

# SPSS 与统计应用分析

## SPSS & the Application and Analysis of Statistics

本书重视应用，研究如何借助统计软件来分析和解释量化数据资料。内容包含行为科学量化研究中统计方法的理论与实践方法应用、基本统计与高等统计的应用分析、单变量与多变量统计的分析、问卷的有效度与问卷分析，更新增了一章，介绍Amos与结构方程模式。

本书内容包括大量SPSS视窗、中英文操作解说、基本量化研究概念的厘清与方法选用，并辅以例题和实际范例说明，使读者对统计分析的前后脉络有完整的概念。本书的报表说明，让读者对呈现的数据有深入而清晰的了解；结果的呈现与整理说明可以作为论文或研究报告的参考。



本书配套光盘可免费下载：[www.dufep.cn](http://www.dufep.cn)

无防伪标志者均为盗版  
举报电话：(0411)84710523

ISBN 978-7-5654-0867-0



9 787565 408670 >

定价：98.00元



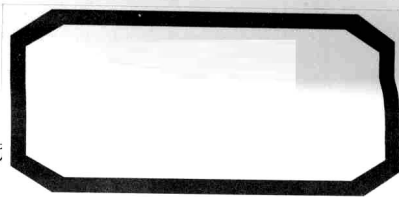
国家出版基金项目  
NATIONAL PUBLICATION FUND PROJECT

经管研究方法系列译丛

“十二五”国家重点图书出版规划项目

当代财经管理名著译库

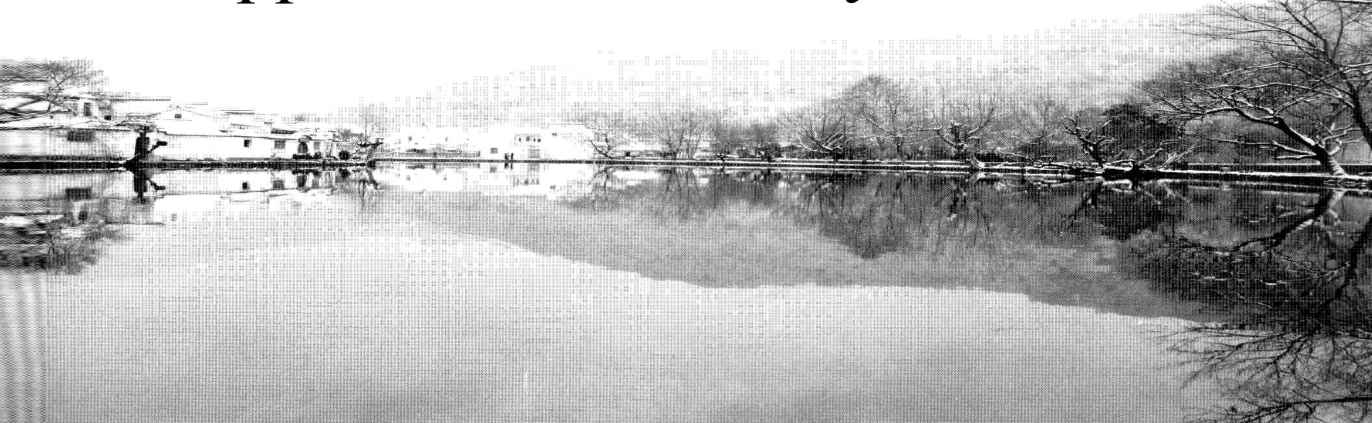
2012年度国家出版基金资助项目



# SPSS

&

## the Application and Analysis of Statistics



# SPSS 与统计应用分析

吴明隆 涂金堂 著

白雪梅 段志民 校

本书不仅是一本统计应用的参考书，更是从事量化研究的工具书！

**FE** 东北财经大学出版社  
Dongbei University of Finance & Economics Press  
大连

## 偏相关、部分相关

7.1 一阶偏相关 .....	222
7.2 二阶偏相关 .....	227
7.3 部分相关 .....	229

## 卡方检验——百分比检验

8.1 基本原理 .....	234
8.2 关联系数 .....	238
8.3 拟合优度检验——期望频数相等 .....	243
8.4 拟合优度检验二——期望频数不相等 .....	250
8.5 百分比同质性检验 .....	253
8.6 独立性检验 .....	258
8.7 改变的显著性检验 .....	264
8.8 费雪尔正确概率检验 .....	267

