词",第二列为具体关键词出现 的"频次"。

SPSSAU 平台支持直接对 txt 文本文档进行分词,因此可以先将 txt 文本文档通过 128/187 SPSSAU 平台的【词云】模块自动完成分词任务,然后导出分词后的词频数据文件,在这 个过程中可人工对词频数据进行适度整理,如将一些近义词进行合并(保留一个关键词后 加总频数)。除此之外,还可使用 SPSSAU 的【文本分析模块】进行深度的文本分析。

2) 上传数据

将整理后的词频数据另存为 Excel 数据文件,通过 SPSSAU 平台的【上传数据】模块 上传至【我的数据】模块。

3) 词云图

通过【词云】模块,使用第 2)步词频数据绘制的词云图,结合研究主题对词云图进行解释,分析方法主要用于观察和分析高频关键词。

6.2.2 实例分析

下面结合问卷填空题进行实例分析,进一步介绍填空题及词云分析。

【例 11-4】txt 文本数据。某科学网站在读者满意度影响因素调查研究中,通过在问 卷中设置填空题"请问您访问该网站的目的是什么?"受访者需要填写具体的来访目的,收 集了 71 条有效作答,现尝试对该填空题进行分析。

1) txt 文本文档

71 条有效作答均为文本格式,每行代表一位受访者的作答,共 71 行文本,保存为一份独立的 txt 文本文档,命名为"词云数据.txt"。例 11-4 词云数据(部分)如图 6-5 所示。

🗐 词云数据.txt - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
统计学小白
学习 统计学,谢谢
有问题请教
宽得 统计学很强大
交流,进步!
喜欢_统计学
学习研究 统计学
分享和学习
%计字新于 你认要 相相言
筑计字 想提尚 你让要要提供了帮助
1931 计字数据分析水明
现仕日子 筑订子。刈奴店分竹很感兴趣

图 6-5 例 11-4 词云数据(部分)

2) 上传文本文档

在仪表盘中先依次单击【可视化】→【词云】模块,再单击分析框上方的【上传 txt 文件】按钮,将提前准备好的"词云数据.txt"文本文档上传到平台中。词云操作界面如图 6-6 所示。

ल व	视化 🖌
散点图	直方图
箱线图	词云
吴差线图	P-P/Q-Q图

图 6-6 词云操作界面 上传后平台会自动完成分词与词频统计,并绘制词云图。 3)词云分析

主要结果即词云图,例 11-4 词云图如图 6-7 所示。



图 6-7 例 11-4 词云图

分词后以关键词形式进行展示,关键词字体越大表示出现的频数越高,或者说该关键 词权重越大。从图 6-7 中可以发现,"统计""统计学""学习"较突出,很容易联想到受访者来 访的目的是以学习统计学知识为主的。

【例 11-5】词频数据。SPSSAU 平台允许用户将词频数据导出或下载,这样一来就可 以借助平台自动完成分词结果,并对关键词进行二次整理,将一些词义相近的关键词词频 进行合并,以提高部分关键词的词频。本例经过重新整理得到名为"例 11-5 词云数据.xls" 的 Excel 数据文档,试使用该词频数据文档进行词云图分析。

1) 上传数据

在【我的数据】模块中单击【上传数据】按钮,将"例 11-5 词云数据.xls"文档上传至 平台并进行分析。数据中包括"关键词"和"频次"两个变量,关键词词频大于 5 的数据如表 6-6 所示。

关键词	频次
统计学	41
学习	34
交流	8
数据分析	7
爱好者	5
问题	5

表 6-6 关键词词频大于 5 的数据

2) 词云图

🕑 可视化 🖌 🖌		开始分析	□ 上传txt文件 Q
散点图	直方图	关键词	分析项(定类)
箱线图	词云		
误差线图	P-P/Q-Q图		
Roc曲线	象限图	频次	加权项【可选】

图 6-8 例 11-5 词频数据与词云分析操作界面

在仪表盘中依次单击【可视化】→【词云】模块,从左侧标题框中将【关键词】变量 拖曳至【分析项(定类)】分析框中,将【频次】变量拖曳至【加权项(可选)】分析框中,最 后单击【开始分析】按钮。例 11-5 词频数据与词云分析操作界面如图 6-8 所示。

3) 结果分析

例 11-5 词云图如图 6-9 所示。将词云图(图)与高频关键词数据(表)结合起来进行解释 和分析,不难发现"学习""统计学"较突出,频次依次有 34 次和 41 次,"交流""数据分析"的 频次分别为 8 次和 7 次,频次大于 5 次的还有"爱好者""问题"这两个关键词。其他较重要 的关键词还包括"知识""新手""请教""研究生"等。



图 6-9 例 11-5 词云图

对照研究目的与本例填空题的内容,可将以上高频关键词的信息进行归纳总结,访问 网站的目的主要是学习统计学知识、交流数据分析的问题,网站的受众包括统计学或数据 分析的爱好者、刚开始学习统计知识的新手或研究生。

6.3 项目分析

量表问卷预调查阶段的分析方法主要包括效度、信度分析及项目分析,以起到优化和 调整问卷设计的目的,最终形成正式量表问卷。本节主要介绍项目分析方法在量表问卷研 究分析中的应用,此处强调,项目分析针对的是量表问卷的数据资料,普通问卷不适合。 通过在仪表盘中依次单击【问卷研究】→【项目分析】模块来实现。

6.3.1 原理介绍

1.项目分析的广义概念

项目分析,从狭义上也可称作项目区分度分析,主要目的在于检验题项的可靠程度。 具体做法是,考察高分与低分两组受访者在每个题项上有无差异,若有差异则说明该题项 具有区分度;若没有差异则说明该题项区分度比较差。没有区分度的题项,要综合其他分 析方法的结果考虑删题或对该题项进行调整。

例如,问卷采用李克特五级量表收集数据,受访者在某题项上的打分为3分或4分, 在总量表高低分两组间该题项无差异,意味着该题项没有区分度、可靠性差,必要时应考 虑删除。

广义上的项目分析,除高低分组区分度外,其内容还包括同质性检验,如每个题项与 量表总得分的相关性、题项与公因子的共同度、总量表的内部一致性信度。其中,题项与 量表总得分的相关性能反映题项用来测量相同维度的同质性。若相关性达到中高程度,则 说明题项与总量表同质;若相关性偏低则考虑删除该题项。

SPSSAU 平台的【项目分析】模块,提供了高低分项目区分度分析及题项与量表总得 分相关性分析,有关因子分析共同度、内部一致性信度的内容在接下来的章节介绍。

2.高低分项目区分度分析的原理

高低分项目区分度分析的基本思想是,将量表总得分划分为高、低分组,使用 t 检验 比较各题项数据在高、低分两组间的差异,可以细分为 4 个步骤,如图 6-10 所示。



图 6-10 高低分项目区分度分析步骤

1) 计算量表总得分

检查各题项是否均为正向计分题,若有反向计分题则要提前进行正向化处理,所有量 表题项加总求和计算总得分数据。

2) 确认百分位数

将量表总得分数据进行升序排列,一般取第 27、73 百分位数的取值作为分割点,将 总得分数据划分为 3 段,低于第 27 百分位数的为低分段,第 27 ~ 73 百分位数的为中分段, 高于第 73 百分位数的为高分段。

3) 高低分分组

创建一个三水平分组变量,一般用数字1代表低分组,数字2代表中分组,数字3代 表高分组,也可以只创建一个二水平分组变量,数字1代表低分组,数字2代表高分组, 不包括中分组。

4) 独立样本 t 检验

以量表题项作为因变量,采用独立样本 t 检验分析高分组与低分组受访者在每个题项 数据上有无差异。对 t 检验结果进行解释,若 p 值小于 0.05 则说明题项在高低分组间存在 统计学差异,认为题项具有区分度;反之,若 p 值大于 0.05 则说明该题项无区分度,可考 虑删除该题项。有时也可将 t 统计量称为决断值 CR, CR 小于 3 时认为题项在高低分组间 无区分度,可考虑删除该题项。

3.题项与量表总得分相关性分析的原理

各题项与量表总得分相关性分析较简单,首先通过加总的方式计算量表所有题项的总 得分,然后分析总得分与每个题项数据的 Pearson 相关系数,根据相关性来优化量表题项。 一般认为 Pearson 相关系数小于 0.4 则表示个别题项与量表是一种低相关关系,同质性较 差,可考虑删除该题项。

6.3.2 实例分析

结合具体案例进一步介绍高低分项目区分度分析和题项与量表总得分相关性分析的应用。

【例 11-6】继续沿用"在线英语学习网站课程购买意愿影响因素"的调查问卷,数据文

档见"例 11-6.xls"。该问卷研究设计了 19 个李克特五级量表题用来研究产品、促销、渠道、 价格、个性化服务和隐私保护对购买意愿的影响。假设该问卷正处于预调查阶段,现在尝 试对这 19 个题项进行项目分析,以考察各题项是否具有区分度。如果发现个别题项区分 度差则考虑删除,或者分析题项与总得分的相关性,相关系数偏低的题项也考虑删除。

1) 数据与案例分析

本例问卷的 19 个题项均为李克特五级题目,变量标题为"q1"到"q19"。预调查阶段主要进行信效度分析和项目分析,本例为项目分析,具体分析内容包括高低分组区分度和题项总得分相关性。

2) 项目分析

前面从原理上介绍了高低分项目区分度分析、题项与量表总得分相关性分析,在实际 分析过程中,这些步骤均由 SPSSAU 平台自动完成,我们只需要解释和分析结果。

将数据导入【我的数据】模块, 在仪表盘中依次单击【问卷研究】→【项目分析】模块, 在标题框中选中 19 个题目变量并拖曳至右侧的【分析项(定量)】分析框中, 在分析框上方的【分位数】下拉列表中选择【27/73 分位法(默认)】。此处注意, SPSSAU 平台提供了两种确认分组标准的百分位数, 一般默认选择 27/73 分位法即可。如果勾选【保存信息】 复选框, 则平台会将高低分组变量另存到数据集中, 以便后续分析使用, 一般情况下无须保存。项目分析操作界面如图 6-11 所示, 最后单击【开始分析】按钮。



图 6-11 项目分析操作界面

3) 结果分析

输出的结果包括高低分项目区分度与量表题项与量表总得分相关性两部分,要分别进 行解释和分析。

(1) 首先来解释和分析高低分项目区分度,结果如表 6-7 所示。

表 6-7 高低分项目区分度分析结果

				1
项目	组别(平均位	直±标准差)	<i>t</i> () 本 断值)	n
	低分组(<i>n</i> =86)	高分组(<i>n</i> =89)		ρ
q1	2.74 ± 0.96	4.21 ± 0.83	10.828	0.000**
q2	2.62 ± 0.90	4.02 ± 0.93	10.183	0.000**
q3	3.02 ± 1.11	4.55 ± 0.69	10.913	0.000**
q4	2.49 ± 0.89	3.71±1.13	7.941	0.000**
q5	1.92 ± 0.83	2.11 ± 0.87	1.506	0.134
q6	2.37±0.91	4.04±0.93	12.046	0.000**
q7	2.55 ± 0.95	4.25 ± 0.79	12.882	0.000**
q8	2.65 ± 1.04	4.08 ± 0.81	10.098	0.000**
q9	2.73 ± 0.95	4.08 ± 0.92	9.519	0.000**
q10	2.69 ± 1.00	4.24 ± 0.77	11.537	0.000**
q11	2.79 ± 0.86	4.10±0.92	9.763	0.000**

q12	2.73 ± 0.83	4.11±0.87	10.704	0.000**
q13	2.78 ± 0.86	4.26 ± 0.76	12.060	0.000**
q14	3.38 ± 1.17	4.80 ± 0.57	10.117	0.000**
q15	3.26 ± 1.15	4.73±0.58	10.657	0.000**
q16	2.88 ± 0.76	4.17±0.73	11.448	0.000**
q17	3.00 ± 0.78	4.12 ± 0.70	9.994	0.000**
q18	2.98 ± 0.70	4.19 ± 0.80	10.709	0.000**
q19	3.06 ± 0.66	4.07 ± 0.78	9.266	0.000**

注: **p<0.01。

从表 6-7 可知, q5 这个题项的 t 检验 p=0.134>0.05, 说明该题项在高分组与低分组 两组间无差异, 即该题项的区分度较差, 可考虑删除 q5 题项。其他题项在高低分两组间 的差异有统计学意义 (p<0.05), 均可保留。

除根据 t 检验的 p 值进行题项筛选及推断之外,还可以使用 t 检验的 t 统计量进行题 项筛选及推断,此处一般将 t 统计量称为决断值 CR,当 CR 值大于 3 时认为题项数据在高 低分组间差异显著,当 CR 值小于 3 时认为题项数据在高低分组间无差异。本例 q5 题项的 决断值 CR 为 1.506<3,也说明该题项在高分组与低分组间无差异,区分度差可考虑删除。 其他题项的 CR 值均大于 3,认为存在良好的区分度。

(2)题项与量表总分相关性。量表题项与量表总得分相关性分析结果如表 6-8 所示,
第 2 列同表 6-7 为 t 检验的结果, 第 4 列为题项与量表总分的相关系数, 第 5 列为相关系数的 p 值。

项目	决断值(CR)	<i>p</i> 值(CR)	与量表总分的相关系数	p 值(与量表总分的相关)
q1	10.828**	0.000	0.657**	0.000
q2	10.183**	0.000	0.621**	0.000
q3	10.913**	0.000	0.595**	0.000
q4	7.941**	0.000	0.514**	0.000
q5	1.506	0.134	0.135*	0.019
qб	12.046**	0.000	0.681**	0.000
q7	12.882**	0.000	0.670**	0.000
q8	10.098**	0.000	0.625**	0.000
q9	9.519**	0.000	0.517**	0.000
q10	11.537**	0.000	0.618**	0.000
q11	9.763**	0.000	0.614**	0.000
q12	10.704**	0.000	0.641**	0.000
q13	12.060**	0.000	0.707**	0.000
q14	10.117**	0.000	0.594**	0.000
q15	10.657**	0.000	0.593**	0.000
q16	11.448**	0.000	0.680**	0.000
q17	9.994**	0.000	0.667**	0.000
q18	10.709**	0.000	0.632**	0.000
q19	9.266**	0.000	0.586**	0.000

表 6-8 量表题项与量表总得分相关性分析结果

注: * p<0.05 ** p<0.01。

可以发现,19个题项中 q5 题项与量表总分的相关系数为 0.135 小于 0.4,说明该题项 与量表的同质性较差,可以考虑删除。其他题项与量表总分的相关系数均大于 0.4,说明同 质性良好。

6.4 效度分析

量表结构效度分析是预调查阶段的一项重要任务,一般使用探索性因子分析。第 6.1 节我们已经介绍了因子分析的原理思想及其在数据降维方面的应用,本节重点介绍因子分 析在量表问卷结构效度中的应用,通过在仪表盘中依次单击【问卷研究】→【效度】模块 来实现。

6.4.1 结构效度

1.效度与结构效度

问卷的效度从广义层面上指测量工具能够准确测出心理或行为特质的程度,针对的是 测验结果是否正确可靠。效度一般分为3种类型:内容效度、效标效度和结构效度。

(1)内容效度:本质上是问卷题目的命题逻辑分析,适用于普通问卷题目,如单选题、 多选题及量表题目,一般用文字来描述内容效度的情况。例如,介绍问卷题目来源,设计 过程,专家讨论、鉴定以形成问卷,说明问卷具有内容效度。

(2)效标效度:一般指问卷数据与外在校标数据间的相关程度,相关程度越高说明效 标效度越好。外在校标,如真实的行为成果(学生考试成绩)、本身就具有良好效度信度的 成熟量表等。

(3)结构效度:量表结构效度是指量表问卷能够准确测出理论或预设维度概念的程度,可测试量表题项和维度结构设计是否合理。在预调查阶段可使用探索性因子分析进行探究,当提取的因子-题项对应关系与理论或预设维度概念-题项对应关系相符合时即说明量表具有良好的结构效度。

例如,初设量表用 25 个题项测量感知有用性(7 题)、感知易用性(6 题)、满意度 (6 题)、忠诚度(6 题)4 个维度,预调查问卷数据进行探索性因子分析,最终提取出 4 个因子,而且每个因子载荷归属的题项与预设的对应关系相符合,说明初设量表结构效度 良好;如果提取因子的个数不是 4 个,或者一部分题项载荷出现"张冠李戴"(设计时用于 测量指定维度的题项在探索性因子分析结果中归属于其他不同的维度)、"纠缠不清"(同一 个题项横跨两个以上因子)的情况,个别题项单独成为一个独立因子等情况时,表明初设 量表的结构效度较差,一些异常题项考虑删除,从而实现对初设量表的优化和修改。

2.探索性因子分析结构效度术语说明

结构效度针对的是量表问卷数据资料,应满足探索性因子分析对数据的要求,普通题型,如单选题、多选题数据不能进行结构效度分析。前面已经介绍过因子分析的思想原理, 此处重点强调探索性因子分析在结构效度分析时的相关内容。

(1) KMO 与巴特利特检验。由于探索性因子分析要求题项间具有一定的相关性基础, 因此需要对是否适合进行因子分析进行研究,可用 KMO (取样适切度)和巴特利特检验进 行判断,若 KMO 值大于最低标准 0.5,巴特利特检验 p 值小于 0.05,则说明数据适合进行 因子分析,反之不适合。

(2)因子个数。提取的因子个数可指定为量表维度概念的个数,如某研究设计的量表需要测量6个变量,预调查阶段通过探索性因子分析来分析结构效度,此时可指定提取的因子个数为6个。若不指定提取的因子个数则由因子分析按特征根大于1的标准来确认因子个数。

(3)因子分析共同度。因子分析共同度是指题项能被因子解释的变异量,反映的是题项与因子的同质性,在量表结构效度分析中,可用于衡量题项的质量,常作为广义项目分 135/187 析中的一个指标。一般因子分析共同度低于 0.4 说明对应的题项与因子的同质性较差要引起关注,如果因子分析共同度低于 0.2 则考虑直接删除。

(4) 载荷系数。载荷反映的是因子与题项的相关关系,载荷值越大说明题项与因子关 联程度越高,也可以理解为题项可作为因子的代表性数据,或题项归属于某个因子。一般 载荷低于 0.4 (有时也可以按 0.5 标准)说明题项与因子没有对应归属关系。

6.4.2 实例分析

下面通过一份量表问卷数据资料分析量表的结构效度。

【例 11-7】某研究自编了一份"学校知识管理量表",在量表设计时要结合相关理论,包括 3 个维度: c1~c6 题项测量"知识创新"、c7~c13 题项测量"知识分享"、c14~c19 题项测量"知识获得",共 19 个题项,试分析该量表的结构效度。数据来源于吴明隆(2010.5),本例数据文档见"例 11-7.xls"。1)数据与案例分析

案例数据有 19 个题项,题项名称依次为 c1~c19。量表设计时所有题项均采用李克特 五级题目,数据可视为定量数据进行统计分析。本例预设的维度有 3 个,如果探索性因子 分析提取的公因子具备合理性(可按预设的 3 个维度概念进行命名、每个因子至少含 3 个 题项),且因子-题项的对应关系与维度-题项的对应关系相符合,则可说明该量表具备良好 的结构效度。在此过程中,若发现个别题项在共同度指标和载荷系数上表现不佳则可考虑 删除。

2) 结构效度分析

数据读入平台后,在仪表盘中依次单击【问卷研究】→【效度】模块,这是专门用来 对量表问卷做结构效度的模块,我们并不需要依次单击【进阶方法】→【因子】模块来完 成。

本例计划采取默认的特征根大于 1 的标准提取因子。将标题框内【c1】~【c19】题 项变量全部拖曳至右侧的【分析项(量表)】分析框中,在分析框上方的【维度个数设置】 下拉列表中选择【维度个数设置(自动)】,即按特征根大于 1 的标准提取公因子。效度分析 操作界面如图 6-12 所示,单击【开始分析】按钮。

▣ 问卷	新究 🗸		开始分析	i i	维度个数设置(自动) ∨	Q
信度	效度		c1			
多选题	单选-多选		c2		3	
多选-单选	多选-多选		c3		5	
项目分析	权重		c4		6 7	
验证性因子分		0	c5	分析项(量表)	8	
析			c6		10	
路径分析	5EM		c7		11	
调节作用	中介作用		c8		13	
调节中介	KANO模型		c9		14 15	

图 6-12 效度分析操作界面

3) 结果分析

探索性因子分析结构效度分析的结果主要包括 KMO 值、巴特利特检验、特征根、共同度、方差解释率、累积方差解释率,以及载荷系数(该模块默认进行正交旋转)。平台能 将以上结果合并成一张表格,在载荷系数的展示上,可根据题项变量名称或因子的顺序进 行排序操作(单击表格上方的【排序】按钮),本例选择按题项变量名称排序。探索性因子 分析效度分析结果如表 6-9 所示。