## 平均数的差异检验——t 检验

9.1 基本原理 .....	272
9.2 单一样本的 t 检验——总体 $\sigma$ 未知和单侧检验 ..	276
9.3 单一样本的 t 检验——总体 $\sigma$ 未知和双侧检验 ..	279
9.4 相依样本的 t 检验 .....	281
9.5 独立样本的 t 检验 .....	284
9.6 学习经验问卷的实例分析 .....	294
9.7 自变量为连续变量的 t 检验 .....	299
9.8 非参数统计——两个独立样本 .....	305
9.9 非参数统计——相依样本的差异检验 .....	311
9.10 家长参与问卷——相依样本实例 .....	314

## 判别分析与 Logistic 回归分析

- 13.1 判别分析····· 510
- 13.2 Logistic 回归分析 ····· 530

## 多元方差分析

- 14.1 独立样本单因素多元方差分析····· 547
- 14.2 独立样本双因素多元方差分析····· 560
- 14.3 双因素多元方差分析——交互作用显著····· 571

## 典型相关分析

- 15.1 理论基础····· 590
- 15.2 执行程序——MANOVA 语法 ····· 595
- 15.3 报表解析····· 597
- 15.4 结果说明····· 605
- 15.5 用 CANCELL 指令执行典型相关····· 608

## 协方差分析

- 16.1 独立样本单因素协方差分析····· 616
- 16.2 独立样本双因素协方差分析····· 640

## 第 5 篇 信效度分析的理论实务

### 项目分析与试题分析

- 17.1 项目分析····· 660

17.2	试题分析	677
------	------	-----

18

## 因子分析与信度检验

18.1	因子分析	688
18.2	信度分析	713

## 第6篇 矩结构分析

19

## Amos 与结构方程模型

19.1	Amos Graphics 窗口的介绍	726
19.2	估计变量间的相关性	749
19.3	路径分析	755
19.4	线性结构模型的验证	771
19.5	一阶验证性因子分析	787
19.6	二阶验证性因子分析	797

# 第1篇

## 定量研究的 理论与实务



对研究方法可以有所偏好，但不能有偏见。量的研究重视研究工具的信效度、数据的正确核查、适当的统计分析方法以及适宜的解释与归纳推断。

## 1.1 研究设计的基本概念

行为及社会科学 (social science) 领域中，定量数据 (quantitative data) 分析与定性研究 (qualitative research) 是研究的两个主要范畴。定量数据分析受到信息科学进步的影响，数据的处理更简便，也较客观，因而社会科学领域中多数专家撰写研究论文时仍倾向于使用定量研究方法。然而，定量研究与定性研究各有其适用的条件与特色，研究者对其中一种方法不能有偏见，应根据自身研究所长、研究目的所需，挑选适合的研究方法。定量研究与定性研究的特点对照如下表 (Merriam, 1988; 邱兆伟, 1995):

关注焦点	定量研究	定性研究
研究焦点	数量、多少、数目	品质、性质、本质
哲理根源	实证论、逻辑经验论	现象学、符号互动论、自然论
相关术语	实验的、实证的、统计的	田野工作、自然取向、主观的
探讨目标	预测、控制、描述、证实、 假设检验 (验证假设)	理解、描述、发现、形成假设
设计特征	先决的、结构的	弹性的、开展的、较无结构的
研究情景	不熟悉的、认为的	自然的、熟悉的
研究样本	大规模的、随机的、代表的	小规模、非随机的、理论性的
资料收集	无生命的工具，如量尺、 测验、调查、问卷、电脑	研究者作为研究工具、 晤谈、观察
分析范式	演绎的 (借助统计方法)	归纳的 (研究者进行归纳)
研究结果	精确的、细化的	综合的、全面的、扩张的

定量研究常与统计学 (statistics) 相关联，统计学是一门应用数量的方法 (quantitative method) 来搜集、整理、整合、分析和解释研究数据 (data)，使其变成有意义的信息 (information)，并由研究样本 (sample) 的性质来推断未知总体 (population) 性质的科学 (余民宁, 1997)。统计学是搜集、整理、分析及推断数字数据 (numerical data) 的科学方法 (朱经明, 2003)。统计学根据其统计推断，可分为描述统计学 (descriptive statistics) 及推断统计学 (inferential statistics)。描述统计学旨在整理凌乱的数据 (data)，使其变成有意义的信息 (information)，使原始数据变得系统、有组织，从而便于解释。描述统计并没有由样本推断总体或由已知推断未知的含义。简单来说，描述统计包括搜集、整理、表现、分析与解释资料，它讨论如何搜集调查数据，以及将所获得的数据，加以整理、呈现、解释与分析。描述统计学帮助人们了解数据的特性，使人们由数据的特性得到某些结论 (林惠玲、陈正仓，