勾护		载荷	廿日帝 (八田之亡美)		
百松	因子 1	因子 2	因子 3	因子 4	共同度(公凶)万左/
cl	0.087	0.062	0.663	0.394	0.606
c2	0.063	0.093	0.810	0.094	0.678
c3	0.055	0.101	0.887	0.076	0.805
c4	0.070	0.009	0.743	-0.053	0.560
c5	0.139	0.058	0.711	-0.194	0.565
c6	0.137	0.157	0.509	-0.092	0.311
c7	0.696	0.311	0.127	-0.147	0.618
c8	0.814	0.181	0.178	0.092	0.735
c9	0.877	0.182	0.078	0.013	0.808
c10	0.803	0.216	-0.016	0.024	0.692
c11	0.728	0.153	0.212	0.174	0.629
c12	0.091	-0.062	-0.003	0.862	0.755
c13	0.886	0.173	0.075	0.001	0.820
c14	0.469	0.643	0.107	-0.101	0.655
c15	0.189	0.945	0.117	-0.037	0.945
c16	0.391	0.644	0.202	-0.086	0.616
c17	0.170	0.930	0.077	0.049	0.903
c18	0.425	0.504	0.261	-0.295	0.591
c19	0.177	0.912	0.025	0.046	0.866
特征根值(旋转前)	7.208	2.834	2.041	1.075	_
方差解释率(旋转前)	37.936%	14.914%	10.744%	5.659%	_
累积方差解释率(旋转前)	37.936%	52.850%	63.594%	69.253%	_
特征根值(旋转后)	4.589	3.992	3.445	1.132	_
方差解释率(旋转后)	24.153%	21.009%	18.131%	5.959%	_
累积方差解释率(旋转后)	24.153%	45.163%	63.294%	69.253%	_
KMO 值	0.855			_	
巴特利特球形值	3079.151			_	
df	171			_	
<i>p</i> 值	0.000				

表 6-9 探索性因子分析效度分析结果

(1) KMO 值与巴特利特检验。本例 KMO=0.855>0.8,认为数据适合进行因子分析。 巴特利特检验的 p 值小于 0.05,也认为数据适合进行因子分析。

(2)特征根值与累积方差解释率。由于【效度】模块默认对公因子进行最大方差法正 交旋转,因此特征根值和累积方差解释率统一看旋转后的结果。按特征根值大于 1 的标准, 用探索性因子分析提取 4 个公因子,累积方差解释率为 69.253%。在社科领域,因子的累 积方差解释率在 60%以上一般是可以接受的。

(3) 共同度。表 6-9 中最右侧一列为共同度指标,也称之为公因子方差,我们发现 c6 题项的共同度为 0.311<0.4,应引起研究者的重视,没有题项共同度低于 0.2 的情况, 初步判断没有需要删去的题项。

(4)因子-题项对应关系。载荷绝对值大于 0.4 时被平台标注为蓝色,平台按载荷大于 0.4 的标准对题项与因子的对应关系进行判断,即一个题项与某个因子的载荷如果大于 0.4 则认为其归属于该因子。

显然"因子 4"仅包括 c12 题项,其下的题项个数少于 3 个,不具有合理性,提示 c12 题项可删除。

存在明显异常的还有 c14 题项和 c18 题项,这两个题项在"因子 1"和"因子 2"上均有较

高的载荷, 按载荷 0.4 标准难以判断题项与因子的对应关系, 为常见的"纠缠不清"现象。此时可考虑按载荷 0.5 标准进行判断, 由此可知, 这两个题项在"因子 1"上的载荷均小于 0.5, 而在"因子 2"上的载荷均大于 0.5, 认为 c14 题项和 c18 题项均归属于"因子 2"。

至此结果已明确: c1~c6 题项与"因子 3"存在对应关系,结合题项内容可命名为"知识 创新"; c7~c13 题项(不包括 c12 题项)与"因子 1"存在对应关系,结合题项内容可命名 为"知识分享; c14~c19 题项与"因子 2"存在对应关系,结合题项内容可命名为"知识获得"。 综合以上分析,我们通过探索性因子分析提取了 4 个因子,删除 c12 题项后其余 18 个题 项与因子的对应关系与预设结构相符(结果略),因此认为删除 c12 题项后量表具有良好 的结构效度。

(5) 其他说明。读者可自行测试删除 c12 题项后按特征根值大于 1 的标准提取因子的结果,与全部 19 个题项指定提取预设维度个数 3 的结果相比较,讨论最终的结构效度 结果,最终结论同样是删除 c12 题项后量表具有良好的结构效度,此处略。

6.5信度分析

项目区分度分析研究的是每个题项的适合性,结构效度分析研究的是量表问卷本身的 准确性,那么对于一次调查数据(量表问卷资料),它结果的稳定性、一致性如何衡量,这 就是信度分析的任务,通过在仪表盘中依次单击【问卷研究】→【信度】模块来实现。

6.5.1 信度系数

信度(Reliability),具体是指量表各维度层面与总量表的稳定性、一致性。它针对的 是量表数据结果而不是量表问卷工具本身,信度分析适用于李克特量表问卷数据资料的定 量数据。

1.信度系数的类型

常用于内部一致性信度系数的指标包括 Cronbach α 系数、折半系数、McDonald omega 系数、theta 系数等。

(1) Cronbach α 系数。在李克特态度量表中,最常用的信度指标之一为 Cronbach α 系数,简称 α 系数。该系数介于 0 到 1 之间, Cronbach α 系数越高代表量表的一致性越高。它与题项个数有关,题项越多 Cronbach α 系数越大,因此维度 Cronbach α 系数低于 总量表 Cronbach α 系数。

(2)折半系数。折半系数将量表题项分为两半,通过计算两半得分的 Cronbach α 系 数及相关系数,进而估计整个量表的信度。此处应注意,所有题项拆分为两半时,每半的 题项个数可能相等也可能不相等,从而分别计算的是等长折半系数和不等长折半系数。 如果是经典量表题项,并且某个维度的题项较多(如大于 5 个)时,则使用折半系数来衡 量信度。

(3) McDonald omega 系数。McDonald omega 系数的计算原理是,利用因子分析浓 缩信息,即维度或量表所有题项进行因子分析并只提取一个公因子,得到载荷系数 loading 后基于载荷计算 McDonald ω 系数。loading 值整体绝对值越大时, McDonald ω 系数值越 高。

(4) theta 系数。theta 系数的原理同 McDonald omega 系数,也是使用因子分析来 研究内部一致性的。其计算公式如下:

$$\theta = \frac{N}{N-1}(1-\frac{1}{\lambda})$$

式中, N 为题项个数; λ 为最大特征根值。可以看到, 当题项个数越多时, theta 系数 值越大, 而且最大特征根越大, theta 系数值也会越大。

2.信度系数解读标准

Cronbach α 系数有多个可供参考的解读标准, 吴明隆(2010.5)综合多位学者的观点 认为, 分层面最低的内部一致性信度系数要在 0.5 以上, 最好能高于 0.6, 而整份量表最低 的内部一致性信度系数要在 0.7 以上, 最好能高于 0.8。

折半系数、McDonald omega 系数与 theta 系数在评价时可参考 Cronbach α 系数的解 读标准。

3.信度分析的使用

一般在预调查阶段与正式量表分析阶段都会使用信度分析,前者用于项目分析,后者 用于报告和评价正式量表信度结果。

(1)预调查阶段项目分析。在量表问卷的预调查阶段,信度分析的作用在于考察初设 量表的信度是否达标,以及哪些题项删除后可提高或改善信度水平。使用 Cronbach α 系 数进行信度分析时,平台会输出"项已删除的 α 系数"和"校正项总计相关性(CITC)"两个 指标值。

项已删除的 α 系数: 指删除某个题项后的新 Cronbach α 系数, 如果其值明显大于删 除题项前的 Cronbach α 系数, 则说明该题项同质性较差, 删除该题项可改善维度或量表 的信度水平, 此时认为对应的这个题项应当删除。例如, "感知有用性"维度, 初设 6 个题 目时该维度的 Cronbach α 系数为 0.68, 第 3 个题目的"项已删除的 α 系数"值为 0.86, 意 味着如果删除第 3 个题项, 剩余 5 个题项构成的新维度 Cronbach α 系数为 0.86, 删除题 项后信度水平得到了明显提高。

校正项总计相关性(CITC): CITC 同样可作为删除题项的依据,它是指某个题项与其他剩余题项总得分数据之间的相关系数,显然 CITC 越低说明对应的题项同质性越差,可考虑删除。如果 CITC 值小于 0.3,则通常考虑删除对应题项。

(2)正式量表问卷信度评价。正式量表一般较少涉及删除题项进行优化,因此此阶段 信度分析的目的主要是报告和评价当前收集的问卷量表的信度水平。如果此阶段仍出现信 度不达标的情况,必要时也可以结合其他分析结论考虑删除某些题项以改善信度水平。

6.5.2 实例分析

下面通过具体实例进一步介绍量表问卷研究分析方法中信度分析的应用。

【例 11-8】继续沿用【例 11-7】的案例,某研究自编的"学校知识管理量表",包括 3 个维度: c1~c6 题项测量"知识创新"、c7~c13 题项测量"知识分享"、c14~c19 题项测量 "知识获得",共 19 个题项,试分析该量表"知识分享"维度的信度及总量表的信度。

1) 数据与案例分析

案例数据有 19 个题项变量,变量名称依次为 c1~c19,均为定量数据,满足信度分析 的基本要求。本例示范对预设的"知识分享"维度进行信度分析,最后报告总量表的信度。 信度系数选择较常用的 Cronbach α 系数和折半系数,若发现有题项同质性差问题则考虑 删除。

2) "知识分享"维度信度分析

数据读入平台后, 在仪表盘中依次单击【问卷研究】→【信度】模块, 将【c7】~ 【c13】题项拖曳至【分析项(量表)】分析框中, 即先分析"知识分享"维度的信度, 在分析 框上方【信度系数】下拉列表中选择【Cronbach α 系数】, 最后单击【开始分析】按钮。 信度分析操作界面如图 6-13 所示。

■ 问卷	研究 🗸		开始分析	Ē	Cronbach α系数 ~ Q
信度	效度		c7		ff半系数 McDonald Omega
多选题	单选-多选		c8		theta系数
多选-单选	多选-多选		c9		
「市日公析	収重		c10		公析顶(景丰)
シロケック			c11		刀/们列(里衣)
· 新	对应分析		c12		
路径分析	结构方程模型 SEM		c13		
调节作用	中介作用				

图 6-13 信度分析操作界面

3) "知识分享"维度结果分析

项已删除的 α 系数、Cronbach α 系数及 CITC 相关系数如表 6-10 所示。

名称	校正项总计相关性(CITC)	项已删除的α 系数	Cronbach α 系数
c7	0.616	0.816	
c8	0.769	0.795	
c9	0.796	0.791	
c10	0.708	0.800	0.841
c11	0.677	0.807	
c12	0.073	0.912	
c13	0.796	0.789	

表 6-10 项已删除的 α 系数、Cronbach α 系数及 CITC 相关系数

注:标准化 Cronbach α 系数: 0.863。

第 4 列"知识分享"维度的 Cronbach α 系数为 0.841>0.8, 说明该维度具有良好的信度。 第 3 列项已删除的 α 系数显示, 删除 c12 题项后剩余 6 个题项构成的新维度 Cronbach α 系数为 0.912, 明显大于原来的 0.841, 提示删除 c12 题项可改善维度信度水平。第 2 列为 校正项总计相关性 (CITC) 值, c12 题项的 CITC 值为 0.073<0.3, 提示该题项与维度的同 质性差,可考虑删除。如果当前数据为预调查阶段的信度分析, 那么可综合项已删除的 α 系数与 CITC 值的提示信息,将 c12 题项删除以改善"知识分享"维度的信度水平。

4) 总量表信度分析

接下来示范折半系数的应用,将【c1】~【c19】题项(对总量表)拖曳至分析框中, 在【信度系数】下拉列表中选择【折半系数】(本例采用 Cronbach α 系数亦可,此处仅示 范折半系数),单击【开始分析】按钮对结果进行分析。

折半信度分析如表 6-11 所示, 平台将 19 个题项拆分为两部分, 前半部分包括 10 个 题项, 后半部分包括 9 个题项, 为不等长的情况, 因此我们主要解读的是不等长折半系数 (Spearman-Brown 系数), 本例折半系数为 0.815, 显然在 0.8 以上, 说明总量表具有良 好的信度。

值 0.817 前半部分 项数 10 0.833 Cronbach a 系数 值 后半部分 项数 9 19 总项数 前半部分和后半部分间的相关系数值 0.688 等长 0.815 折半系数(Spearman-Brown 系数) 不等长 0.815 Guttman Split-Half 系数 0.812

表 6-11 折半信度分析

6.6验证性因子分析

量表数据资料本质上是对潜变量数据的测量,并对其进行研究分析,应当采取潜变量 数据分析方法。潜变量数据分析方法常见的有验证性因子分析、结构方程模型等。

本节介绍验证性因子分析,通过在仪表盘中依次单击【问卷研究】→【验证性因子分 析】模块来实现。

6.6.1 方法概述

1.潜变量与显变量

潜变量是指不能直接观测到的变量,常用于对特殊行为或心理态度的测量,如老年人 心理抑郁,不能直接提问"你感到抑郁吗?",而是需要通过一系列跟抑郁有关的同质化题 项来测量抑郁。此时抑郁即为潜变量,用来测量它的题项被称为显变量或观测变量。

再如通过因子分析从数学、物理、化学成绩等提取到"理科因子",从语文、历史、政治、地理成绩提取到"文科因子",理科与文科是潜在的概念,并不能直接观测理科和文科的成绩。

显变量是可以直接观测的变量,如量表中的具体题项得分,或者常见的身高、体重等 可测定取值的数据。潜变量和显变量是一对统计术语,有时候也分别称它们为潜在变量和 外显变量。



图 6-14 验证性因子分析模型示意图

2.验证性因子分析概念

验证性因子分析的英文全称为 Confirmatory Factor Analysis, 简称 CFA。它的应用目的 在于,验证潜变量与相对应题项间的关系是不是契合研究者初设的理论关系。用统计术语 来说,即验证显变量分析得到的方差协方差矩阵与预设模型(研究者的理论模型)的方差 协方差矩阵是否一致,若模型契合评价达标则认为量表的结构与实际数据契合,所设计的 题项可以有效测量潜变量因子。

一般可用路径图来反映验证性因子分析模型题项与因子间的测量关系,图 6-14 所示为验证性因子分析模型示意图。在该模型中,量表题项 A1~A3 测量的是第一个因子 F1, 另外 4 个题项 B1~B4 测量的是第二个因子 F2。一个题项只受一个因子影响,没有交叉载荷的情况。

在验证性因子分析模型示意图中,用椭圆表示潜变量或因子,用矩形框表示量表题项 或观测指标,用双箭头表示因子之间的协方差或相关关系、从因子指向题项的单箭头表示 直接影响关系。例如,B1~B4 题项具有共同概念,这些共同概念提取后即为潜变量 F2, 潜变量可以解释题项的共同特质。直接影响关系表现为回归路径系数,标准化回归路径系 数也被称为载荷系数,一般用 λ 表示,载荷系数的平方即潜变量可解释观测题项的程度, 解释不了的剩余部分则为测量误差,一般用 e 表示。

验证性因子分析可以分为一阶验证性因子分析和高阶验证性因子分析两种,一阶验证 性因子分析是基础测量模型也较常见,当我们发现由众多题项测量得到的一阶验证性因子 之间存在较高的相关性(一般认为相关程度在 0.7 以上),在一阶验证性因子之上还可以继 续提取共同特质,这样的共同特质即二阶验证性因子,整个测量模型称为二阶验证性因子 分析模型。同理,如果二阶验证性因子仍然存在明显的相关,则可尝试提取三阶验证性因 子。

3.与探索性因子分析的区别

因子分析包括探索性因子分析和验证性因子分析两种。对于量表问卷数据来说,探索 性因子分析用于探索及发现潜变量因子与题项之间的结构关系,先对数据进行分析再探索 得到一个结构模型;与之相反,验证性因子分析先由研究者提出理论上的潜变量与题项的 结构关系,再通过量表数据分析来验证该结构关系的真实性。

在使用上,探索性因子分析主要用在量表预调查阶段,探索初设量表的结构效度,通 过共同度、载荷系数等为删除、优化题项提供依据。验证性因子分析主要用在正式量表分 析阶段,对已知的量表结构效度进行验证和评价。

此处注意,在量表问卷分析的整个过程中不应将探索性因子分析和验证性因子分析主 观割裂,二者既有区别又有联系。例如,在新量表编制过程中,一般会先进行探索性因子 分析,从而发现公因子与题项的对应归属关系,确立结构效度。该结果一方面可用于结构 效度分析,另一方面也可作为验证性因子分析结构关系的基础。最后用另外一批数据进行 验证性因子分析,并对前面已经确立的结构进行验证。

6.6.2 验证性因子分析步骤

验证性因子分析过程较其他分析方法要复杂一些,并不能通过一次分析就获得最终结果, 具体分析时还需要结合模型识别及模型拟合指标对模型进行适当的修正。图 6-15 所示为 验证性因子分析的一般步骤,主要包括五大环节。



图 6-15 验证性因子分析的一般步骤

1) 模型设定与识别

模型设定即按研究者的理论支持确立研究变量(维度/构念/因子/潜变量)、题项,以 及研究变量与题项间测量的对应关系,如用 a1~a5 题项测量维度"感知有用性",用 a6~ a10 题项测量维度"感知易用性"。

如果引用成熟量表,由于量表的维度和题目关系早已得到论证,这种情况下一般可以 直接进行模型设定并进行验证性因子分析。如果是新开发编制的量表,则建议先通过探索 性因子分析探索结构效度,有不合适的题项考虑删除,甚至先对个别维度的定义进行修订, 再进行验证性因子分析。

在开始估计参数之前,应检查模型是否可识别,所谓的识别即理论模型是否存在合适的解。验证性因子分析模型识别有 3 种结果,分别是不可识别、恰好识别、过度识别,其中恰好识别与过度识别均属于模型可识别的范畴。一般采用"t 法则"来判断模型是否可识别,具体来说,指模型中参数的个数不能超过方差协方差矩阵的内部元素数,即 t≤p(p+1)/2,t 代表模型中需要估计的参数个数,p 代表观测变量(在量表中即题项)的个数。

例如,7 个李克特五级量表题测量两个维度,待估计的参数包括5 个路径系数(每个 因子下任意一个路径系数固定为1,两个因子则有两个路径系数固定为1,所以待估计的 路径系数为7-2=5 个、7 个残差、2 个因子的方差,以及1 个相关系数,参数个数 t=5+7+2+1=15,小于方差协方差矩阵的元素数 p(p+1)/2=7×8/2=28,所以属于过度识别, 模型是可识别的。

通常来说,我们并不需要刻意去计算判断模型是否可识别,一个取巧的办法是直接查 阅统计软件工具输出的模型自由度 df,如果 df 大于等于 0,则意味着模型可识别,因为方 差协方差矩阵元素数减去待估计参数等于自由度 df。

此外,验证性因子分析与结构方程模型识别规则还包括:每个因子指定一个测量指标的未标准化回归路径系数为1;每个因子至少包括3个测量题项。如果遇到模型无法识别的情况,那么可以考虑结合模型理论删除个别自由参数,或者固定某些自由参数。

2) 模型参数估计

在验证性因子分析模型中,我们主要关注的参数包括未标准化回归路径系数(标准化 值即载荷系数)、测量残差及因子间相关系数。

有多种估计模型参数的方法,其中较常用的是最大似然法(ML),它要求数据为定量数据,且服从多元正态分布。在实际研究及分析中,对于量表问卷数据资料的正态性要求可适当放宽标准,如以各题项偏度系数、峰度系数作为依据进行大致判断,各题项数据不是严重偏离正态分布即可认为采用最大似然法估计的结果是稳健的。SPSSAU 平台默认采用最大似然法估计模型参数,能正常输出模型拟合结果,表明模型能够识别。

在参数的解释与分析方面,使估计的每个参数达到显著水平(p<0.05)是保证模型内 在质量的重要基础,如果某个参数不显著(p>0.05),那么提示对应的题项要考虑优化或删 除。待估计的参数包括测量路径上的未标准化回归系数、因子间的协方差、因子方差、测 量误差。