2003)。

通过样本所搜集的数据结果来推断总体的特性，且附带陈述这种推断正确的可能性和犯错的概率 (probability) 有多大，就是推断统计学 (inferential statistics)，推断统计学又称归纳统计学 (inductive statistics)。推断统计学包括估计与检验 (test，或译为检定)，估计又包含点估计及区间估计。估计或检验的目的是希望由样本的性质，来推断总体的性质。样本的统计指标或量数称作统计量或统计数 (statistic)，一般用英文字母来表示，如  $M$  (表示平均数)、 $SD$  (表示标准差)、 $r$  (表示相关系数)、 $b$  (表示回归系数)；说明或表示总体真实性质的统计指标或量数，一般称作母数或参数 (parameter)，一般用希腊字母表示，如  $\mu$  (表示平均数)、 $\sigma$  (表示标准差)、 $\rho$  (表示相关系数)、 $\beta$  (表示回归系数)。通常，参数是未知的，研究者必须用统计量来估计参数，其估计数值的大小就称为估计值 (estimate)。推断统计的意义可以用图 1—1 表示：

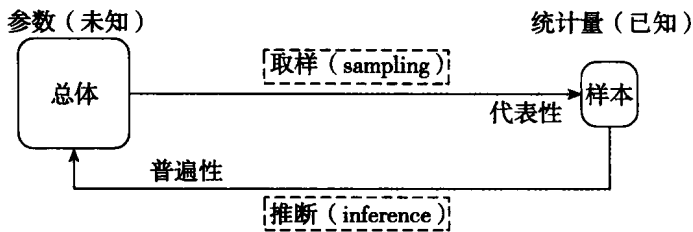


图 1—1

社会科学领域许多所谓的理论，都是根据抽象的普通研究方法建构出来的，各种理论的观点也未必完全一致，因而社会科学领域多用较低层次的理论来解释许多特殊的现象，如投票行为、犯罪行为、攻击行为等。社会科学理论要更具说服力与解释力，就应多以直接实证研究结果为立论基础，但这样的方式，也会使得理论类推受到很大的限制。因而在社会科学理论部分，研究者应多从其他理论文献中获取与研究主题最有密切关系的数据，将其归纳整合为有用信息。理论阶段包括文献资料的搜集、整理与归纳，研究者的研究主题最好与研究者的实务经验、学术理论、相关研究或重要议题有关。

有关行为及社会科学领域的研究设计，心理学家倾向于采用实验设计，社会学家较偏爱采用调查研究或相关研究。不论是实验设计或调查/相关研究，或数据/数据分析，都要以计算机为工具，根据变量属性，选用适当的统计方法来检验假设，资料搜集的常用方法包括访问、问卷调查及观察等。实验研究、调查研究、相关研究或后设分析研究，均为实证研究模式。实验研究的单向直线模式见图 1—2 (这种模式也是行为及科学研究中被研究者采用最多的模式)：当研究采取不同的方式搜集到有效的数据后，要分析数据、验证假设就要采用适当的统计方法，如果统计方法选用有误或对统计结果报表解释错误，则会造成研究的偏误。因而实证或定量的研究与统计方法的应用息息相关。