因子载荷即标准化回归系数,它反映了题项能被因子解释的变异程度,可依据因子载 荷了解各题项在因子中的相对重要性,经验法要求因子载荷大于 0.5,理想情况是大于 0.7, 载荷越大,越能说明题项可用来测量因子。

对于因子间相关系数,如果其大于 0.7 则表明因子间存在较强的相关性,意味着可以 从这些一阶验证性因子中继续提取共同特质,提示可考虑进行二阶验证性因子分析。

3) 模型拟合评价

验证性因子分析模型是否适配或是否达标,需要对比模型拟合指标的适配标准。常用 的指标包括卡方自由度比、RMSEA(近似残差均方根)、SRMR(标准化残差均方根)、GFI (拟合优度指数)、CFI(比较拟合指数)、NFI(标准拟合指数)和 TLI(非标准拟合指数) 等,还可以选用 AGFI(调整后的拟合优度指数)、AIC、BIC、PGFI等其他指标。吴明隆 (2010.10)总结的验证性因子分析模型及结构方程模型常用拟合指标及指标解读标准,如 表 6-12 所示。

指标类型	指标名称	一般接受范围				
绝对适配指标	卡方值	卡方检验p>0.05 适配良好				
	卡方自由度比	<3 模型适配良好				
	RMR	<0.05 模型适配良好				
	SRMR	<0.1				
	RMSEA	<0.1 基本可接受; <0.08 模型适配良好				
	GFI/AGFI	>0.9, 越接近 1 模型拟合越佳				
增值适配指标	NFI	>0.9, 越接近 1 模型拟合越佳				
	RFI	>0.9, 越接近 1 模型拟合越佳				
	IFI	>0.9, 可能大于 1				
	TLI	>0.9, 可能大于 1				
	CFI	>0.9, 可能大于 1				
简约适配指标	AIC/BIC	一般用于多个模型拟合的比较,取值小的模型拟合更佳				
	PNFI	>0.5 模型可以接受				
	PGFI	>0.5 模型可以接受				

表 6-12 验证性因子分析模型及结构方程模型常用拟合指标及指标解读标准

为了方便理解和解读各指标,一般将表 6-12 中的指标划分为 3 种。

(1) 绝对适配指标。绝对适配指标包括卡方值、卡方自由度比、RMR、SRMR、 RMSEA、GFI/AGFI。

关于模型适配的卡方检验,原假设观察数据的方差、协方差矩阵与预设模型的方差、 协方差矩阵相等,经卡方检验可推断该假设成立与否,当 p>0.05 时接受原假设,即认为 预设模型与量表数据结构适配。在实际分析中,卡方自由度比作为一个很重要的指标,一 般该值小于 3 大于 1,表明模型适配良好。

RMR、SRMR、RMSEA 这 3 个指标和测量残差有关, SRMR 可理解为 RMR 的标准化值, RMR<0.05, SRMR<0.1 认为模型的适配可接受。RMSEA 的使用较多, 一般认为 RMSEA 大于 0.1 时, 模型适配欠佳, 小于 0.08 时模型适配良好。

GFI、AGFI 这两个指标类似于回归分析中的 R2 与调整后的 R2 两个指标,一般标准为 大于 0.9 说明模型适配良好。

(2) 增值适配指标。增值适配指标包括 NFI、RFI、IFI、TLI、CFI 这 5 个指标, SPSSAU 平台整合的拟合指标中有 CFI、NFI、NNFI、TLI、IFI, 注意 TLI 和 NNFI 是等价的。 在具体取值上这几个指标大多介于 0~1, 其中 TLI(NNFI)、CFI、IFI 可能会出现大于 1 的情况。以上几个指标的一般标准为大于 0.9 说明模型拟合良好。

(3)简约适配指标。简约适配指标包括 AIC/BIC、PNFI、PGFI 等。 AIC/BIC 这两个指标的取值越小越能说明模型适配良好,可用于模型间的比较。PNFI 和 PGFI 的一般标准为大于 0.5 说明模型适配良好。

以上指标适用于验证性因子分析模型、结构方程模型,此处应注意,对模型拟合的评价,应综合多个指标的达标情况而定,一般不依据单个指标决定模型的拟合质量。例如, 模型总体适配卡方检验有时候会出现 p 值小于 0.05 的情况,这并不意味着模型拟合就一定 失败,还应当综合其他多项指标的达标情况最终确定拟合评价。

4) 模型修正(非必须)
模型是否需要修正是根据模型拟合评价来考虑的,并非所有的模型都需要修正。如果 模型拟合指标表现不佳,如卡方自由度比大于 3、GFI 小于 0.9、RMSEA 大于 0.1 等,则可 考虑结合模型修正指数(Modification Indices, MI)结果的提示对模型进行适当的处理。

MI 值较大时表示应建立变量间的协方差关系或删除某个题项以减少参数估计。一般 认为 MI 值大于 5 时模型才具有修正的必要,而且建议从最大的 MI 值开始逐一进行修正。 对于验证性因子分析模型的修正可以从因子-题项的对应关系与题项间的协方差关系这两个 方面入手。同时模型修正应建立在理论、经验及验证性因子分析模型的规范之上,不能为 了"修正"而修正。

5) 分析与应用

经过模型识别、参数估计、模型修正等过程,最终会得到一个达到适配标准的验证性 因子分析模型,根据研究分析的目的,对结果进行分析和应用。例如,根据载荷系数判断 结构效度并解释题目保留的情况,以及依据载荷系数计算组合信度(CR)、平均方差提取 量(AVE),并进行聚敛和区分效度的评价。再如将验证性因子分析用于多模型比较、共同 方法偏差检验等应用。验证性因子分析应用方向如表 6-13 所示。表中前 4 项是验证性因 子分析的常规应用目的,除此之外,有一些研究还将验证性因子分析用于量表问卷的共同 方法偏差的分析。

验证性因子分析 统计结果的应用	说明	重点结果解释
县丰建物为审	潜变量与题项对应关系是否契合预设或	因子载荷应大于 0.5, 理想情况是大于 0.7,
里农知何双反	理论设计的关系	且模型拟合达标
CR	测量信度水平	CR 值应大于 0.6, 最好大于 0.7
聚合效度	同一个维度概念下的题项归属该维度的 程度	AVE 值大于 0.5
区分效度	不同维度概念间的题项的不完全相关性	AVE 平方根满足相关性要求
共同方法偏差	问卷研究中常见的系统误差问题	验证性因子分析单因子模型拟合不达标

表 6-13 验证性因子分析应用方向

(1)量表结构效度。验证性因子分析模型总体需要达到拟合标准,载荷系数才具有显 著性。经验法要求因子载荷大于 0.5,理想情况是大于 0.7 (刘红云, 2019)。

(2) CR。CR 可作为潜变量的信度指标,反映潜变量的内在质量,由因子载荷计算得 到。刘红云(2019)认为 CR 值应大于 0.6,最好大于 0.7,该值越接近于 1 表示潜变量概 念信度/内在质量越好。

(3)聚合效度。聚合效度也称收敛效度,强调的是用来测量同一维度概念的题项会出现在同一个潜变量因子下的程度,一般用 AVE 值和 CR 值来衡量。AVE 值与 CR 值也是利用因子载荷进行计算的,AVE 作为聚合效度的指标,该值越大表示题项越能体现潜变量因子的同质性,一般要求该值大于 0.5。

(4)区分效度。区分效度是与聚合效度对应的概念,同个维度概念下的题项强调聚合效度,而不同维度概念间的题项则强调区分效度。一个良好的量表结构应该同时有良好的聚合效度和区分效度。一般用维度概念的 AVE 平方根与维度概念间的相关系数进行比较,若 AVE 平方根大于该维度与其他维度两两间的相关系数,则认为具有良好的区分效度。

(5) 共同方法偏差。共同方法偏差简称 CMV, 它是量表问卷从设计到实测过程中可 能出现的系统误差, 通常采用 Harman 单因素法进行检测。针对探索性因子分析, 如果所 有题项提取的第一个公因子的方差解释率大于 40% (也有说法是大于 50%), 则说明存在共 同方法偏差问题;针对验证性因子分析, 所有题项只用测量一个 CMV 因子, 如果该验证 性因子分析的拟合指标不良, 即所有题项并不能归属于同一个维度概念, 则说明量表数据 无共同方法偏差。本节仅在此处提出验证性因子分析可用于共同方法偏差的检测识别,共同方法偏差的具体内容、分析方法暂不展开,感兴趣的读者请参阅其他资料。

6.6.3 验证性因子分析实例分析

结合具体案例进一步介绍验证性因子分析的应用。

【例 11-9】继续沿用【例 11-7】的量表(对题目做少许删减)研究自编的"学校知识管理量表",根据相关研究理论和既往研究结论初设了3个维度:A1~A4题项测量"知识创新"、B1~B5题项测量"知识分享"、C1~C5题项测量"知识获得",共14个题项,数据文档见"例 11-9.xls",试分析该量表的结构效度、聚合效度、收敛效度。

1) 模型设定与识别

根据某种理论或参考成熟量表结构,由 14 个李克特量表题测量 3 个变量。预设模型 结构为: A1~A4 题项测量"知识创新"、B1~B5 题项测量"知识分享"、C1~C5 题项测量"知 识获得"。此处知识创新、知识分享、知识获得统一称作因子(Factor)。本例可通过一阶验 证性因子分析验证已知结构关系的适配情况,以及评价量表的聚合效度、收敛效度。

该预设模型的待估计参数个数为 31, 小于方差、协方差矩阵的元素个数 p(p+1))/2= 14×15/2=105, 从必要条件来看, 模型是可识别的。这一步通常模型是会满足的, 不是必 须要通过计算来验证的, 稍后也可以直接根据验证性因子分析输出的模型自由度来判断。

接下来将模型测量关系通过 SPSSAU 平台进行设定,数据读入平台后,在仪表盘中依 次单击【问卷研究】→【验证性因子分析】模块。SPSSAU 平台最多可分析 12 个因子,最 高支持二阶验证性因子分析,因子名称默认为"Factor1"到"Factor12",允许修改名称。在标 题框中选中【A1】~【A4】题项将其拖曳至右侧的【Factor1(量表题)】分析框中,

同样的操作,将【B1】~【B5】题项拖曳至【Factor2(量表题)】分析框中,将 【C1】~【C5】题项拖曳至【Factor3(量表题)】分析框中,建立起预设的一阶因子与题项 的对应关系。

在分析框右侧可直接修改因子名称,不做修改则默认为"Factor"加数字顺序号。本例将 Factor1 修改为"知识创新",Factor2 修改为"知识分享",Factor3 修改为"知识获得"。

如果分析目的是二阶验证性因子分析,则需要勾选【因子名称【可选】】窗格下方的 【二阶模型】复选框,并给二阶因子命名。本例仅示范一阶验证性因子分析,不勾选【二 阶模型】复选框。【二阶模型】复选框下面的【测量项协方差关系【结合 MI 指标设置】】 设定要等模型执行后根据 MI 指标的情况返回进行修正设定。

SPSSAU 科研数据分析方法与应用

开始分析 💼 输出M	MI>5 ~ Q	
A1	因子名称【可选】	
B1	知识创新	知识分享
C1	知识获得	Factor4
	Factor5	Factor6
Factor4(量表證)	Factor7	Factor8
Factor5(最表題)	Factor9	Factor10
Factor6(量表题)	Factor11	Factor12
Factor7(显表题)	□ 二阶模型 二阶因子名称	
Factor8(量表题)	测量顶协方差关系【结合MI指标设置】	
Factor9(量表题)	第1项	第2项
2 2	请选择 ∨ ↔	请选择 Y + -

图 6-16 验证性因子分析操作界面

在分析框上方的下拉列表中可设置 MI 的起始标准,一般选择【输出 MI>5】,即输出 MI 指标在 5 以上的潜在修正项目,最后单击【开始分析】按钮。验证性因子分析操作界面 如图 6-16 所示。

2) 模型参数估计结果解释

待估计的模型参数包括载荷系数、测量残差、因子间相关系数及因子方差。

(1)载荷系数的估计。因子载荷系数及显著性检验如表 6-14 所示。表格前两列显示 了 3 个因子与题项的测量关系,第 3~6 列为非标准化回归系数(也可以叫作非标准载荷 系数、非标准路径系数)的估计值与显著性检验,第 7 列为标准化回归系数也叫作因子载 荷系数。

Factor	测量项	非标准载荷系	标准误	- (CD 店)		因子载荷系数
(潜变量)	(显变量)	数(Coef.)	(Std. Error)	Z(GR 徂)	p	(Std. Estimate)
知识创新	A1	1.000	—	—	—	0.634
知识创新	A2	1.402	0.145	9.673	0.000	0.808
知识创新	A3	1.680	0.167	10.077	0.000	0.970
知识创新	A4	0.925	0.123	7.527	0.000	0.592
知识分享	B1	1.000	—	—	—	0.592
知识分享	B2	1.076	0.131	8.189	0.000	0.689
知识分享	B3	1.564	0.152	10.289	0.000	0.991
知识分享	B4	1.324	0.159	8.339	0.000	0.707
知识分享	B5	1.616	0.157	10.287	0.000	0.991
知识获得	C1	1.000	—	—	—	0.680
知识获得	C2	1.347	0.107	12.607	0.000	0.975
知识获得	C3	0.932	0.106	8.836	0.000	0.655
知识获得	C4	1.293	0.104	12.416	0.000	0.956
知识获得	C5	1.271	0.105	12.074	0.000	0.925

表 6-14 因子载荷系数及显著性检验

为了让模型识别验证性因子分析(包括结构方程模型),一般默认将因子与第一个题项的未标准化回归路径系数固定为 1,不需要检验显著性。例如,本例"知识创新→A1" "知识分享→B1""知识获得→C1"的非标准载荷系数统一被固定为 1.0000。

重点解读的结果是,第6列回归系数显著性检验的p值及第7列因子载荷系数。若回 归系数显著(p<0.05),则说明因子对题项的解释有统计学意义。本例3个因子对所有14 147/187 个题项的回归系数均小于 0.05, 说明非标准载荷系数的参数具有显著性。14 个题项的因子 载荷系数介于 0.592~0.991, 均大于 0.5, 多数在 0.7 以上, 说明预设的题项可以测量潜变 量。

非标准估计系数 标准误 标准估计系数 项 z р (Coef.) (Std. Error) (Std. Estimate) A1 0.254 0.027 9.411 0.000 0.598 A2 0.178 0.024 7.427 0.000 0.347 ÷ ÷ ÷ ÷ ÷ ÷ 0.016 0.008 2.088 0.037 0.018 B3 ÷ ÷ ÷ ÷ ÷ ÷ C5 0.110 0.013 8.244 0.000 0.144 知识创新 0.000 0.171 0.035 4.843 1.000 知识分享 0.361 0.079 4.596 0.000 1.000 知识获得 0.405 0.075 5.429 0.000 1.000

表 6-15 测量残差与回归残差(篇幅有限仅展示部分结果)

(2)测量残差的估计。本次验证性因子分析的测量残差与回归残差如表 6-15 所示。 表格较大,此处仅展示其中部分结果。一般要求残差估计值的标准误为正数且达到显著水 平(p<0.05),本例中表格的第 3 列标准误均为正数且 p 值均小于 0.05,参数具有显著性。</p>

(3)因子间相关系数的估计。维度间的协方差及相关系数如表 6-16 所示。第 3 列 "非标准估计系数"即维度间的协方差,最后一列"标准估计系数"则为维度间的相关系数。如 果显著性 p 值小于 0.05,则说明相关关系具有统计学意义。

表 6-16 维度间的协方差及相关系数

Factor	Factor	非标准估计系数	标准误	7	n	标准估计系数
Factor	Factor	(Coef.)	(Std. Error)	2	ρ	(Std. Estimate)
知识创新	知识分享	0.049	0.020	2.498	0.012	0.197
知识创新	知识获得	0.056	0.021	2.670	0.008	0.212
知识分享	知识获得	0.152	0.035	4.339	0.000	0.396

本例"知识创新""知识分享""知识获得"3个维度间的标准估计系数介于 0.197~ 0.396, 相关性相对较弱,因此不考虑二阶验证性因子分析。

综上,本例载荷系数、测量残差、因子间相关系数等参数均有统计学意义,说明模型的内在质量良好。

3) 模型拟合评价

接下来需要对模型拟合质量进行总体评价,主要针对一系列常用的拟合指标的达标情况进行解释及分析。表 6-17 所示为 SPSSAU 平台整合后的模型拟合指标。本例模型的自由度 df=74.000>0,说明模型可识别,和前面通过"t 法则"计算的判断结果一致。

常用 指标	χ ²	df	p	卡方自由 度比	GFI	RMSEA	RMR	CFI	NFI	NNFI
判断 标准	_	_	>0.05	3	>0.9	<0.10	<0.05	>0.9	>0.9	>0.9
值	246.290	74.000	0.000	3.328	0.839	0.108	0.094	0.933	0.907	0.917
其他 指标	TLI	AGFI	IFI	PGFI	PNFI	SRMR	RMSEA 90% CI			
判断 标准	>0.9	>0.9	>0.9	>0.5	>0.5	<0.1	—			
值	0.917	0.771	0.933	0.591	0.738	0.104	0.093~0.123			

表 6-17 SPSSAU 平台整合后的模型拟合指标

注: Default Model: $\chi^2(91)=2656.033$, $p=1.000_{\circ}$

表 6-17 中的结果分为"常用指标"和"其他指标"两部分结果,本书推荐优先解释及分析 "常用指标"是否达标。模型适配的卡方检验 p<0.05,拒绝原假设,表明模型不适配。卡方 自由度比为 3.328>3, RMSEA=0.108>0.1、RMR=0.094>0.05、GFI=0.839<0.9,这5 项 常用指标均未达标。其他指标中的 AGFI=0.771 也未达标。综合以上指标,认为模型适配 欠佳。

4) 模型 MI 修正

前面经过模型拟合评价认为模型适配欠佳,可通过 MI 修正项的提示进行适当的修正。 修正时应注意结合理论支持和模型规则,不能为了"修正"而修正。

(1) MI 指标之因子-题项的对应关系。因子与题项的 MI 如表 6-18 所示。第 4 列 "MI 值"为修正指数, 第 5 列"Par Change"指修正参数后期望参数的改变量。

测量项	关系	因子	MI 值	Par Change
C1	测量	知识分享	28.653	0.481
C3	测量	知识分享	12.419	0.315
B1	测量	知识获得	13.877	0.376
B2	测量	知识获得	5.348	0.194
B4	测量	知识获得	5.726	0.235

表 6-18 因子与题项的 MI

注:表格中的 MI 值均大于 5。

表 6-18 中最大的 MI 值为 28.653, Par Change 为 0.481, 意味着如果我们创建"知识 分享→C1"的影响关系则可使卡方统计量减少 28.653, 但是本例认为该修正项目是不合适 的(C1 题项是为测量"知识获得"而设计的, 在内容上与"知识获得"同质), 同理按 MI 值降 序进行判断, 在表 6-18 中没有发现适合的修正项目。

(2) MI 指标之题项间的协方差关系。题项间的 MI(部分结果) 如表 6-19 所示。由于该表格较大,我们将完整表格复制粘贴到 Excel 按 MI 从大到小降序排列,截取前 5 个 MI 值。

测量项 1	关系	测量项 2	MI 值	Par Change
B5	\leftrightarrow	B3	74.617	0.364
C5	\leftrightarrow	C4	32.374	0.077
C3	\leftrightarrow	C1	22.728	0.162
B2	\leftrightarrow	B1	20.635	0.180
C4	\leftrightarrow	C1	19.850	-0.068

表 6-19 题项间的 MI(部分结果)

注: 表格中的 MI 值均大于 5。

最大的 MI 值为 74.617, 提示创建 B3 题项和 B5 题项两个题项间的协方差共变关系有 助于降低卡方值,同一个因子下的题项反映的是同一个主题,因此存在相关性,所以构建 B3 题项与 B5 题项的相关关系在理论上是合适的,该表的其他 MI 值均以此方式去解读。 研究者可根据模型拟合、MI 修正综合决定是否有必要进行修正,修正过程是逐次进行的, 不能一次对多个项目进行修正。例如,按 MI 值 74.617 提示进行修正,是指重新进行一次 验证性因子分析,并且要建立 B3 和 B5 的协方差参数估计。

A1	÷	因子名称【可选】				
31	* *	知识创新			知识分享	
01	<u>^</u>	知识获得			Factor4	
		Factor5			Factor6	
Factor4(證表說)		Factor7			Factor8	
Factor5(量表题)		Factor9			Factor10	
Factor6(量表證)		Factor11			Factor12	
Factor7(量表题)			因子名称]	
Factor8(圖表題)		测量项协方差关系	【结合MI指标设置	1		
		第1项		1	第2项	

图 6-17 验证性因子分析测量残差修正的操作界面

(3) 按 MI 进行 B3↔B5 的协方差修正。从仪表盘中单击【验证性因子分析】模块,其他操作设置和前面的设置一致。在【测量项协方差关系【结合 MI 指标设置】】的下方, 【第 1 项】下拉列表中选择【B3】,【第 2 项】下拉列表中选择【B5】,其他设置不变,最 后单击【开始分析】按钮。验证性因子分析测量残差修正的操作界面如图 6-17 所示。

SPSSAU 平台会专为新增的 B3↔B5 两个题项的协方差关系输出对应的结果。题项间的 协方差及相关系数如表 6-20 所示。

表 6-20 题项间的协方差及相关系数

显变量	显变量	非标准估计系数 (Coef.)	标准误(Std. Error)	z	p	标准估计系数 (Std. Estimate)
B3	B5	0.258	0.040	6.394	0.000	0.277

协方差显著性检验 p 值小于 0.05, 即协方差有统计学意义, B3 与 B5 之间的相关性是显著的, 具体的相关系数(标准估计系数)为 0.277。

其他的输出结果和第一次验证性因子分析相同,我们主要来看在原模型基础上增加 B3 与 B5 的协方差关系后,模型的适配有无改善,其他表格均以第 2 次修正后的结果为准进 行解读,解读过程可参考首次分析的内容。

常用 指标	χ ²	df	р	卡方自由 度比 χ ² /df	GFI	RMSEA	RMR	CFI	NFI	NNFI
判断 标准	_		>0.05	<3	>0.9	<0.10	<0.05	>0.9	>0.9	>0.9
值	185.076	73.000	0.000	2.535	0.876	0.088	0.069	0.956	0.930	0.946
其他 指标	TLI	AGFI	IFI	PGFI	PNFI	SRMR	RMSEA 90% CI			
判断 标准	>0.9	>0.9	>0.9	>0.5	>0.5	<0.1	_			
值	0.946	0.821	0.957	0.609	0.746	0.078	0.072~0.103			

表 6-21 模型拟合指标达标情况(修正后)

注: Default Model: $\chi^2(91)=2656.033$, $p=1.000_{\circ}$

卡方自由度比由原来的 3.328 下降到 2.535,小于 3 达标; RMSEA 由 0.108 下降到 0.088,小于 0.1 达标; GFI 由 0.839 提高到 0.876 接近 0.9,基本达标; SRMR 由 0.104 下 150 / 187

降到 0.078, 小于 0.1 达标; AGFI 由 0.771 提高到 0.821 接近 0.9; CFI、NFI、NNFI、IFI 这 4 个原本在 0.9 以上的指标修正后指标值仍然有所提升。

增加 B3 与 B5 的协方差关系后有 9 个以上的拟合指标达标或基本达标,修正后的模型 适配基本达到拟合要求,各参数估计值达到显著性水平,测量误差均为正数且显著。综合 认为,预设的模型结果得到了数据验证,模型与数据是适配的。

5) 分析与应用

接下来以修正的第二次验证性因子分析结果为最终模型,分析本例量表的聚合效度和 区分效度。SPSSAU 平台计算并提供了 AVE 值、CR 值,以及 AVE 平方根与因子相关系数 的比较结果。

(1) AVE 与 CR。本例 AVE 与 CR 指标如表 6-22 所示。

Factor	AVE 值	CR 值
知识创新	0.587	0.845
知识分享	0.647	0.901
知识获得	0.723	0.927

表 6-22 AVE 与 CR 指标

"知识创新""知识分享""知识获得"3 个因子的 CR 值均大于 0.8, 按 0.6 标准说明各因子的信度水平较高,量表内在质量良好。3 个因子的 AVE 值均大于 0.5, 按 0.5 标准说明因子内题项的同质性高,结合 CR 值与 AVE 值认为本例量表具有良好的聚合效度。

(2) 区分效度。一个良好的量表结构应该同时有良好的聚合效度和区分效度。一般用 因子的 AVE 平方根与因子相关系数进行比较,若 AVE 平方根大于该因子与其他因子两两 间的相关系数,则认为具有良好的区分效度。

模型区分效度如表 6-23 所示。对角线上加粗的数字即为 3 个因子的 AVE 平方根值, 其他黑色数字为因子间的 Pearson 相关系数。例如,本例"知识创新"的 AVE 平方根为 0.766,它大于同一列的 0.211 ("知识创新"与"知识分享"的 Pearson 相关系数)和 0.217

("知识创新"与"知识获得"的 Pearson 相关系数);"知识分享"的 AVE 平方根为 0.804,它大于同一列的 0.547 ("知识分享"与"知识获得"的 Pearson 相关系数),并且大于同一行的 0.211;"知识获得"的 AVE 平方根为 0.850,大于同一行的 0.217 和 0.547。所以,本例量表具有良好的区分效度。

项	知识创新	知识分享	知识获得
知识创新	0.766		
知识分享	0.211	0.804	
知识获得	0.217	0.547	0.850

表 6-23 模型区分效度

注:斜对角线加粗数字为 AVE 平方根值。

综合以上结果解释与分析, 在本例 14 个题项中测量 3 个因子的测量模型和数据适配, 参数估计值均具有显著性, 载荷系数均大于 0.5, 多数大于 0.7, 3 个因子的 CR 值均大于 0.7, 3 个因子的 AVE 值均大于 0.5, 且 AVE 平方根大于因子间相关系数, 总体上认为测量 模型适配且结构效度良好。

6.7路径分析

线性回归被广泛用于多个自变量对一个因变量的预测,在实际分析中存在许多更加复 151/187 杂的变量关系,如多个自变量同时对多个因变量的影响,或一个变量有两种身份既是自变 量又是因变量。这种同时存在多个因变量的变量关系研究,可使用路径分析。本节介绍路 径分析的应用,通过在仪表盘中依次单击【问卷研究】→【路径分析】模块来实现。

6.7.1 方法概述

1.基本概念

路径分析(Path Analysis),也称作通径分析,可通过构建路径图来研究变量间的影响 关系,验证模型假设。各研究变量按照某种理论假设通过单箭头或双箭头进行连接,构成 路径图,单箭头所指的方向即两个变量的影响路径,双箭头则表示两个变量相关。通过估 计路径系数(标准化回归系数),分析路径系数的正负方向及显著性,从而验证理论假设, 分析变量的结构关系。

路径分析有两种, 第一种是显变量路径分析, 即路径图中的各变量均为可观测变量, 如果是量表问卷数据, 则一般为各维度或量表题项加总的数据; 第二种是潜变量路径分析,



即路径图中各变量为潜变量,此时同时存在测量模型和路径分析模型,共同构成结构方程 模型,结构方程模型的内容在 11.8 节进行详细介绍。

图 6-18 路径分析示意图

图 6-18 所示为路径分析示意图,有4个研究变量,5个单箭头代表了5对变量的关系(假设箭头均为正向影响关系):A 正向影响 D(A→D);A 正向影响 B(A→B);C 正向影 响 B(C→B);C 正向影响 D(C→D);B 正向影响 D(B→D),以上这5种路径称为直接影响路径。此外还包括 A 通过先影响 B 再影响 D(A→B→D),以及 C 通过先影响 B 再影响 D(C→B→D)两条间接影响路径。

2.效应分解

吴明隆(2010.10)指出,在路径分析中,变量间的影响效应(Effects)包括直接效应 (Direct Effect)与间接效应(Indirect Effect),两种效应的总量称为外生变量对内生变量影 响的总效应(Total Effects)值。

通常把直接影响关系的标准化回归系数称为路径系数,也就是直接效应值,图 6-18 中用 a1~a4 与 b 进行标注,如 a1 是 A→B 的直接效应值,b 是 B→D 的直接效应值。间 接效应值被定义为间接路径前后两个或多个直接效应值的乘积,如 A 先通过 B 再影响 D(A→B→ D)的间接效应值,等于 a1×b; C 先通过 B 再影响 D(C→B→D)的间接效应值,等 于 a3×b。作用到同一个因变量上的直接效应与间接效应的和称为总效应,如 A→D 的总效 应=A→D 的直接效应+A→B→D 的间接效应。

3.内生变量与外生变量

路径分析强调变量间的关系事先要有理论支撑,单箭头方向代表变量间的影响关系, 箭头左侧为自变量,箭头右侧为因变量。在实际中一般不使用自变量和因变量的称呼,取 而代之的术语是外生变量和内生变量。简单来说,凡是被箭头指向的变量称为内生变量 (因变量)、没有被箭头指向的变量则称为外生变量(自变量),在路径分析中一个变量既可能是外生变量也可能是内生变量。图 6-18 所示的路径关系中,A→B,B 被箭头所指所以是内生变量,但在 B→D 关系中 B 作为外生变量,此时 B 既是内生变量又是外生变量。

4.路径分析步骤

实际路径分析时主要包括以下 5 个步骤。

(1) 模型设定。一般根据某种理论或借鉴成熟模型的框架结构,结合当前研究主题的 既往研究成果、专业知识,对要研究的变量进行定义,设定研究变量间的影响关系,并将 这种结构关系绘制成路径图。应注意,本节介绍的路径分析特指观测变量数据资料的情况, 所以路径图中的矩形框表示观测变量,其他元素,如单箭头表示影响关系,双箭头表示相 关关系。

(2)路径系数估计与检验。路径图中路径系数的估计方法常用的有两种,第一种是传统多元线性回归,如传统上使用 SPSS 软件通过线性回归进行路径分析;第二种是最大似然估计,结构方程模型软件工具常采用该方法进行路径分析。SPSSAU 平台的【路径分析】 模块默认使用的是最大似然估计,按路径图上的影响关系拟合回归方程,有几个内生变量 即可拟合几个回归方程。一般用标准化回归系数作为路径图中箭头方向的路径系数,根据 显著性 p 值判断路径系数是否有统计学意义 (p<0.05 时显著,即影响关系成立),对于无统计学意义的路径可根据研究目的或理论依据考虑删除。

(3) 模型拟合评价。对路径分析模型的整体拟合评价,可从回归分析 R2 及卡方自由 度比、RMSEA、RMR、GFI、CFI、NFI 和 NNFI 等指标入手,可参考前面验证性因子分析模 型拟合评价介绍的内容。

(4)模型修正(非必须)。模型拟合较差时可结合 MI 进行必要、适当的修正。从以下两个方面考虑修正,一方面要结合"回归影响关系"的 MI 指标及理论专业知识,重新调整模型中变量的影响关系;另一方面要根据"模型协方差"的 MI 指标,创建外生变量间的协方差关系。

修正时应注意以理论和实际意义为导向,而不是以数据结果为导向。

(5)结果分析报告。对修正后(如果有修正的话)的路径分析结果进行综合分析和报告,具体来说,主要包括验证前面的影响关系、研究假设是否成立。

6.7.2 实例分析

下面以一个量表问卷研究为例介绍路径分析的应用。

【例 11-10】为提高顾客满意度和忠诚度,研究者设计了一份售后服务满意度量表问卷,共有 22 个李克特五级量表题,用于测量"服务态度""促销活动""个性服务""满意度" "忠 诚度"5 个变量,试根据理论模型进行路径分析。问卷数据来源于李金林和马宝龙 (2007),但对问卷描述和数据均有修改及编辑,本例数据文档见"例 11-10.xls"。1)模型设定

在本例数据中,"服务态度""促销活动""个性服务""满意度""忠诚度"5 个变量的数据即各 维度题项加总的和,属于显变量。根据相关理论和文献认为"服务态度""个性服务""促销活 动"会直接正向影响"忠诚度"。此外,"服务态度"和"个性服务"又会通过"满意度"间接影响 "忠诚度"。5 个变量间结构关系的路径图如图 6-19 所示。



153 / 187

图 6-19 5 个变量间结构关系的路径图

由路径图可知,"满意度"既是内生变量又是外生变量。整个模型包括两个回归方程, 第一个是以"满意度"作为内生变量,"服务态度""个性服务"作为外生变量;第二个是以"忠诚 度"作为内生变量,其他 4 个变量作为外生变量。两个回归方程中包括以下 6 个直接影响关 系(路径)。

- (1) 服务态度→忠诚度。
- (2) 个性服务→忠诚度。
- (3) 促销活动→忠诚度。
- (4) 满意度→忠诚度。
- (5) 服务态度→满意度。
- (6) 个性服务→满意度。

将数据文档"例 11-10.xls"导入平台后,在仪表盘中依次单击【问卷研究】→【路径分析】模块,通过【自变量 X】下拉列表、【因变量 Y】下拉列表及中间的单箭头或双箭头将路径图的 6 个直接影响关系输入平台。此处应注意,【影响→】为单箭头,表示影响关系,包括正向影响和负向影响;【相关→】为双箭头,表示相关关系,通常来说默认用【影响→】即可。

例如,服务态度→忠诚度直接影响路径,先在【自变量 X】下拉列表中选择【服务态 度】,中间箭头默认用【影响→】,然后在【因变量 Y】下拉列表中选择【忠诚度】。设置完 一组关系后,单击【因变量 Y】下拉列表右侧的【+】按钮,可增加并设定更多关系。如 果设定了多余的关系,则单击【-】按钮删去。本例的其他 5 个直接影响关系均是同样的 操作。路径分析操作界面如图 6-20 所示。

影响→ ¥ 影响→ ¥	忠诚度	~	+	-
影响→V	忠诚度			
			+	-
影响→ ∨	忠诚度	~	+	-
影响→V	忠诚度	~	+	-
影响→∨	满意度	~	+	-
影响→ ✔	满意度	~	+	-
	影响→ × 影响→ × 影响→ ×	影响→ 思诚度 影响→ 思诚度 影响→ 满意度 影响→ 满意度	影响→ 思城度 影响→ 思城度 影响→ 满意皮 影响→ 满意皮	影响→ 思诚度 + 影响→ 思诚度 + 影响→ 满意度 + 影响→ 满意度 +

图 6-20 路径分析操作界面

在【MI 指标】下拉列表中选择【输出 MI>5】,即命令输出 MI 指标的结果,以便我们 根据模型拟合的情况按 MI 指标的提示进行适当或必要的修正。下一步勾选界面最底部的 【外生变量自动协方差关系】复选框,使平台自动对路径图中的外生变量建立协方差(相 154/187 关)关系,该操作类似于前面的【相关↔】,前者由平台自动完成,后者由用户指定。本例 的外生变量包括"服务态度""个性服务""促销活动",它们之间可能存在相关性,因此要勾选 【外生变量自动协方差关系】复选框,最后单击【开始分析】按钮。

2) 路径系数估计与检验

根据前面的模型设定,本例路径分析需要估计各路径系数、协方差/相关系数,以及残差,并做参数的显著性检验。

(1) 单箭头路径系数的估计与检验。在路径分析中,我们较为关心的是模型路径系数 及显著性检验,如表 6-24 所示。路径分析的原理是回归分析,回归系数显著性检验 p 值

作为路径系数有无统计学意义的判断依据,通常采用标准化回归系数表示两个变量间的影 响关系,此处直接称为标准化路径系数。

Х	→	Y	非标准化路径系数	非标准化路径系数 SE z(CR 值) p 林		标准化路径系数	
服务态度	+	忠诚度	0.072	0.058	1.233	0.218	0.121
个性服务	+	忠诚度	0.131	0.063	2.076	0.038	0.211
促销活动	+	忠诚度	0.076	0.033	2.301	0.021	0.164
满意度	+	忠诚度	0.433	0.126	3.425	0.001	0.433
服务态度	+	满意度	0.296	0.036	8.264	0.000	0.498
个性服务	→	满意度	0.301	0.037	8.058	0.000	0.486

表 6-24 模型路径系数及显著性检验

注: →表示路径影响关系。

本例中"服务态度"对"忠诚度"的影响没有统计学意义(p=0.218>0.05),模型设定时提出的关系假设是不成立的。其他 5 组直接影响关系均有统计学意义(p 值均小于 0.05),而且非标准化路径系数均为正数,所以这 5 组直接影响关系的假设成立。

表 6-24 中的标准化路径系数为 6 条路径的直接影响效应,如"个性服务"对"满意度"的 直接效应值为 0.486,"满意度"对"忠诚度"的直接效应值为 0.433。

本例还存在"服务态度"通过"满意度"影响"忠诚度","个性服务"通过"满意度"影响"忠诚 度"这两条间接效应。例如,"个性服务→满意度→忠诚度",间接效应值为"个性服务→满意 度"效应值 0.486 与"满意度→忠诚度"效应值 0.433 的乘积,即 0.486× 0.433=0.210。同理, "服务态度→满意度→忠诚度"的间接效应值为 0.498×0.433=0.215。

(2)协方差/相关系数的估计与检验。3 个外生变量间的协方差及相关系数结果如表 6-25 所示。

表 6-25 3 个外生变量间的协方差及相关系数结果

_	х	Y	非标准估计系数 (Coef.)	标准误 (Std. Error)	z	p	标准估计系数 (Std. Estimate)
_	服务态度	促销活动	7.025	1.927	3.645	0.000	0.398
_	个性服务	促销活动	8.611	1.924	4.475	0.000	0.510
-	个性服务	服务态度	7.735	1.553	4.982	0.000	0.586

3 个外生变量之间的协方差均具有统计学意义(p<0.05),表明两两变量之间存在相关性。"服务态度"与"促销活动"、"个性服务"与"促销活动"、"个性服务"与"服务态度"间的相关系数分别为 0.398、0.510、0.586,可见 3 个外生变量间存在中等程度的相关。

3) 模型拟合评价

针对路径分析模型拟合的评价, SPSSAU 平台能提供回归模型拟合度 R2 及路径分析模型总体评价指标(卡方自由度比、RMSEA、RMR、GFI、CFI、NFI 和 NNFI 等, 参考前面验证性因子分析的解释和分析结果)。

(1)模型拟合度 R2。分别以"满意度""忠诚度"为因变量的回归拟合优度如表 6-26 所示, 给出了以"忠诚度""满意度"为因变量的两个回归方程 R2,代表各自的回归方程可解释因变 量变异的百分比。本例分别为 64.2%、76.9%,相对来说模型的解释能力良好。

项	R方值
忠诚度	64.2%
满意度	76.9%

表 6-26 分别以满意度、忠诚度为因变量的回归拟合优度

路径分析主要用于研究变量间的影响关系有无统计学意义,此处 R2 的意义较小,该结果了解一下即可。

(2) 模型总体拟合指标。路径分析和验证性因子分析相同,可使用卡方自由度比、 RMSEA、RMR、GFI、CFI、NFI和 NNFI 等指标对模型拟合质量进行综合评价。

路径分析模型拟合指标如表 6-27 所示。本例卡方检验 p=0.187>0.05、卡方自由度比 =1.744<3、RMSEA=0.088<0.1、SRMR=0.016<0.1、GFI、AGFI、CFI、NFI、TLI、IFI 均大于 0.9,绝大多数指标是达标的,综合认为模型拟合良好,理论设定的模型与实际样本数据适 配。

常用指标 卡方自由度比 χ²/df GFI RMSEA RMR CFI NFI NNFI χ^2 df р 判断标准 >0.05 <3 >0.9 < 0.10 < 0.05 >0.9 >0.9 >0.9 ____ ____ 值 1.744 1 0.187 1.744 0.998 0.088 0.162 0.998 0.994 0.976 其他指标 TLI AGFI IFI PGFI PNFI SRMR RMSEA 90% CI 判断标准 >0.9 >0.5 < 0.1 >0.9>0.9>0.5值 0.976 0.966 0.998 0.067 0.099 0.016 $0.078 \sim 0.301$

表 6-27 路径分析模型拟合指标

注: Default Model: $\chi^2(10)=316.176$, $p=1.000_{\circ}$

如果模型总体上拟合欠佳或前面有路径系数不显著,则考虑将不显著的路径系数删除, 并依据 MI 指标进行适当修正。本例模型适配良好,暂不考虑删除不显著的路径系数。

4) 模型 MI 修正(非必须)

模型协方差 MI 修正指标、回归影响关系 MI 修正指标如表 6-28 和表 6-29 所示。由于 MI 值全部小于 5,因此没有提供可修正的影响关系。总体模型拟合达标,本例模型无须 修正。

表 6-28 模型协方差 MI 修正指标

项	关系	项 MI 值		Par Change
注: MI	[值全部小于5 或其他	也原因,因	而无输出。	

表 6-29 回归影响关系 MI 修正指标

项	项 关系 项		MI 值	Par Change	
注: MI	值全部小于5或其6	也原因,因	而无输出。		

5) 结果分析与报告

综合以上结果认为模型拟合良好,预设模型与实际样本数据适配。模型假设的6条直

接影响路径中有 5 条成立, 且均为正向影响关系。把 6 个直接效应值标记到路径图上, 本 例的路径模型结果, 如图 6-21 所示。



图 6-21 本例的路径模型结果

根据吴明隆(2010.10),总效应可分解为直接效应加间接效应,直接效应即标准化路 径系数,而间接效应为间接路径前后两个或多个直接效应值的乘积。对本例路径模型中各 项效应值进行汇总,如表 6-30 所示。