学者 Kirk (1995) 认为，研究者进行实验研究时，应当把握以下四大原则：

1. 要有适当的统计检验力，使  $\mu$  和  $\mu_0$  间的差异，有实质意义存在。
2. 使用最少的资源、人数。

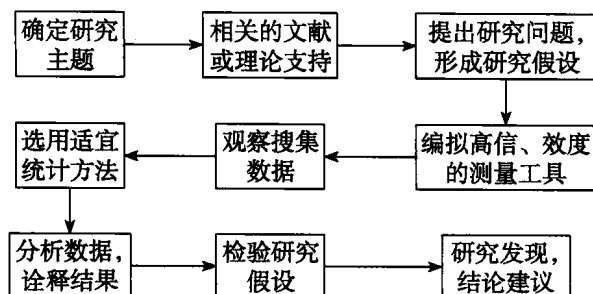


图 1—2

3. 提供适当的保护，以防止第一类错误。
4. 无关干扰变量的影响要最小。

在实验研究情境中，各种变量均可以分析，且彼此独立。实验情境中有自变量与因变量，自变量又称实验变量（experimental variable），是研究设计中可操作的变量；而因变量（dependent variable）则是根据自变量得来的，可对其进行观察或评估，以决定实验效果，其又称为结果变量（outcome variable）或效标变量（criterion variable）。至于干扰变量（extraneous variable），也属自变量之一，但该变量并非实验处理变量，干扰变量的变异可能会影响因变量的变异。干扰变量通常可区分为情境变量（situational variables）与个体变量（individual variables）。情境变量的例子包括时间、实验情境的物理环境、人为的偏见等；个体变量包括个人身心特质的差异、心理态度的感受、动机、智力等。进行实验设计时最重要的就是要把握以下 3 个方面：使实验变量的变异量最大，控制无关干扰变量对因变量变异量的影响，使误差的方差最小（林生传，2003；林清山，2003）。严谨的实验设计应采用真正的实验设计，能利用各种策略去控制实验的情境，减少或排除干扰因素的影响，以提高实验的效度，这种方法采用的是实验控制（experimental control）。但在行为及科学研究中，比较接近实际情境的研究，则是准实验设计，研究者明知干扰因素会影响实验结果，但无法在实验进行时加以排除或控制。这种情况下，只能在实验之后采用统计分析的方法，把影响结果的因素抽离出来，这种方法采用的是统计控制（statistical control），统计控制使用的是协方差分析（analysis of covariance）。不论是实验控制还是统计控制，要得知自变量对因变量变化的影响是否显著，都要运用定量统计分析方法。

在实验设计的选择方面，除把握 Kirk（1995）所提的以上 4 个原则外，还要考虑以下 5 个方面：

1. 这个设计能够有效计算出实验效果与误差效果值，因为由这两个效果值才能得知实验处理是否有效。
2. 所搜集的数据足以产生可靠的结果，如果所搜集的数据有偏误，则统计结果自然不正确。
3. 设计是否可运用适当的统计方法加以检验，而检验结果是否具有足够的统计检验力？定量数据如果不能用统计方法检验，则无法验证。另外，若统计检验力太低，也代表正确判定率不高。
4. 在实验情境限制下，这种设计是否是最有效率、也最富经济化的设计？

方便易行为抽样的主要目的，往往是为了争取时效或达到某种特殊目的，实时进行访谈或实施问卷调查（林生传，2002）。立意抽样与便利抽样，考虑的因素多是时间、受试样本与人力成本。采取这两种抽样方式，由于并不是随机抽样法，在研究推论——外在效度上，要特别留意。便利抽样或立意抽样的原理与程序见图 1—4：

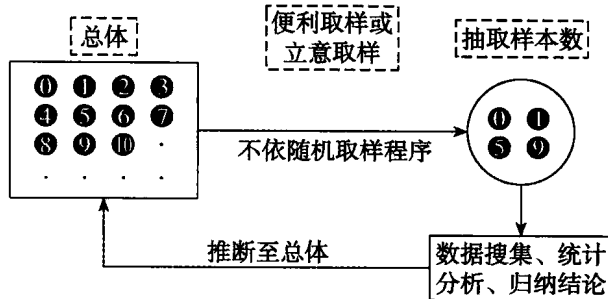


图 1—4

## 2. 分层随机抽样 (stratified random sampling)

研究设计中，如果受试者总体间的差异很大（异质性很高），或某些样本点很少，则为顾及小群体的样本点也能被抽取，采用分层随机抽样比较适宜。在实施上，研究者根据研究关注的准则，先将总体分成互斥的若干层（不同的小群体），使各层间尽可能异质，而各层内尽可能同质，然后从每层中采用随机抽样方式，按一定比例各抽取若干样本。

分层随机抽样适用的条件在于总体的结构比较复杂，总体有很高的异质性。此时如果采取随机抽样方法，恐怕某些主要群组的样本数会很少，如某社会学家想调查不同成年人的婚姻状态与其生活满意度的关系。如果该社会学家根据某一行政区的成人采取随机抽样的方法抽取样本，可能无法真实反映婚姻状态与其生活满意度的关系，因为婚姻状态的组别间（已婚、未婚、离异、丧偶四组）人数差异可能很大，进而会造成统计上的偏误，此时就应采取分层随机抽样方法。分层随机抽样法又可分为比例分配抽样（proportional allocation sampling）与等量分配抽样（equal allocation sampling）。比例分配抽样根据各分层的人数，选取一定的比例样本。如研究者想从各分层中选取 10/100 的样本，甲分层中的总体为 350 人，则甲分层中应随机抽取 35 人，而乙分层中的总体为 400 人，则乙分层中应随机抽取 40 人，抽取的样本数均为各分层总体的 10/100。比例分配抽样又如依照学校班级数将学校分成大型学校（50 个班以上）、中型学校（25 ~ 49 个班）、小型学校（24 个班以下），要抽取 1 000 位小学教师，总体人数为 6 000 人，则分层抽取比例分配如下：