影响关系	直接效应值	间接效应值	总效应值
服务态度→忠诚度	0.121	0.215	0.336
个性服务→忠诚度	0.211	0.210	0.421
促销活动→忠诚度	0.164	/	0.164
满意度→忠诚度	0.433	/	0.433
服务态度→满意度	0.498	/	0.498
个性服务→满意度	0.486	/	0.486

表 6-30 例 11-10 直接效应值、间接效应值、总效应值汇总表

6.8结构方程模型

前面已经介绍了验证性因子分析与路径分析,前者是测量模型,用具体可观测的题项 来测量潜变量因子;后者是结构模型,用于研究变量间的关系,变量可以是显变量也可以 是潜变量。在统计学上,测量模型与结构模型合在一起能构成结构方程模型。本节主要介 绍结构方程模型的应用,通过在仪表盘中依次单击【问卷研究】→【结构方程模型 SEM】 模块来实现。

6.8.1 方法概述

1.基本概念

结构方程模型(Structural Equation Model, SEM)主要用于分析及研究潜变量之间的 关系,由于潜变量不能直接观测,在问卷中常用一系列量表题进行测量,因此结构方程模 型由两个模型构成,即测量模型和结构模型,其中测量模型即前面介绍的验证性因子分析, 用于测量概念、评价结构效度并考察测量误差;结构模型即潜变量间的路径分析,用来分 析潜变量之间的影响关系。



图 6-22 结构方程模型示意图

相较于路径分析,结构方程模型突出的优势在于先考虑测量误差再研究潜变量之间的 关系,在普遍存在潜变量的心理、教育等领域得到了广泛应用。

图 6-22 所示为结构方程模型示意图。假设潜变量 A 为因子 Factor1、潜变量 B 为因子 Factor2、潜变量 C 为因子 Factor3、潜变量 D 为因子 Factor4, Factor1~Factor4 这 4 个潜 变量因子是受访者的心理感受或态度,无法直接观测到数据,所以要设计一定数目的量表 题进行测量。用 A1~A5 题项测量 Factor1、B1~B6 题项测量 Factor2、Y1~Y3 题项测量 Factor3、Z1~Z3 题项测量 Factor4,这 4 个测量过程就是 4 个测量模型。而 Factor1 影响 Factor3(A→C)、Factor2 影响 Factor3(B→C)、Factor3 影响 Factor4(C→D),这 4 个潜变量 因子之间的影响关系为结构模型,先用具体的题项测量潜变量,再研究潜变量间的影响关 系,这就是结构方程模型。

2.结构方程模型路径图

路径图是结构方程模型的重要图示,能清晰地表达观测题项与潜变量、潜变量之间的 关系。在路径图中用椭圆表示潜变量,如图中的 A、B、C、D;用矩形框表示显变量,如 图中的 A1~A5;小圆为误差项,它包括测量误差,如图中的 e1~e17,以及回归误差,如 图中的 e18 和 e19。

单箭头表示变量间假设的影响关系,箭头由外生变量(自变量)指向内生变量(因变量)。外生变量之间可创建双箭头,表示潜变量间的协方差或相关性关系。在结构方程模型 修正时,也可以创建测量误差间的双箭头,表示误差间的协方差或相关性关系。

3.结构方程模型分析步骤

结构方程模型从原理上的分析步骤和验证性因子分析基本一致,严格来说,验证性因子分析本就是结构方程模型的特例。图 6-23 所示为结构方程模型的分析步骤,主要包括 5 个环节。



图 6-23 结构方程模型的分析步骤

1) 模型设定与识别

模型要研究哪些变量、变量间的影响关系,以及每个变量由哪些题项测量,这些都属

于模型设定的部分,变量间的关系需要有一定的理论依据或得到既往研究中某个研究成果 的支持。在结构方程模型中,因子的个数是确定的,结构方程模型路径图中的每个箭头都 可以提出一个具体的研究假设。

在模型识别上和前面介绍的验证性因子分析类似,要求模型能够计算出唯一解。有两 个必要条件,第一个是观测变量(在量表数据中即题项)的方差、协方差矩阵元素数不能 少于自由参数的个数;第二个是每个潜变量都应该指定任意一个测量题项的路径系数为 1。

2) 模型参数估计与检验

在结构方程模型中,我们关注的参数既包括验证性因子分析中的载荷系数、测量残差 及因子间相关系数,又包括潜变量因子间影响关系的路径系数。在 SPSSAU 平台中,结构 方程模型参数估计的方法默认使用最大似然估计,它要求数据为连续型定量数据资料,且 满足多元正态分布。载荷系数、测量残差、相关系数显著性检验与验证性因子分析模型一 致,潜变量因子间的路径系数显著性检验,则主要依据模型理论支持和研究假设来判断是 否满足要求或是否能验证提出的研究假设。

例如,研究假设认为因子 F1 正向影响因子 F2,如果该路径系数估计值具有显著性 (p<0.05),则说明该研究假设成立,相反如果没有显著性,则说明该研究假设不成立。

3) 模型拟合评价

结构方程模型样本数据是否拟合适配,有多种拟合指标进行评价,与前面介绍的验证 性因子分析模型拟合评价时的内容一致,包括卡方自由度比、RMR、SRMR、RMSEA、 GFI/AGFI 等绝对适配指标、NFI、RFI、IFI、TLI、CFI 这 5 个增值适配指标,以及 AIC/BIC、 PNFI、PGFI 等简约适配指标。有关内容见 11.6 节,对模型拟合的评价,应综合多个指标 的达标情况而定,一般不能由单个指标决定模型拟合的质量。

张伟豪等人(2020)认为,结构方程模型适配度检验是对模型的整体拟合进行检验, 模型中可能某些部分较差,某些部分较好,所以需要有其他不同性质的拟合指标补充说明。 由于没有一个或一组指标是公认最好的,因此最好的做法是从绝对适配指标、增值适配指 标及简约适配指标这3种类型指标中各选一个或两个作为代表性指标。

4) 模型 MI 修正(非必须)

结构方程模型如果和样本数据拟合欠佳,可考虑利用 MI 的提示对模型进行修正。模型修正意味着要调整模型设定,所以此前的步骤要重复一次。另外,模型修正一定要结合理论和研究成果的支持,不能完全按照 MI 进行修正,否则可能会破坏模型理论的基本假设,得不偿失。修正过程是逐次进行的,不能一次对多个项目进行修正。

5) 分析与报告

综合评价模型拟合情况, 解释和分析模型参数估计及显著性检验, 对模型假设进行验证, 对分析结果进行总结和报告。

6.8.2 实例分析

接下来使用某研究参考成熟量表编制的"顾客满意度量表",通过实例介绍结构方程模型的应用。

【例 11-11】某研究参考成熟量表编制的"顾客满意度量表", 共 12 个李克特五级量表 题目, 收集了 201 份问卷。根据相关研究理论和既往研究结论初设了 4 个维度: A1~A4 题项测量"感知质量", B1~B3 题项测量"感知价值", C1~C3 题项测量"顾客满意", D1~ D2 题项测量"顾客忠诚"。试根据理论模型进行结构方程模型分析。本例数据文档见 "例 11-11.xls", 案例和数据均为模拟获得, 仅用于方法操作和解释, 其分析结果不代表真实研 究结论。



图 6-24 本例结构方程模型理论模型路径图

1) 模型设定与识别

以下理论与假设仅作为结构方程模型的案例示范实现结构方程模型的过程。根据相关 理论、文献资料及研究目的,提出以下假设:"感知质量(Factor1)"正向影响"顾客满意

(Factor3) ""感知质量(Factor1)"正向影响"感知价值(Factor2)""感知价值(Factor2)" 正向影响"顾客满意(Factor3)""顾客满意(Factor3)"正向影响"顾客忠诚(Factor4)"。 我们要研究的变量有 4 个, A1 ~ A4 题项测量 Factor1, B1 ~ B3 题项测量 Factor2, C1 ~ C3 题项测量 Factor3, D1 ~ D2 题项测量 Factor4。本例结构方程模型理论模型路径图如图 6-24 所示。

由模型路径图可知,本例模型中变量间的影响关系包括: Factor1→Factor3、 Factor1→ Factor2、Factor2→Factor3、Factor3→Factor4。模型自由参数个数少于 12 个题 项的协方差矩阵元素数,模型可识别。

将数据文档"例 11-11.xls"导入平台, 在仪表盘中依次单击【问卷研究】→【结构方程 模型 SEM】模块。

首先从标题框中将【A1】~【A4】题项拖曳至右侧【Factor1(量表题)】分析框中,同 理将【B1】~【B3】题项拖曳至【Factor2(量表题)】分析框中、【C1】~【C3】题项拖曳 至【Factor3(量表题)】分析框中、【D1】~【D2】题项拖曳至【Factor4(量表题)】分析框中, 设定好测量模型。

接下来在【设置模型关系】窗格,设定 4 个因子变量间的影响关系。在【第 1 项】下 拉列表中选择【Factor1】,中间的关系类型下拉列表选择【影响→】,在【第 2 项】下拉列 表中选择【Factor3】,即表示 Factor1→Factor3。单击【+】按钮能增加更多关系设定。同 样的操作,设定 Factor1→Factor2、Factor2→Factor3、Factor3→Factor4 的影响关系。在 其他相对复杂的结构方程模型中,如果有多个外生变量,一般建议设定外生变量间的相关 关系,只需要在中间的关系类型下拉列表中选择【相关→】即可。结构方程模型操作界面 如图 6-25 所示。

SPSSAU 科研数据分析方法与应用

开始分析 💼 輸出M	II>5 ~ Q
A1	设置模型关系
B1	第1项 第2项
C1	Factor1 マ Factor3 マ + -
••••••••••••••••••••••••••••••••••••••	Factor1 マ Factor2 マ + -
D1	Factor2 ~影响→~ Factor3 ~ + -
Factor5(量表题)	Factor3
Factor6(量表證)	影响→ 相关↔
Factor7(最表踐)	设置量表二阶结构【高级设置,可选】 🌑
Factor8(昼表题)	量表Scale1 请选择 ▼
Factor9(显表题)	at at a state of the state of
Factor10(量表題)	设置名称【可选】 🔍
Factor11(量表题)	感知质量 感知价值 Scale1
Factor12(量表證)	顾客满意 顾客忠诚 Scale2

图 6-25 结构方程模型操作界面

由于本例无二阶结构关系,因此无须设定【设置量表二阶结构】。为了方便结果分析, 我们可以在【设置名称【可选】】窗格设置因子的名称。Factor1~Factor4 依次设置为"感知 质量""感知价值""顾客满意""顾客忠诚"。

在分析框上方设置 MI 指标, 在下拉列表中选择【输出 MI>5】, 即只输出 MI>5 的修 正项目, 最后单击【开始分析】按钮。

2) 模型参数估计与检验

在结构方程模型中我们主要关注的是潜变量因子间的路径系数,其模型回归系数及显 著性检验如表 6-31 所示。表中除表头外共有 16 行,前 4 行即因子间的结构模型关系,感 知质量→感知价值、感知质量→顾客满意、感知价值→顾客满意、顾客满意→顾客忠诚的 路径系数(标准化回归系数)依次为 0.509、0.337、0.357、0.485,路径影响关系均显著, 具有统计学意义(p<0.05),模型设定时提出的因子间正向影响关系假设成立。

Х	\rightarrow	Y	非标准化回归系数	SE	z(CR 值)	р	标准化回归系数
感知质量	+	感知价值	0.472	0.121	3.909	0.000	0.509
感知质量	+	顾客满意	0.307	0.127	2.424	0.015	0.337
感知价值	+	顾客满意	0.350	0.136	2.567	0.010	0.357
顾客满意	→	顾客忠诚	0.404	0.129	3.121	0.002	0.485
感知质量	+	A4	1.029	0.188	5.461	0.000	0.575
感知质量	+	A3	1.013	0.184	5.514	0.000	0.585
感知质量	+	A2	1.065	0.190	5.607	0.000	0.605
感知质量	→	A1	1.000	—	—	—	0.581
感知价值	+	B3	1.355	0.230	5.881	0.000	0.743
感知价值	+	B2	0.989	0.180	5.503	0.000	0.557
感知价值	+	B1	1.000	—	—	—	0.592
顾客满意	→	C3	1.163	0.229	5.081	0.000	0.573
顾客满意	→	C2	1.220	0.232	5.265	0.000	0.636

表 6-31 模型回归系数及显著性检验

_	顾客满意	+	C1	1.000	_	_	_	0.556
_	顾客忠诚	t	D2	1.274	0.396	3.219	0.001	0.668
_	顾客忠诚	1	D1	1.000			—	0.598

注: →表示回归影响关系或测量关系。

后 12 行的结果为测量模型关系,其结果可按前面介绍的验证性因子分析进行解读。 显然本例大多数路径系数大于 0.5,说明所有测量关系均显著,具有统计学意义 (p<0.05), 说明模型的内在质量较好。如果发现有较多路径系数没有显著性 (p>0.05),则说明测量模 型适配一般,提示要重新结合理论进行关系设定以调整模型。

常用 指标	χ ²	df	р	卡方自由 度比	GFI	RMSEA	RMR	CFI	NFI	NNFI
判断 标准	—	—	>0.05	<3	>0.9	<0.10	<0.05	>0.9	>0.9	>0.9
值	83.683	50	0.002	1.674	0.938	0.058	0.036	0.916	0.821	0.889
其他 指标	TLI	AGFI	IFI	PGFI	PNFI	SRMR	RMSEA 90% CI			
判断 标准	>0.9	>0.9	>0.9	>0.5	>0.5	<0.1	_			
值	0.889	0.903	0.919	0.601	0.622	0.063	0.035~0.079			

表 6-32 模型拟合指标

注: Default Model: $\chi^2(66)=466.207$, $p=1.000_{\circ}$

3) 模型拟合评价

和前面验证性因子分析、路径分析相同,可使用卡方自由度比、RMSEA、RMR、GFI、 CFI、NFI 和 NNFI 等指标对模型拟合质量进行综合评价。

模型拟合指标如表 6-32 所示。本例卡方自由度比=1.674<3, RMSEA=0.058<0.1, SRMR<0.1, GFI、AGFI、CFI、IFI 均大于 0.9, TLI 接近 0.9, 多数常用指标达到了拟合标准, 综合认为模型拟合良好, 根据理论设定的模型与实际样本数据适配。

如果模型总体上拟合欠佳或前面有路径系数不显著,则可考虑将不显著的路径系数删 除,并依据 MI 指标进行适当修正。

4) 模型 MI 修正

如果模型拟合评价认为模型适配欠佳,可考虑通过 MI 修正项的提示进行适当修正。 一般而言,对于结构方程模型的修正可以从影响关系与误差、协方差关系两个方面入手。 针对影响关系的 MI 修正指标与针对模型协方差关系的 MI 修正指标,分别如表 6-33 和表 6-34 所示。

项	关系	项	MI 值	Par Change
· · · · · · · ·	工业估入初小工作	* 甘心 丙日	1 田山工校山	

注:由于 MI 值全部小于 5 或其他原因,因此无输出。

表 6-34 针对模型协方差关系的 MI 修正指标

测量项 1	关系	测量项 2	MI 值	Par Change
B2	\leftrightarrow	A3	6.035	-0.085
B2	\leftrightarrow	C2	7.411	0.095
B1	\leftrightarrow	D1	11.092	0.081
B1	\leftrightarrow	C2	8.927	-0.098

注:表格中的 MI 值均大于 5。

显然在 MI>5 的条件下无须对影响关系进行修正,而在测量项方面,最大 MI 值为 11.092 提示进行修正,它是指重新设定一次结构方程模型,模型设定时要增加建立 B1 题 项和 D1 题项的协方差相关关系,可以使卡方值降至 11.092,但考虑 B1 题项和 D1 题项分 别属于不同的维度,在特质上是独立的或没有理论支撑的,因此"B1↔D1"的协方差关系并 不符合逻辑,同理其他修正项目的协方差关系也不符合逻辑,本例暂无可执行的修正。根 据 MI 修正指标对测量误差进行协方差修正在实际分析研究中较常见,如果模型适配拟合 不佳,并且存在合理且必要的修正项目,则根据理论驱动原则进行必要修正。

第1项		第2项	
Factor1	▶ 影响	→ V Factor3	~ + -
Factor1	✔ 影响	→ V Factor2	~ + ·
Factor2	▶ 影响	→ ¥ Factor3	~ +
Factor3	▶ 影响	→ V Factor4	~ +
A1 置量表二阶结构【高	→ 相关 影响 3级设置,相关	⇔ ✓ A4 → ⇔	~ + ·

图 6-26 结构方程模型修正操作界面

例如,我们假设模型拟合不佳,最大 MI 值提示增加"A1↔A4"协方差关系,只需要再 次打开【结构方程模型 SEM】模块,在【设置模型关系】窗格增加一条"A1↔A4"的协方差 关系。单击【+】按钮后,在【第 1 项】下拉列表中选择【A1】,在中间的关系类型下拉 列表中选择【相关↔】,在【第 2 项】下拉列表中选择【A4】,即表示设定"A1↔A4"的协方 差相关。结构方程模型修正操作界面如图 6-26 所示,仅用于操作示范。

5) 结果分析与报告

多数常用拟合指标达标或基本达标、各参数估计值达到显著性水平、测量误差均为正数且显著,说明模型适配良好。4 条影响关系路径系数均显著,具有统计学意义(p<0.05),模型设定时提出的因子间正向影响关系假设成立。本例结构方程模型路径图(含参数)如图 6-27 所示。



图 6-27 本例结构方程模型路径图 (含参数)

例 11-11 直接效应值、间接效应值、总效应值表如表 6-35 所示。

影响关系	直接效应值	间接效应值	总效应值
感知质量→感知价值	0.509		
感知质量→顾客满意	0.337	0.182	0.691
感知价值→顾客满意	0.357	—	—
顾客满意→顾客忠诚	0.485	—	_

表 6-35 例 11-11 直接效应值、间接效应值、总效应值表

6.8.3 结构方程模型分析讨论

结构方程模型由测量模型与结构模型构成,对应的有验证性因子分析和路径分析过程。 在理想状态下,验证性因子分析和结构方程模型分析都需要适配拟合指标,但在实际分析 时,"事与愿违"的现象也时有发生。

通常情况下,我们是将量表中所有的题项按预设模型结构来做验证性因子分析的,在 实际分析中,有一些研究的验证性因子分析结果在合理修正后仍然难以在拟合指标上完全 达标。个别研究的变量较多,甚至要研究二阶、三阶潜变量间的关系,变量间影响关系很 复杂,实际分析时难免出现难以拟合适配的情况。导致模型拟合差或难以拟合适配的原因 可能是多方面的,要有针对性地解决这些问题。结构方程模型拟合差时的分析策略如表 6-36 所示,以下策略可根据具体情况进行分析讨论和参考使用。

策略	内容说明
	(1) 量表问卷的样本量是否足够,有可能的话应当增加受试者数量
数据质量	(2) 量表数据是否近似正态分布或服从正态分布
	(3) 更加严苛地筛查、剔除无效问卷,以排查异常问卷对分析的干扰
	(4) 从源头重视共同方法偏差的控制
	针对验证性因子分析模型:
	① 变量关系较多的验证性因子分析模型无法拟合时,考虑将量表题项按自变量、因变量、
	其他变量拆分后对"细分"的验证性因子分析模型进行分析
	② 引进、开发新量表的验证性因子分析无法拟合时,要重新考虑理论支撑和预设结构,根
公长甘西	据某些理论或文献资料合并个别维度
J 1013X+J	针对结构方程模型:
	① 含有高阶潜变量的结构方程模型无法拟合时,考虑题项打包、简化模型,将高阶潜变量
	转换为一阶测量结构
	② 在有足够结构效度、信度支持下,将题项加总或求平均分,将各潜变量转换为显变量,进
	行显变量路径分析

表 6-36 结构方程模型拟合差时的分析策略

6.9中介效应分析

在心理学、医疗卫生等领域的影响关系研究中,已经不再局限于只考虑自变量对因变 量的影响,许多研究还会涉及其他变量,如中介变量、调节变量。本节介绍中介变量与中 介效应分析,通过在仪表盘中依次单击【问卷研究】→【中介作用】模块来实现。

6.9.1 中介变量与中介效应

中介(Mediation)变量,是指在考虑自变量 X 对因变量 Y 的影响时,如果 X 通过影 响变量 M 来影响 Y,则称 M 为中介变量。而 X 通过中介变量 M 对 Y 产生的影响就是中介 效应。例如,张晓等人(2009)在家庭收入对儿童一般社会能力的影响关系研究中,发现 家庭社会文化环境发挥了中介效应,即家庭收入通过影响社会文化环境从而影响儿童的一般社会能力。

本书介绍的中介变量、中介效应,只限定于中介变量和因变量均为定量数据的情况, 中介变量或因变量为定类数据时,请参考其他专业书籍。

1.中介效应模型



图 6-28 X→Y 影响关系路径图 图 6-29 简单中介效应路径图

单独研究自变量 X 与因变量 Y 的影响或预测时,其影响关系路径图如图 6-28 所示。 对应的回归方程为: $Y = cX + e_1$,回归系数 c 即自变量 X 影响因变量 Y 的路径系数,它 是在不考虑其他变量影响下的 X 与 Y 之间的总效应, e1 为因变量 Y 中不能被回归模型解 释的残差。

最简单的中介效应模型即一个自变量、一个中介变量、一个因变量,构成的路径图如 图 6-29 所示。

在中介模型路径中包括两个线性回归方程,第一个是以中介变量为因变量的回归方程, 其模型为*M=aX+e*₂,回归系数 a 即自变量 X 影响中介变量 M 的路径系数,e2 表示无法被 该模型解释的残差;第二个是以 Y 为因变量,X 和 M 为自变量的回归,其模型为 *Y=c'X+bM+e*₃,偏回归系数 c'是在中介变量 M 保持不变时自变量 X 对因变量 Y 的路径系 数,在中介效应分析中将 c'称为在控制其他变量下 X 对 Y 的直接效应,偏回归系数 b 是自 变量 X 保持不变时中介变量 M 对因变量 Y 的路径系数,e3 表示无法被该模型解释的残差。

2.中介效应分解

自变量 X 通过中介变量 M 对因变量 Y 产生的影响效果称为中介效应,中介效应属于间接效应,其路径关系为 X→M→Y,前半段路径为 X 对 M 的影响,其路径系数为 a,后 半段路径为 M 对 Y 的影响,其路径系数为 b,整个中介路径上的效应用前半段路径系数 a 乘以后半段路径系数 b 来计算,即用乘积项 a×b 来估计中介效应量,我们直接写为 ab。

前面介绍了回归方程中的路径系数 c、a、b、c′,以及中介效应 ab,当 M 和 Y 均为定 量数据时,具有以下关系: c=ab+c′,即 X 对 Y 的总效应可分解为直接效应与间接效应, 根据这一关系,我们可以计算中介效应在总效应中的占比。此处应注意,中介效应一定是 间接效应,但反过来间接效应不一定完全等价于中介效应。

3.中介效应类型

根据中介变量的个数,只有1个中介变量时的中介效应称为简单中介效应,中介变量 多于1个时的中介效应称为多重中介效应。 多重中介效应又可以分为平行中介效应和链式中介效应,平行中介效应指两个或多个 中介变量是平等关系,互相独立。两个平行中介变量的模型如图 6-30 (a)所示。链式中 介效应指两个或多个中介变量具有影响关系。两个中介变量的链式中介模型如图 6-30 (b) 所示,其中中介变量 M1 影响中介变量 M2,整条中介影响路径串联为 X→M1→M2→Y。



(a)两个平行中介变量的模型 (b)两个中介变量的链式中介模型 图 6-30 平行中介模型与链式中介模型

6.9.2 中介效应检验流程与实例

1.中介效应检验方法

中介效应分析的核心任务是检验中介效应是否不等于 0。目前有多种检验思路,如传统上的"依次回归法",逐个检验回归系数 c≠0、a 和 b 同时不等于 0、c²是否等于 0,这种方法被证明检验功效较低,已不再推荐使用;再如 Sobel 检验,它的思想是直接检验中介效应原假设 ab=0 是否成立,但是该方法受样本量和 ab 服从正态分布要求的限制和影响,导致其使用范围较小,而且相对保守,检验功效一般。

当前接受度较高、普遍使用的中介检验方法是 Bootstrap 置信区间(CI)法。该方法 的思想也是直接检验中介效应原假设 ab=0 是否成立,关键操作是先进行 Bootstrap 有放回 的重复抽样获得 Bootstrap 样本,用该样本的数据计算路径系数 a 和 b,再计算得到 ab, 重复抽样多少次就会得到多少个 ab 数据,从而获得 ab 的分布,取该分布的第 2.5 和第 97.5 百分位数,得到 ab 统计量的 95% CI,该区间也称为非参数百分位 Bootstrap CI,如果 该区间不包括数字 0,则表明 ab≠0,认为中介效应是成立的;否则如果区间内包括数字 0, 则认为中介效应不存在。

举个例子,假设某研究获得有效量表问卷 500 份,设定 Bootstrap 抽样 5000 次,每次 都从这 500 份问卷中有放回的重复抽取 500 个样本,用每次获得的 Bootstrap 样本数据计 算 ab 值,5000 次抽样可计算 5000 个 ab 值,由小到大排序,取第 2.5 和第 97.5 百分位数, 作为 ab 值的 95% CI,如果该区间不包括 0 则认为 ab 不等于 0,那么 ab 中介效应是存在 的。

关于抽样次数,如果样本量小于等于 500,则抽样次数可设定为 5000 次;如果样本量 介于 500~2000(含)范围内,则抽样次数可设定为 1000 次;如果样本量在 2000 以上, 则抽样次数可设定为 50。

2.中介效应检验流程

以 Bootstrap CI 作为检验方法, 温忠麟和叶宝娟(2014) 在《中介效应分析_方法和模型发展》一文中提出了中介效应检验流程, 如图 6-31 所示。



图 6-31 中介效应检验流程

(1) 检验方程式*Y* = *cX* + *e*₁中的回归系数 c, 即检验总效应是否显著, 如果总效应 c 显著 (p<0.05), 则按中介效应立论, 否则按遮掩效应立论。c 显著与否均不影响继续向下 分析, 只决定最终分析结论的解释。

(2) 检验 $M=aX+e_2$ 中的回归系数和 $Y=c'X+bM+e_3$ a 中的偏回归系数 b, 即间接效应 (中介效应) 的前半段路径系数和后半段路径系数是否显著, 如果 a 和 b 同时显著, 则说 明中介效应存在, 报告 ab 的 Bootstrap CI, 继续执行第(4) 步检验直接效应; 如果 a 和 b 至少有一个不显著, 则继续执行第(3) 步。

(3) Bootstrap CI 法检验 ab 是否为 0, 放回重复抽样次数一般可取 1000~5000 次, 如果 CI 不包括 0 则说明 ab 不为 0, 即中介效应存在,继续执行第(4)步检验直接效应; 如果 CI 包括 0 则说明 ab 可能为 0, 即中介效应不存在,分析结束。

(4) 到这一步时,已经确认中介效应存在,需要检验 *Y=c'X+bM+e*₃中的直接效应 c', 如果 c'不显著,则说明只有中介效应,按中介效应解释结果;如果 c'显著,则说明直接效 应存在,继续执行第(5)步。

(5)到这一步时,已经确认中介效应、直接效应都存在,需要根据二者的符号方向决定如何立论和解释结果。例如, ab 和 c'的符号,如果二者同符号(指均为正或均为负),则说明中介效应属于部分中介效应,报告中介效应占总效应的比例为 ab/c;如果二者符号相反,则以遮掩效应立论,报告中介效应与直接效应的比例的绝对值|ab/c'|。

遮掩效应是一种现象, 在数理角度上表现为间接效应 ab 的系数符号, 与直接效应 c⁻ 的系数符号相反, 在实际研究中要结合自身专业知识进行阐述。

3.简单中介效应实例分析

下面进行案例实践,就一个简单中介效应实例来进一步介绍中介检验流程。

【例 11-12】某研究提出如下研究假设:积极劳资关系压力会通过增强个体的积极情绪而提高个体的工作满意度,案例数据文档见"例 11-12.xls",试检验中介效应是否存在。 案例数据为模拟获得,分析结果只用于讲述方法,不代表实际的效应大小。

1) 数据与案例分析

根据案例描述可知,"积极劳资关系压力"作为自变量 X,"积极情绪"作为中介变量 M, "满意度"作为因变量 Y,均为定量数据,构成了一个简单中介效应,其预设模型路径图如图 6-32 所示。



图 6-32 简单中介效应预设模型路径图

分析目的在于检验"积极情绪"在"积极劳资关系压力"对"满意度"的影响关系中是否起中 介效应。

2) 中介效应分析

数据读入平台后, 在仪表盘中依次单击【问卷研究】→【中介作用】模块, 将 3 个变 量按角色依次移入右侧对应的【因变量 Y】【自变量 X】【中介变量 M】分析框中, 如果研 究中还设定了控制变量, 则将其移入【控制变量】分析框中。简单中介效应可理解为只有 一个中介变量的平行中介, 因此在分析框上方的下拉列表中选择【平行中介(默认)】。 Bootstrap 抽样次数常用 1000~5000 次, 也可以由平台自动根据样本量决定合适的抽样次 数, 本例选择 2000 次。简单中介效应分析操作界面如图 6-33 所示, 最后单击【开始分析】 按钮。

■ 问卷	新究 🗸		T-44/\+r	-	平行中介(默认)	~	
					2000次	~	
信度	效度		满意度			因	E.
多选题	单选-多选		积极带资关系压力				
多选-单选	多选-多选	L	1/10/350/3/12/3				
项目分析	权重					自刭	E.
验证性因子分 析	对应分析	e					
路径分析	结构方程模型 SEM						
调节作用	中介作用		积极情绪			中介引	· 在 文
调节中介	KANO模型						

图 6-33 简单中介效应分析操作界面

3) 结果分析

按前面介绍的中介效应检验流程温忠麟和叶宝娟(2014),对输出的结果进行解释和 分析。

(1)对总效应 c 进行检验。中介效应 3 个回归模型参数估计与检验如表 6-37 所示, 呈现了中介效应的 3 个回归模型结果,第 2 列是 X→Y 的总效应回归方程结果,第 3 列是 X→M 的中介前半段路径回归方程, 第 4 列是 X、M→Y 的中介后半段路径与直接效应回归 方程。

由第 2 列可知, "积极劳资关系压力"对"满意度"的回归系数为 0.256, 在 α=0.01 水平 下显著, 即总效应 c 是显著的, 按中介效应立论。

项	满意度	积极情绪	满意度	
告券	2.985**	3.008**	2.308**	
币刻	(10.581)	//ק.ж./д //т.ж./ј ј 26 2.985** 3.008** (10.581) (11.405) 0.256** 0.256** (3.854) (4.123) 260 260 0.054 0.062 0.051 0.058	(6.809)	
和超带次关系压力	0.256**	0.256**	0.199**	
恢恢力页天东压力	(3.854)	(4.123)	(2.953)	
和超情绪			0.225**	
17/1/2 11 21			(3.453)	
样本量	260	260	260	
R^2	0.054	0.062	0.096	
调整 R ²	0.051	0.058	0.089	
<i>F</i> 值	F(1,258)=14.855, p=0.000	F(1,258)=16.996, p=0.000	F(2,257)=13.704,p=0.000	

表 6-37 中介效应 3 个回归模型参数估计与检验

注: ** p<0.01 括号里面为 t 值。

(2) 对前、后半段路径系数 a 和 b 进行检验。由表 6-37 中第 3 列可知, "积极劳资 关系压力"对"积极情绪"的回归系数为 0.256, 在 α=0.01 水平下显著(标记**), 即前半段 路径系数 a 显著; 由第 4 列可知, "积极情绪"对"满意度"的回归系数为 0.225, 在 α=0.01 水平下显著, 即后半段路径系数 b 显著。此时 a 和 b 同时显著, 说明中介效应存在, 应报 告 ab 的 Bootstrap CI。中介效应检验如表 6-38 所示, 继续执行中介效应检验流程的第(4) 步检验直接效应。

由表 6-38 可知, 中介效应 ab=0.256×0.225=0.058, ab 的 95% Cl 为[0.019,0.096], 该 区间内不包括 0, 表明中介效应存在。

项	<i>c</i> 总效应	а	b	ab 中介效 应值	<i>ab</i> (Boot SE)	ab (z 值)	<i>ab</i> (p 值)	ab (95% BootCI)	<i>ć</i> 直接 效应	检验 结论
积资压积省意	0.256**	0.256**	0.225**	0.058	0.020	2.892	0.004	0.019~ 0.096	0.199**	部分 中介

表 6-38 中介效应检验

注: ** p<0.01。

(3) 中介效应检验流程第(4) 步检验直接效应。本例总效应、中介效应、直接效应、 效应占比等结果如表 6-39 所示。第 5 列(直接效应 c)中,"积极劳资关系压力"对"满意 度"的回归系数为 0.199,在 α=0.01 水平下显著,直接效应 c[·]显著,接下来继续执行中介 检验流程中的第(5)步。

表 6-39 总效应、中介效应、直接效应、效应占比等结果

项	检验结论	<i>c</i> 总效应	ab 中介效应	c [′] 直接效应	效应占比 计算公式	效应占比
积极劳资关系压力≥ 积极情绪≥满意度	部分中介	0.256	0.058	0.199	ab/c	22.656%

(4) 中介效应检验流程第(5) 步可通过比较中介与直接效应的符号来决定如何立论 和解读结果。本例中介效应 ab=0.058,为正数,而直接效应 c´=0.199,也为正数,二者方 向相同,说明中介效应属于部分中介效应,中介效应占总效应的比例 ab/c=0.058/0.256=22.656%。

6.9.3 多重中介效应分析与实例

前面我们是以只有一个中介变量的简单中介效应模型来介绍中介效应检验流程的,但 实际上检验流程同样也适用于多个自变量或多个中介变量的模型。当中介变量不止一个时, 相应的模型称为多重中介效应模型,包括平行中介效应模型和链式中介效应模型。在多重 中介效应模型中,会出现多条中介路径,一般称之为特定中介路径。多条中介效应值的加 和为总间接效应或总中介效应,总中介效应与直接效应的加和为总效应,如果各特定中介 效应存在,也可以分别计算其在总效应中的占比。

对多重中介进行检验和分析时,研究者应明确感兴趣的是哪条特定中介路径,多条特 定中介效应的加和建议用总间接效应表述。下面通过一个实例介绍多重中介中的链式中介 效应应用。



图 6-34 链式中介效应模型路径图

【例 11-13】在陈世民等人(2014)的研究中,基于有关理论提出"自我提升幽默"通过"情绪幸福"和"社会支持"两个中介变量来提升"生活满意度"。其链式中介效应模型路径图 如图 6-34 所示,试进行链式多重中介效应检验。

数据文档为"例 11-13.xls",案例使用的为模拟数据,分析结果只用于讲述方法,不代 表实际的效应大小。

1) 数据与案例分析

根据案例描述可知, X 表示自变量(自我提升幽默), M1 表示中介变量(情绪幸福), M2 表示中介变量(社会支持), Y 表示因变量(生活满意度), 均为定量数据资料。该模型中包括 3 个特定中介路径: ①X→M1→Y; ②X→M2→Y; ③X→M1→M2→Y。其中第 3 条中介路径中, M1 与 M2 串联, 形成一条长的链式中介路径。

对于①X→M1→Y、②X→M2→Y 这两条特定中介路径的解释和分析与上例简单中介效 应完全一致,读者可参考上例进行结果分析。本例将重点解释和分析 X→M1→M2→Y 链式 中介路径的效应是否存在。

2) 中介效应分析

数据读入平台后, 在仪表盘中依次单击【问卷研究】→【中介作用】模块, 将 4 个变 量拖曳至相应的分析框中, 中介类型选择下拉列表中的【链式中介】, Bootstrap 抽样次数 选择下拉列表中的【5000次】。链式中介效应分析操作界面如图 6-35 所示, 最后单击【开 始分析】按钮。

■ 问義	新研究 🗸	开始分析			链式中介	Ň O
			开始方机	11	5000次	^ ¥
信度	效度		Υ			因变量Y
多选题	单选-多选		Y			
多选-单选	多选-多选		^			
项目分析	权重					自变量X
验证性因子分 析	对应分析	e				
路径分析	结构方程模型 SEM					
调节作用	中介作用		M1			
调节中介	KANO模型		M2			
NPS	PSM					中介变量M

图 6-35 链式中介效应分析操作界面

3) 结果分析

(1)对总效应 c 进行检验。链式中介效应回归模型参数估计与检验如表 6-40 所示,
呈现了多重中介效应的 4 个回归模型结果,由第 4 列结果(X→Y 总效应方程)可知, X 对
Y 的回归系数为 0.384,在 α=0.01 水平下显著,即总效应 c 是显著的,按中介效应立论。

项	M1	M2	Y(总效应方程)	Y
告粉	1.658**	1.124**	2.328**	1.592**
市奴	(12.367)	(6.183)	(16.890)	(10.076)
v	0.576**	0.294**	0.384**	0.146**
Λ	(15.424)	(5.408)	(10.011)	(3.111)
M1		0.370**		0.260**
IVI I		(6.403)		(5.155)
MO				0.176**
IVIZ				(4.223)
样本量	400	400	400	400
R^2	0.374	0.312	0.201	0.312
调整 R ²	0.373	0.308	0.199	0.307
E 佶	F(1,398)=237.900,	F(2,397)=89.957,	F(1,398)=100.218,	F(3,396)=59.813,
F 1且	<i>p</i> =0.000	<i>p</i> =0.000	<i>p</i> =0.000	<i>p</i> =0.000

表 6-40 链式中介效应回归模型参数估计与检验

注: ** p<0.01 括号里面为 t 值。

(2) 对 X→M1、M1→M2、M2→Y 这 3 个回归系数进行检验。由表 6-40 中的第 2 列 可知, X 对 M1 的回归系数为 0.576,我们可以将该回归系数设为 a1,显然 a1 显著;由第 3 列结果可知, M1 对 M2 的回归系数为 0.370,用 d 表示,显然 d 显著;由第 5 列结果可 知,M2 对 Y 的回归系数为 0.176,用 b2 表示,显然 b2 显著。此时路径系数 a1、d、b2 同时显著,说明链式中介效应存在,应报告链式中介效应值即乘积项 a1db2 及其 Bootstrap Cl。接下来继续执行中介效应检验流程中的第(4)步以检验直接效应。

(3) 中介效应检验流程第(4) 步检验直接效应。在表 6-40 中第 5 列结果中, X 对 Y 的回归系数为 0.146, 即直接效应 c[·]显著, 接下来继续执行中介效应检验流程中的第(5) 步。

(4) 中介效应检验流程第(5) 步,比较链式中介与直接效应的符号,以决定如何立 论和解读结果。简单中介、链式中介效应检验如表 6-41 所示。

项	Effect	Boot SE	BootLLCI	BootULCI	z	р
X⇒M1⇒Y	0.150	0.036	0.102	0.246	4.159	0.000
X⇒M2⇒Y	0.052	0.019	0.027	0.103	2.680	0.007
X⇒M1⇒M2⇒Y	0.037	0.014	0.018	0.074	2.657	0.008

表 6-41 简单中介、链式中介效应检验

注: BootLLCI 指 Bootstrap 抽样 95%区间下限, BootULCI 指 Bootstrap 抽样 95%区间上限。灰色底纹为链式中介,其余为平行中介。

由表 6-41 可知, X→M1→M2→Y 的链式中介效应, 其 Bootstrap 法标准误为 0.014, 链式中介效应值 Effect 为 0.037 (a1db2=0.576×0.37×0.176, 因小数点计算存在微小差异), 95% CI 为[0.018,0.074], CI 内不包括 0, 说明链式中介效应存在, 结论和前面的一致。

此外 X→M1→Y 和 X→M2→Y 这两个特定中介路径, 它们的 Bootstrap 抽样 95% CI 依 次为[0.102,0.246]、[0.027,0.103], 不包括 0, 说明这两个中介效应同样都是存在的。

本例中链式中介效应 Effect=0.037,为正数,而直接效应 c´=0.146,也为正数,二者 方向相同,说明链式中介效应属于部分中介效应。链式中介效应占总效应的比例为 a1db2/c=0.037/0.384=9.6%。

6.10 调节效应分析

调节效应和中介效应一样,也是对影响关系研究的延伸。中介效应反映的是自变量影 响因变量的机制,而调节效应研究的是自变量在何种条件下如何影响因变量。本节介绍调 节变量与调节效应分析,通过在仪表盘中依次单击【问卷研究】→【调节作用】模块来实 现。

6.10.1 调节变量与调节效应

调节变量 (Moderator), 是指在研究自变量 X 对因变量 Y 的影响时,如果外来的第三 个变量 W 会影响 X 与 Y 关系的大小或方向,则称 W 为调节变量。此时 X 对 Y 的影响关系 受到 W 的调节,依赖于 W 的不同条件, X 对 Y 关系的强弱或方向的变化程度反映的就是 调节效应。

例如,某研究最初假设缩小知识鸿沟(Y)的方法是尽可能多地开展信息普及活动 (X),但是研究结果却恰好相反,受教育程度(W)会扩大而不是缩小信息普及对知识鸿 沟的影响。因为相比受教育程度低的人,那些接触到信息普及活动的受教育程度高的人更 容易学到相关知识。

1.调节效应模型

调节效应在统计上等价于交互作用,可将调节效应直接理解为自变量 X 与调节变量 W 的交互作用 XW 对因变量 Y 有影响,但是在具体研究中,调节效应与交互作用是有区别的。 二者的不同之处在于,调节变量是研究者根据某理论或研究确立的,调节变量和自变量均 有理论支撑,因此二者的身份角色不能互换,但是在交互作用中,两个变量的身份角色是 平等的,可以调换。



(a)调节效应概念图 (b)模型路径图 图 6-36 调节效应概念图与模型路径图

图 6-36 所示为调节效应概念图与模型路径图,图中的自变量 X 与调节变量 W 的关系, 不要求必须相关。

自变量或调节变量至少有一个是定量数据时,调节效应一般采用分层线性回归进行检验,包括两个线性回归方程模型:

 $Y=aX+bW+cXW+e_2 ft{5}ft{2}(2)$

方程(1)指分层线性回归的第一层放入 X、W 两个变量,做 Y 对 X、W 的回归;方 程(2)指分层线性回归的第二层放入交互项 XW(X 与 W 中心化或标准化后的数据乘积), 做 Y 对 X、W 及 XW 的回归。

调节效应存在与否有两种检验方法,第一种是依据交互项的偏回归系数 c,如果 c 显 著,则表示调节效应显著;第二种是依据两个方程的 R2 变量 (R2),如果 R2 变量显著, 则表示调节效应显著。

2.调节效应场景

在数据类型上, 自变量 X 与调节变量 W 既可以是定类数据, 如性别、受教育程度, 也可以是定量数据, 如满意度评分、抑郁评分。对于因变量为定类数据的情况本书不进行 介绍。根据 X 与 W 的数据类型不同, 调节效应可以分为以下 4 种场景, 如表 6-42 所示。

4 种场景	X 定类 W 定类	X定量 W定类	X 定类 W 定量	X定量 W定量	
调节检验方法	方差分析	分层或分组线性回	公已经州同山	分层线性回归	
		归,建议用分层回归	力压线性回归		
数据预处理	无	X 中心化/标准化	X哑变量处理	x と w 由 き 化 / 長速化	
		W哑变量处理	W 中心化/标准化	X 与 W 中心化/ 协催化	
判断依据	若交互项 XW 显	(1) 若交互项	,则调节效应显著		
	著则调节效应显著	(2) 若分层回归方程 ΔR^2 显著,则调节效应显著			

表 6-42 调节效应的 4 种场景

(1) X 与 W 均为定类数据。第一种场景可以使用考察交互作用的方差分析,如果交 互项 XW 显著则认为调节效应显著。

(2) X 与 W 至少有一个为定量数据。此情况包括了表中后 3 种场景。

因变量为定量数据时,采用线性回归考察影响关系,而遇到定类自变量时应事先将其 转换为哑变量进行回归,因此不管是自变量 X 还是调节变量 W,如果为定类数据,则统一 进行哑变量转换。遇到定量数据,在调节效应分析时一般采取中心化(变量数据减去其平 均值)或标准化(Z 得分法)处理。

通过哑变量、中心化或标准化处理后,对调节效应的检验,可统一采用以下两种方法

中的一种: ①若交互项 XW 回归系数 c 显著, 则调节效应显著; ②若分层回归方程 R2 显著, 则调节效应显著。

6.10.2 简单斜率与斜率图

将回归方程 Y=aX + bW + cXW + e₂ 中的 X 单独提取出来进一步整理,可得如下方程: Y=bW + (a + cW)X + e₂ 方程 (3)

方程(3)中自变量 X 的回归系数为(a+cW),也就是 X 对 Y 的回归斜率,此时它是 W 的一元回归函数。我们将(a+cW)称为简单斜率,指自变量 X 对因变量 Y 的影响关系是如何 受变量 W 调节的。当 W 取不同值时简单斜率有不同的影响程度,即调节变量在不同条件 时,自变量对因变量的影响关系的程度或大小会发生改变。

为方便研究和结果解释,通常采用选点法对简单斜率的显著性进行检验。所谓选点法 即固定调节变量 W 的几个取值,计算该条件下简单斜率的 t 统计量和 p 值,从而判断简单 斜率的显著性。W 的固定取值点,一般根据平均值 M、M+1 倍 SD(标准差)、M-1 倍 SD 这 3 个条件取值,对应中等调节水平、高调节水平、低调节水平,常采用其中的高调节点、 低调节点绘制简单斜率图。



简单斜率图示例如图 6-37 所示。调节变量 W 取高水平(M+1SD)时, 自变量 X 对因变量 Y 的影响是正向的, 即 X 取值越高 Y 相应越大; 而当 W 取低水平(M-1SD)时, X 对 Y 的影响变为负向的, 即 X 取值越高 Y 相应越小, 直观地体现了调节变量在不同条件时, 如何影响简单斜率的程度或方向。

以上介绍的是调节变量取定量数据时的简单斜率。当调节变量为分类数据时,选点法 简单斜率直接取分类数据的水平值。例如,将性别作为调节变量,选点法可取性别为男性、 性别为女性这两个不同条件下,X对Y的斜率变化。

调节效应显著时,还需要对 W 取不同条件时简单斜率的显著性做检验,可直接依据 t 检验 p 值进行检验,若 p 值小于 0.05 则说明在调节变量 W 的某个选点条件下,自变量对 因变量的影响是显著的,反之不显著。

6.10.3 调节效应分析步骤与实例

1.调节效应分析步骤

调节效应分析总体上包括三大部分,第一部分是数据处理,第二部分是调节效应检验, 第三部分是简单斜率分析,具体分为以下 5 个步骤,如图 6-38 所示。



图 6-38 调节效应分析步骤

1) 数据处理

首先根据自变量、调节变量的数据类型,确认其属于调节效应四大场景中的哪一类。 如果自变量、调节变量不同时为定类数据,此时定类数据应转换为哑变量,定量数据应先 进行中心化或标准化操作,再将处理的自变量与调节变量计算乘积得到二者的交互项。

例如,常见的自变量、调节变量均为定量数据的场景,先将自变量与调节变量进行中 心化得到 Zx 和 Zw,再将二者相乘得到交互项 Zx×Zw。如果自变量、调节变量其中有一个 是定类数据,则定类数据进行哑变量转换后与定量数据的中心化或标准化数据依次相乘得 到多个交互项。

哑变量、中心化、标准化均可在 SPSSAU 平台【调节作用】模块通过勾选相关复选框 自动实现,实际分析时并不需要用户额外操作,因变量一般不做中心化或标准化处理。

2) 建立调节效应模型

以 Y 为因变量,以 Zx、Zw、交互项 Zx×Zw 为自变量建立线性回归模型。在实际分析 时,采取分层回归的方式,如第一层放入 Zx、Zw,第二层放入交互项 Zx×Zw。如果在实 际研究中还需要考虑控制变量,则可将所有的控制变量放入第一层,Zx、Zw 放入第二层, 交互项 Zx×Zw 放入第三层;或控制变量、Zx、Zw 统一放入第一层,交互项 Zx×Zw 放入第 二层,总的原则是交互项 Zx×Zw 要单独放入最后一层。

3) 调节效应检验

调节效应存在与否有两种检验方法,第一种是依据交互项 Zx×Zw 的回归系数,如果其显著(t 检验的 p<0.05),则表示调节效应显著或存在;第二种是比较有交互项 Zx×Zw 的回归方程与无交互项 Zx×Zw 的回归方程,如果两个方程的 R2 变量(R2)显著(F 检验的 p<0.05),则表示调节效应存在。

4) 简单斜率分析

当调节效应存在时,应当进一步考察调节变量取不同值时自变量影响因变量的程度和 方向,一般通过选点法进行简单斜率分析,以检验调节变量取低(均值-1 倍标准差)、中 (均值)、高(均值+1 倍标准差)水平时,简单斜率的显著性情况。当简单斜率的 t 检验 p 值小于 0.05 时,表示该条件下自变量对因变量的影响是显著的,反之不显著。

5) 绘制斜率图

斜率图可直观地展示简单斜率分析的结果,有助于解释和分析调节效应,因此一般建 议根据前面的回归方程及简单斜率分析,绘制斜率图,综合前面的结果进行调节效应的解 释和分析。SPSSAU 平台【调节作用】模块可输出该斜率图的结果,以及绘图所需的坐标 点数据,用户可直接采用该斜率图,或利用坐标点数据通过其他工具绘制斜率图,最终进 行分析总结。

2.调节效应实例分析

下面通过具体实例进一步介绍调节效应的应用并进行解释和分析。

【例 11-14】根据某些理论和研究成果,研究者提出如下研究假设:"工作经验"可以 调节"受教育年限"对"当前薪金"的影响,即假设"工作经验"对"受教育年限"与"当前薪金"之 间的关系具有调节作用,试对该调节进行检验和分析。数据来源于 SPSS 统计软件自带的 数据集,对案例说明及数据均进行了编辑及修改,本例仅展示分析过程,不代表真实的变 量关系。

1) 数据与案例分析

本例数据文档见"例 11-14.xls","受教育年限"作为自变量,"当前薪金"作为因变量, "工作经验"作为调节变量,均为定量数据。研究者要研究的是工作经验对受教育年限与当 前薪金的调节作用是否成立,本例将对"受教育年限""工作经验"做中心化处理,并计算二者 的交互项,通过分层线性回归建立调节效应模型,完成调节效应检验与简单斜率分析。

2) 调节效应分析

数据读入平台后,在仪表盘中依次单击【问卷研究】→【调节作用】模块,将【受教 育年限】【当前薪金】【工作经验】拖曳至对应的分析框中。此处应注意,如果自变量或调 节变量为分类数据,则平台会自动完成哑变量转换。如果研究中有控制变量,则统一移入 【控制变量[可选]】分析框中,本例无。

"受教育年限""工作经验"均为定量数据,因此要先在分析框上方的下拉列表中选择【X 定量 Z 定量(默认)】及【中心化(默认)】,对"受教育年限""工作经验"进行中心化处理,平台 会自动计算二者的交互项。调节效应操作界面如图 6-39 所示,最后单击【开始分析】按 钮。

€ 问卷	研究 🗸	开始分析	Ō	X定量Z定量(! ~ 中心化(默认) ~	Q
多选-单选	多选-多选	当前薪金			Y(定量)
项目分析	权重				
验证性因子分析	对应分析	受教育年限			自变量X
路径分析	结构方程模型 SEM	工作经验			调节变量Z
调节作用	中介作用			控	制变量【可选】
调节中介	KANO模型				

图 6-39 调节效应操作界面

3) 建立调节效应模型

为了方便展示,本节以分层线性回归的形式展示调节效应的分析结果。分层回归模型 与调节效应检验如表 6-43 所示。SPSSAU 平台默认为调节效应建立了 3 个回归方程模型, 模型 1 是指自变量"受教育年限"对因变量"当前薪金"的回归关系,可以忽略。重点是模型 2 和模型 3,模型 2 构建的是自变量"受教育年限"、调节变量"工作经验"对"当前

薪金"的回归模型; 模型 3 是在模型 2 基础上新增加了二者的交互项后共同对"当前薪 金"的回归模型。

项	模型 1	模型 2	模型 3	
常数	34419.568**	34419.568**	33850.360**	
	(58.391)	(58.596)	(56.920)	
受教育年限	3909.907**	4020.343**	4131.246**	
	(19.115)	(19.085)	(19.762)	

表 6-43 分层回归模型与调节效应检验

てたほみ		12.071*	0.805	
工作经验		(2.078)	(0.127)	
受教育年限×			-7.492**	
工作经验			(-4.064)	
样本量	474	474	474	
R^2	0.436	0.441	0.460	
调整 R ²	0.435	0.439	0.457	
<i>F</i> 值	F(1,472)=365.381,p=0.000	F(2,471)=186.132,p=0.000	F(3,470)=133.679,p=0.000	
ΔR^2	0.436	0.005	0.019	
ΔF 值	F(1,472)=365.381,p=0.000	F(1,471)=4.316,p=0.038	F(1,470)=16.512,p=0.000	

注: 1.因变量: 当前薪金。

2.* p<0.05 ** p<0.01 括号里面为 t 值。

4) 调节效应检验

研究者可以根据交互项回归系数的显著性,以及含有交互项模型与不含交互项模型 R2 变化量的显著性这两种依据来检验调节效应是否存在,不管采用哪种,其结论都是一致的。 在表 6-43 模型 3 中,我们看到"受教育年限×工作经验"交互项回归系数为-7.492,t=-4.064,p<0.01,在 α=0.01 水平下认为交互项具有统计学意义,交互项对"当前薪金"的影 响显著,因此"工作经验"的调节效应成立。模型 2 变化到模型 3 时,该过程 R2 的变化量显 著 (F=16.512,p<0.01),同样认为调节效应成立。

5) 简单斜率分析

"工作经验"对"受教育年限"与"当前薪金"的关系存在调节作用, 接下来我们通过选点法 进一步考察"工作经验"在低、中、高不同取值水平时, "受教育年限"对"当前薪金"的简单斜 率显著性及变化。简单斜率分析如表 6-44 所示。

调节变量水平	简单斜率	标准误	t	р	95% CI	
低水平(-1SD)	4914.780	302.338	16.256	0.000	4322.209	5507.352
平均值	4131.246	209.054	19.762	0.000	3721.509	4540.984
高水平(+1SD)	3347.712	265.251	12.621	0.000	2827.830	3867.595

表 6-44 简单斜率分析

工作经验在低、中、高不同取值水平时,简单斜率依次为 4914.780、4131.246、 3347.712,均具有统计学意义(t1=16.256、t2=19.762、t3=12.621,均p<0.01)。

6) 斜率图分析

"工作经验"对"受教育年限"与"当前薪金"的调节斜率图,如图 6-40 所示。



图 6-40 "工作经验"对"受教育年限"与"当前薪金"的调节斜率图

结合简单斜率分析表中不同调节水平下的简单斜率及图 6-40 中的 3 条直线的变化, 可以发现,"受教育年限"正向影响"当前薪金","受教育年限"越高,"当前薪金"相应越高。 这种关系受"工作经验"的调节,"工作经验"调节水平由低水平变到中等水平再变到高水平时, "受教育年限"对"当前薪金"的回归系数逐渐降低。通俗理解即"受教育年限"对"当前薪金"的 影响被"工作经验"削弱。

6.11 有调节的中介分析

在一些复杂的影响关系模型中,会同时包含中介变量和调节变量,常见的研究形式包括有调节的中介、有中介的调节。相对而言,有调节的中介应用更为广泛,本节介绍有调 节的中介,通过在仪表盘中依次单击【问卷研究】→【调节中介】模块来实现。

6.11.1 方法概述

1.有调节的中介

有调节的中介,可以通俗理解为自变量 X 通过中介变量 M 对因变量 Y 的间接影响受 第四个变量 W 的影响,此时中介效应的大小或方向受另一个调节变量 W 的影响,中介效 应的表现是有条件的,所以也可以把有调节的中介称为调节中介。



图 6-41 有调节的中介模型示例

举一个研究的实例,有调节的中介模型示例如图 6-41 所示。"日常性学业复原力"对 "感恩"与"学业成就"的关系具有部分中介效应,而"压力性生活事件"对"日常性学业复原力" 178/187 的中介作用具有调节效应,具体表现为"日常性学业复原力"对"学业成就"的影响,随"压力性生活事件"的增加而降低,所以"感恩"对青少年"学业成就"的影响是有调节的中介效应。

"日常性学业复原力"的中介作用受"压力性生活事件"的影响,或者说"压力性生活事件" 的调节作用通过中介路径传递并影响"学业成就",该模型深化了"感恩"与"学业成就"的直接 关系,不仅可回答"感恩"怎样影响"学业成就",而且可回答这种影响"何时"更强或更弱,这 类模型就是典型的有调节的中介模型。

有调节的中介,研究的重点是中介效应,需要回答的是这个中介效应是否受到调节。 所以从大的方面讲,它包括两大部分内容,第一部分是证明中介效应存在,第二部分是检 验中介效应受到调节。

2.有调节的中介类型

按中介效应的前半段路径、后半段路径及直接路径是否受到调节,有调节的中介模型 分为7种。SPSSAU平台参考 Andrew F.Hayes 学者开发的 Process 宏程序,采取 Process 的模型编号,模型具体包括模型5、模型7、模型8、模型14、模型15、模型58、模型 59。为方便理解,以下统一按只有一个自变量、一个中介变量、一个调节变量、一个因变 量的情况进行讨论,其中模型59的概念图与统计图如图6-42所示。





(a) 模型 59 的概念图(b) 模型 59 的统计图图 6-42 模型 59 的概念图与统计图

1) 模型 5

模型 5 只调节中介模型中的直接路径,严格意义上来说并不是有调节的中介,按简单 中介和简单调节效应进行分析和解释。

2) 模型 7 与模型 8

模型 7 只调节中介模型中的前半段路径,模型 8 能同时调节中介模型中的前半段路径 与直接路径,有调节的中介效应为 $M = (a_1 + a_3 W)b$ 。M 和 W 分别表示中介和调节变量, 其中 a1 是 X 对 M 的回归系数, a3 是交互项 XW 的回归系数, b 是 M 对 Y 的回归系数。

3) 模型 14 与模型 15

模型 14 只调节中介模型中的后半段路径,有调节的中介效应为 $M = a(b_1 + b_3 V)$, M 和 V 分别表示中介和调节变量,其中 a 是 X 对 M 的回归系数, b1 是 M 对 Y 的回归系数, b3 是交互项 MV 的回归系数。

模型 15 能同时调节后半段路径与直接路径,有调节的中介效应为 $M = a(b_1 + b_2 V)$ 。 M 和 V 分别表示中介和调节变量,其中 a 是 X 对 M 的回归系数, b1 是 M 对 Y 的回归系 数, b2 是交互项 MV 的回归系数。

此处注意, b2 和 b3 均是中介 M 与调节 V 的交互项回归系数, 差别只在于编号不同, 其意义一致。模型 14 可视为模型 15 的特例, 所以这两个模型的中介效应我们统一使用 $M = a(b_1 + b_2 V)$ 来解释。

4) 模型 58 与模型 59

模型 58 能同时调节中介模型中的前半段及后半段路径, 有调节的中介效应为

 $M = (a_1 + a_3 W)(b_1 + b_3 W)$ 。M 和 W 分别表示中介和调节变量,其中 a1 是 X 对 M 的回归系数, a3 是交互项 XW 的回归系数, b1 是 M 对 Y 的回归系数, b3 是交互项 MW 的回归系数。模型 59 能同时调节中介模型中的前半段、后半段路径及直接路径,有调节的中介效应为

 $M = (a_1 + a_3 W)(b_1 + b_2 W)$ 。M 和 W 分别表示中介和调节变量,其中 a1 是 X 对 M 的回归系数, a3 是交互项 XW 的回归系数, b1 是 M 对 Y 的回归系数, b2 是交互项 MW 的回归系数。

此处注意, b2 和 b3 均是中介 M 与调节 W 的交互项回归系数, 差别只在于编号不同, 其意义一致。模型 58 可视为模型 59 的特例, 所以这两个模型的中介效应我们统一使用 $M = (a_1 + a_3 W)(b_1 + b_2 W)$ 来解释。

5) 特例说明

温忠麟和叶宝娟(2014)将有调节的中介分成两类:只调节间接效应、同时调节间接 效应和直接效应,并指出图 6-42 所示的模型(Process 模型编号 59),是较一般的有调节 的中介模型,其余模型都是其特例。

例如, 模型 7 实际是当 b2 =0 (后半段路径被调节的交互项回归系数为 0) 的特例, 模型 15 是 a3 =0 (前半段路径被调节的交互项回归系数为 0) 的特例。因此, 模型 59 是 较一般的有调节的中介模型,此时有调节的中介效应为 $M = (a_1 + a_3 W)(b_1 + b_2 W)$,其中 M 和 W 分别表示中介和调节变量, a1 是 X 对 M 的回归系数, a3 是中介前半段路径交互 项 XW 的回归系数, b1 是 M 对 Y 的回归系数, b2 是中介后半段路径交互项 MW 的回归 系数。后面我们将针对调节中介的检验在这个模型表达式基础上进行解释和分析。

3.调节中介的检验

将 $M = (a_1 + a_3 W)(b_1 + b_2 W)$ 进行数学变换,得到下面的表达式:

 $(a_1 + a_3W)(b_1 + b_2W) = a_1b_1 + (a_1b_2 + a_3b_1)W + a_3b_2W^2$

只要上面表达式与 W 有关,或者说随 W 变化,就可以说明中介效应能被调节。在上 式中,跟 W 有关的系数乘积共有 3 个,分别是 a1b2 、 a3b1 与 a3b2 ,所以检验有调节 的中介,就可以简化为检验这 3 个系数乘积是否等于 0,即检验 a1b2 =0、a3b1 =0、a3b2 =0,只要其中有一个证明被拒绝,就可以说明中介效应能被调节。

关于有调节的中介检验方法,常用的有依次检验、系数乘积区间检验、中介效应差异 检验等方法,且有研究表明它们是替补关系而并非竞争关系。从检验力、操作难度、结果 丰富性等方面综合认为,可按依次检验、系数乘积区间检验、中介效应差异检验的顺序进 行检验,如果前一个检验方法结果显著,则如无特别需要,可不用继续做后面的检验。

1) 依次检验

依次检验是指分别按不同系数乘积(包括 a1b2 、a3b1 与 a3b2) 去检验其中两个系数同时不等于 0,若满足则说明中介效应能被调节。例如,针对模型 15,它涉及的是系数乘积 a1b2 =0 的假设,所谓依次检验即检验 a1 \neq 0 且 b2 \neq 0,也就是检验 X 对 M 的回归系数不等于 0,并且 M 与 W 的交互项 MW 的回归系数不等于 0。再如模型 7,它涉及 a3b1 这个系数的乘积,依次检验的是 a3 \neq 0 且 b1 \neq 0,也就是检验 X 对 M 的交互项回归 系数不等于 0,并且 M 对 Y 的回归系数不等于 0。

如果 a1b2 、 a3b1 与 a3b2 任一系数乘积的两个回归系数同时不等于 0, 则说明中介
效应是能被调节的。在结果报告中,可报告调节变量在均值及均值加减一倍标准差取值下 的中介效应情况。

2) 系数乘积区间检验

如果依次检验无法对有调节的中介做出是否存在的结论,可考虑使用系数乘积区间检验方法继续进行检验。所谓系数乘积区间检验,是指直接检验 a1b2 、a3b1 与 a3b2 这 3 个系数乘积是否等于 0,采用的方法是 Bootstrap CI 法,若区间内不包括 0 则说明系数乘积不等于 0,中介效应被调节是成立的。

在 SPSSAU 平台中,对应的输出结果为调节中介作用指数(Index of Moderated Mediation),此处简称为 Index of MM 法。具体分析时,可直接查看 Index 指数区间是否 包括数字 0,如果不包括则说明具有调节中介作用;反之,则说明没有调节中介作用。同 样,在结果报告中,可报告调节变量在均值及均值加减一倍标准差取值下的条件中介效应 情况。

3) 中介效应差异检验

如果系数乘积区间检验无法确认调节中介作用是否成立,则可考虑继续进行中介效应 差异检验。在 SPSSAU 平台输出的结果中,可依据"条件间接效应表"的结果,对不同调节 变量取值条件下,间接效应有无显著性改变或效应值方向改变来判断调节中介作用是否成 立。例如,在低调节水平时,条件间接效应显著,而在高调节水平时,条件间接效应不显 著,说明存在有调节的中介作用,反之不成立;再如在低调节水平时,条件间接效应为正 值,而在高调节水平时,条件间接效应为负值,也可说明存在有调节的中介作用。

6.11.2 有调节的中介作用实例

下面通过具体实例进一步介绍有调节的中介作用。

【例 11-15】张晓等人(2009)的研究表明,在"家庭收入"对儿童"一般社会能力"的预测中,"家庭智力文化取向"发挥了中介作用,"家庭控制性"发挥了调节作用。根据一定的理论提出新的假设:"家庭智力文化取向"中介作用在后半段路径上,受"家庭控制性"的调节。 模拟的数据文档见"例 11-15.xls",试进行有调节的中介作用检验。案例数据为模拟获得, 只用于方法示范,分析结果不代表真实关系和效应。

1) 数据与案例分析

根据案例描述可知, 自变量 X 为"家庭收入", 因变量 Y 为"一般社会能力", "家庭智力 文化取向"为中介变量 M, 而"家庭控制性"作为调节变量 W 调节中介的后半段路径。以上 4 个变量均为定量数据资料, 本例有调节的中介作用模型的路径图如图 6-43 所示。



图 6-43 有调节的中介作用模型的路径图

显然,"家庭控制性"只能调节间接效应不能调节直接效应,对应的 Process 模型编号为 模型 14,有调节的中介效应为 a b1 b3W ,其中 a 是 X 对 M 的回归系数,为了方

便和前面介绍的通用表达式进行比对及理解,式中的 b3 我们改用 b2,因此调节中介作用 为 a b1 b2W , b2 是交互项 MW 的回归系数。只要证明 ab2 不等于 0,则表明有 调节的中介作用成立。

2) 有调节的中介作用分析

数据读入平台后, 在仪表盘中依次单击【问卷研究】→【调节中介】模块, 将 4 个变 量拖曳至对应的分析框中, 在分析框上面的下拉列表中依次选择【Model 14】(调节中介 类型)、【5000 次】(Bootstrap 抽样次数)、【±标准差(默认)】(不同调节取值水平), 最后 单击【开始分析】按钮。有调节的中介效应分析操作界面如图 6-44 所示。



图 6-44 有调节的中介效应分析操作界面

3) 结果分析

我们按依次检验、系数乘积区间检验、中介效应差异检验的顺序对调节中介作用进行 检验。本例为了示范不同检验方法的应用,分别给出了以上 3 个检验方法的解释和分析, 在实践中若前一个检验方法显著,则已无必要继续做后面的检验。

(1) 依次检验。模型 14 调节的中介作用为 $M = a(b_1 + b_2 W)$, a 是自变量 X 对中介 变量 M 的回归系数, b2 是交互项 MW 的回归系数。依次检验的原假设: ab2 =0, 变成检 验 a =0 且 b2 =0。在调节中介作用回归模型表中可检验 a 和 b2 的显著性。回归模型参数 估计与检验如表 6-45 所示。

项	一般社会能力				家庭智力文化取向			
	β	SE	t值	p值	β	SE	t值	p值
常数	-116.045	39.778	-2.917	0.004**	19.952	1.499	13.311	0.000**
家庭收入	30.664	2.174	14.105	0.000**	1.239	0.110	11.231	0.000**
家庭控制性	-27.011	15.202	-1.777	0.076				
家庭智力文化取向	0.058	0.946	0.062	0.951				
家庭智力文化取向×家庭控制性	1.089	0.388	2.809	0.005**				
样本量	757			757				
R^2	0.301			0.143				
调整 R ²	0.297			0.141				
<i>F</i> 值	F (4,752)=81.037,p=0.000			F (1,755)=126.143,p=0.000				

表 6-45 回归模型参数估计与检验

注: ** p<0.01。

表 6-45 中包括两个回归方程,第一个方程是以"一般社会能力"Y 为因变量的,我们关

心的是 M 与 W 二者交互项"家庭智力文化取向×家庭控制性"回归系数 b2 的显著性,显然 b2 为 1.089,与 0 相比差异显著 (t=2.089,p=0.005<0.01);第二个方程是以"家庭智力 文 化取向"M 为因变量的,我们关心的是 X 对 M 的回归系数 a 的显著性,显然 a 为 1.239, 与 0 相比差异显著 (t=11.231,p<0.01)。综上, a \neq 0 且 b2 \neq 0,拒绝 ab2 =0 的原假设,因此 ab2 \neq 0,即表明"家庭智力文化取向"对"家庭收入"与"一般社会能力"的中介效应受 "家 庭控制性"的调节,调节中介作用成立。

就本例而言,此时已无必要再做后面的系数乘积区间检验和中介效应差异检验,可以 直接报告调节变量取低、中、高水平时条件中介效应的情况。

(2)系数乘积区间检验。如果依次检验法无法确认有调节的中介作用是否成立,则考虑采用系数乘积区间检验进行判断和分析。本例为模型 14,系数乘积为 ab2,通过表 6-45 的结果,可计算 ab2 =1.239×1.089=1.350,该值与 0 相比有无差异则需要通过 Bootstrap CI 法进行检验。在 SPSSAU 平台中对应的输出结果为调节中介作用指数 Index of MM,如表 6-46 所示。

表 6-46 调节中介作用指数 Index of MM

	调节变量	中介变量	Index	BootSE	BootLLCI	BootULCI
_	家庭控制性	家庭智力文化取向	1.350	0.530	0.390	2.472

显然,本例 Index=1.350,其 Bootstrap 置信区间法 95% CI 为[0.390,2.472],该区间不 包括 0,即系数乘积 ab2 不等于 0,拒绝 ab2 =0 的原假设,表明本例有调节的中介效应成 立 (此处仅做结果解读示范),结论同依次检验法一致。

(3) 中介效应差异检验。如果系数乘积区间检验还无法确认有调节的中介效应是否显 著,则考虑做中介效应差异检验。在 SPSSAU 平台的输出结果中,可依据"条件间接效应表" 中的结果,直观地比较当调节变量 W 分别取低、中、高水平时,各条件中介效应的显著性 及方向变化情况,如果显著性有改变,或方向有改变,则认为有调节的中介效应成立。

本例条件间接效应(Conditional Indirect Effect)结果如表 6-47 所示。

中介变量	水平	水平值	Effect	BootSE	BootLLCI	BootULCI
家庭智力文化取向	低水平 (-1SD)	0.156	0.283	1.188	-2.198	2.497
	平均值	1.831	2.544	0.835	0.955	4.205
	高水平(+1SD)	3.506	4.805	1.249	2.497	7.412

表 6-47 条件间接效应结果

注: BootLLCI指 Bootstrap 抽样95%区间的下限, BootULCI指 Bootstrap 抽样95%区间的上限。

调节变量取平均值-1SD 时为低水平,取平均值为中等水平,取平均值+1SD 时为高水 平,表 6-47 中的每行中介效应 Effect 值称为调节中介效应值,其 Bootstrap CI 不包括 0, 表示该条件中介效应显著。显然,本例中当"家庭控制性"取低水平时,Bootstrap 抽样 95% CI 为[-2.198,2.497],包括 0,W 取低水平时的条件中介效应不显著,同理我们发现,W 取 高水平时 Bootstrap 抽样 95% CI 均不包括 0,说明高调节条件下的中介效应显著。由于从 低调节到高调节,调节中介效应的显著性改变,因此认为"家庭控制性"对中介效应发挥了 调节作用,调节中介成立。

4) 结果小结

本例中我们示范了依次检验、系数乘积区间检验及中介效应差异检验 3 种方法,早在 依次检验时已经检验到"家庭控制性"对"家庭智力文化取向"中介效应的调节,综合 表 6-45 和表 6-46 的结果分析得到:有调节的中介成立,在家庭控制程度较高的家庭中,"家庭收 入"能通过"家庭智力文化取向"显著影响儿童的"一般社会能力",此时的中介效应值为 4.805, Bootstrap 抽样 95% CI 为[2.497,7.412]; 而在家庭控制程度较低的家庭中, "家庭智力文化 取向"并不能起到中介作用。

参考文献

[1] HAYES A F.Introduction to mediation,moderation,and conditional process analysis:A regression-based approach[M]. 2nd ed.New York: The Guilford Press, 2018.

[2] HIGGS N T. Practical and Innovative Uses of Correspondence Analysis[J]. The Statistician, 1991, 40(2): 183-194.

[3] Huang ZX. Clustering Large Data Sets With Mixed Numeric and Categorical Values[C]. Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining Conference, Singapore, 1997: 21-34.

[4] JAMES L R, WOLF G, DEMAREE R G. Estimating interrater reliability in incomplete designs[J]. Journal of Applied Psychology, 1981, 69.

[5] KAISER H F, RICE J. Little jiffy, mark iv[J]. Educational and Psychological Measurement, 1974, 34(1): 111-117.

[6] LANDIS J R, KOCH G G. The measurement of observer agreement for categorical data[J]. Biometrics, 1977, 33(1): 159-174.

[7] LEBRETON J M, SENTER J L. Answers to 20 Questions About Interrater Reliability and Interrater Agreement[J]. Organizational Research Methods, 2007, 11(4): 815–852.

[8] 陈平雁. SPSS 13.0 统计软件应用教程[M]. 北京:人民卫生出版社,2005.

[9] 陈强. 计量经济学及 Stata 应用[M]. 北京:高等教育出版社, 2015.

[10] 陈世民, 方杰, 高树玲, 等. 自我提升幽默与生活满意度的关系: 情绪幸福和社会支持的链式中介模型[J]. 心理科学, 2014, 37 (2): 377-382.

[11]初立苹,粟芳,我国财产保险公司融资效率的 DEA 比较分析[J].保险研究,2013,(4): 22-32.

[12] 丛晓男. 耦合度模型的形式、性质及在地理学中的若干误用[J]. 经济地理, 2019, 39 (4): 18-25.

[13] 邓雪, 李家铭, 曾浩健, 等. 层次分析法权重计算方法分析及其应用研究[J]. 数学的实 践与认识, 2012, 42 (7): 93-100.

[14] 董大钧. SAS 统计分析应用[M]. 北京: 电子工业出版社, 2014.

[15] 董志永,赵文露,陆应军,等.双参数磁共振 pi-rads v2 联合临床指标对前列腺移行带临床显著癌的诊断价值[J].医学影像学杂志,2021,031 (11):1930-1935.

[16] 都伟欣, 崔颖杰, 卢锦标, 等. 重组结核分枝杆菌 11kda 蛋白皮肤试验与体外干扰素 y 检测方法的比较[J]. 中国生物制品学杂志, 2015, 28 (11): 1183-1186.

[17] 杜强, 贾丽艳. SPSS 统计分析从入门精通[M]. 2 版. 北京: 人民邮电出版社, 2009.

[18] 方积乾. 卫生统计学[M]. 7版. 北京:人民卫生出版社, 2012.

[19] 冯国双. 白话统计[M]. 北京: 电子工业出版社, 2018.

[20] 郭显光. 熵值法及其在综合评价中的应用[J]. 财贸研究, 1994, 5(6): 56-60.

[21] 何勇男. 哈尔滨城市旅游品牌定位研究[D]. 哈尔滨:哈尔滨理工大学, 2012.

[22] 贾俊平. 统计学:基于 SPSS[M]. 北京:中国人民大学出版社, 2014.

[23] 晋瑞, 张冰, 王琳, 等. 非肿块强化乳腺癌 mri 征象与分子分型的相关性分析[J]. 实用 放射学杂志, 2019, 35 (11): 1759-1762.

[24] 李金林,马宝龙.管理统计学应用与实践[M].北京:清华大学出版社,2007.

[25] 李磊,汤学兵,陈战波. 中部六省会城市区域经济竞争力评价研究[J]. 商业经济研究,

2017, 728 (13): 122-125.

- [26] 李晓松. 卫生统计学[M]. 8 版. 北京: 人民卫生出版社, 2017.
- [27] 李志辉,杜志成. MedCalc 统计分析方法及应用[M].北京:电子工业出版社,2018.

[28] 李宗富, 张向先. 政务微信公众号服务质量的关键影响因素识别与分析[J]. 图书情报工作, 2016, 60 (14): 84-93.

[29] 刘红云. 高级心理统计[M]. 北京:中国人民大学出版社,2019.

[30] 卢纹岱,朱红兵. SPSS 统计分析[M]. 5 版. 北京:电子工业出版社, 2015.

[31] 卢纹岱. SPSS for Windows 统计分析[M]. 3 版. 北京: 电子工业出版社, 2006.

[32] 潘晓平, 倪宗瓒. 组内相关系数在信度评价中的应用[]]. 华西医科大学学报, 1999

(1): 62-63+67.

[33] 沈浩,柯惠新.对应分析在新产品名称测试中的应用[J].市场与人口分析,1999,5 (4):26-30.

[34] 宋冬梅, 刘春晓, 沈晨, 等. 基于主客观赋权法的多目标多属性决策方法[J]. 山东大学 学报 (工学版), 2015, 45 (4): 1-9.

[35] 孙振球, 徐勇勇. 医学统计学[M]. 4 版. 北京: 人民卫生出版社, 2014.

[36]王娜. 绿色金融、经济增长与生态环境的耦合协调发展水平测度研究[J]. 中国集体经济, 2019, 618 (34): 21-24.

[37] 王彤. 医学统计学与 SPSS 软件应用[M]. 北京:北京大学医学出版社, 2008.

[38] 王孝玲. 教育统计学[M]. 4 版. 上海: 华东师范大学出版社, 2007.

[39]王新会.公立医院医疗服务绩效与生产效率评价研究[J].现代医院管理,2016,14(5) : 32-35.

[40] 王祎,遇琪,姜蕾,等. 层次分析法在电视节目评价中的应用[J]. 广播与电视技术, 2022,49 (8):13-17.

[41] 温忠麟, 叶宝娟. 有调节的中介模型检验方法: 竞争还是替补?[J]. 心理学报, 2014, 46 (5): 714-726.

[42] 温忠麟, 叶宝娟. 中介效应分析: 方法和模型发展[J]. 心理科学进展, 2014, 22 (5): 731-745.

[43] 闻浩, 陆梦洁, 刘玉秀, 等. 定量测量 bland-altman 一致性评价方法研究及临床应用 [J]. 医学研究生学报, 2015, 28 (10): 1107-1111.

[44] 吴晗,陈大伟.基于 ISM 和 AHP 方法的建筑施工高处坠落事故风险因素分析[J].建筑安全,2018,33 (10):52-56.

[45] 吴明隆. 问卷统计分析实务: SPSS 操作与应用[M]. 重庆: 重庆大学出版社, 2010.

[46] 吴明隆. 结构方程模型: AMOS 的操作与应用[M]. 2 版. 重庆: 重庆大学出版社, 2010. [47] 武松, 潘发明. SPSS 统计分析大全[M]. 北京: 清华大学出版社, 2014.

[48] 许汝福. Logistic 回归变量筛选及回归方法选择实例分析[J]. 中国循证医学杂志, 2016, 16 (11): 1360-1364.

[49] 颜虹, 徐勇勇. 医学统计学[M]. 3 版. 北京: 人民卫生出版社, 2010.

[50]杨维忠,张甜,刘荣. SPSS 统计分析与行业应用案例详解[M].北京:清华大学出版社, 2015.

[51] 叶宝娟,杨强,胡竹菁.感恩对青少年学业成就的影响:有调节的中介效应[J].心理发展与教育,2013,29(2):192-199.

[52] 余红梅,罗艳虹,萨建,等.组内相关系数及其软件实现[J].中国卫生统计,2011,28 (5):497-500.

[53] 余松林. 医学统计学[M]. 北京:人民卫生出版社, 2002.

[54] 张厚粲,徐建平.现代心理与教育统计学[M].5版.北京:北京师范大学出版社,2020.
[55] 张明芝,李红美,吕大兵.实用医学统计学与 SAS 应用[M].苏州:苏州大学出版社,2015.

[56]张伟豪,徐茂洲,苏荣海.与结构方程模型共舞[M].重庆:重庆大学出版社,2020.
[57]张文彤,董伟.SPSS统计分析高级教程[M].2版.北京:高等教育出版社,2013.
[58]张文彤,邝春伟.SPSS统计分析基础教程[M].2版.北京:高等教育出版社,2011.
[59]张文彤.SPSS11.0统计分析教程:高级篇[M].北京:北京希望电子出版社,2002.
[60]张文彤.SPSS11.0统计分析教程:基础篇[M].北京:北京希望电子出版社,2002.
[61]张晓,陈会昌,张银娜,等.家庭收入与儿童早期的社会能力:中介效应与调节效应
[J].心理学报,2009,41(7):613-623.

[62] 郑宗成,张文双,黄龙,等.市场研究中的统计分析方法:专题篇[M].广州:广东经济 出版社,2012.

[63] 周登远. 临床医学研究中的统计分析和图形表达实例详解[M]. 2 版. 北京: 北京科学技术出版社, 2017.

[64] 周俊. 问卷数据分析——破解 SPSS 的六类分析思路[M]. 北京: 电子工业出版社, 2017.[65] 朱海腾. 多层次研究的数据聚合适当性检验: 文献评价与关键问题试解[J]. 心理科学进展, 2020, 28 (8): 1392-1408.