	小型学校 (24 个班以下)	中型学校 (25 ~ 49 个班)	大型学校 (50 个班以上)	合计
学校数 (单位: 所)	20	25	35	80
教师人数	600	1 800	3 600	6 000
抽取教师人数	100	300	600	1 000
依教师人数抽取比例	$\frac{1}{10} = \left(\frac{600}{6\,000}\right)$	$\frac{3}{10} = \left(\frac{1\,800}{6\,000}\right)$	$\frac{6}{10} = \left(\frac{3\,600}{6\,000}\right)$	$\frac{10}{10} = 1$ (1 000 位)

有人会对总体进行普查以搜集资料，因而也就无从了解抽样误差率的精确值（周文钦，2004）。在95%的显著性水平下（ $\alpha = 0.05$ ），一般用下列公式来估计抽样误差率：

$$\text{抽样误差率} = \frac{1}{\sqrt{N}} \quad (N \text{ 为样本大小})$$

举例而言，在调查研究中，有效样本数为800，在95%的显著性水平下，抽样误差率 =  $\frac{1}{\sqrt{800}} \doteq \pm 0.035355 \doteq \pm 0.035 = \pm 3.5\%$ ，因而有效样本数为800，在95%的显著性水平下，其抽样误差率约为 $\pm 3.5\%$ 。当抽样有效观察值增加至1600人时，抽样误差率 =  $\pm 0.025 = \pm 2.5\%$ 。可见，当样本人数变大时，其抽样误差率会相对地变小。当N增至无限大时，抽样误差率会趋近于0（ $= \frac{1}{\sqrt{\infty}} \doteq \frac{1}{\infty} \doteq 0$ ），此时所用的抽样方法即是普查。

在行为科学研究中，有效样本数越多，则抽样所得的数据结果的误差越小，有学者将这种误差称为统计误差（储存滋，1992），统计误差主要源于对总体的定义不当，以及采用不适当的抽样调查并用所得的样本数据来代表总体。因而在定量研究中，选取具有代表性及一定数目的样本非常重要。样本选取时，最重要的是样本要有代表性，才能从受试样本推断到其总体的性质。受试样本要有代表性，最重要的方法就是要把握随机抽样与随机分配的两个原则。

另外，进行研究与数据处理时，应考虑以下几个问题：

### 1.1.1 抽样的样本多大才算够大？

在研究设计中，受试者的个数为多少才算具有代表性？有多少个样本才算够大（large enough）？这个问题，社会科学研究领域中，似乎无一致结论。其中，学者Sudman（1976）提出的看法，可供研究者参考。Sudman（1976）认为：

1. 初学者进行与前人相类似的研究时，可参考别人的样本数，作为自己抽样的参考。

2. 如果是地区性的研究，则平均样本人数在500~1000人之间比较适合；而如果是全国性研究，则平均样本人数在1500~2500人之间比较适宜。

学者Gay（1992）对于样本数的选择，提出以下看法：

1. 进行描述研究（descriptive research）时，样本数最少占总体的10%，如果总体较小，则最小的样本数最好为总体的20%。

2. 进行相关研究（correlational studies）的目的在于研究变量间有无关系存在，受试者至少为30人。

3. 进行因果比较研究（causal-comparative studies）与许多实验研究，各组的人数至少为30人。

4. 如果实验研究设计得当，有严密的实验控制，则每组受试者至少为15人，但

等距变量来设计。这类等距变量也可转化为不同类别，它虽然不是真正的等距变量，但多假定其具有真正等距变量的性质，这样才能进行有意义的数据统计分析，并归纳出合理的结论（Bryman & Cramer, 1997）。

另外，在社会科学领域中，另一个常见的变量为二分变量（dichotomous variables），这种变量被视为名义变量或次序变量，亦属于离散变量（discrete variable，测量数据的属性只能用一种特定的数值来表示，并且无法进行无限分割，又称非连续变量），对象属性只分为两大类，如男、女，及格、不及格等。如果离散变量有三大类或三个水平，则称为三分变量（trichotomous variable）。如就学校规模变量而言，1代表大型规模学校、2代表中型规模学校、3代表小型规模学校；就家庭社会地位而言，可将其划分为高社会地位、中社会地位、低社会地位，上述均为三分名义变量。如果变量包含3个以上类别或3个以上水平的离散变量，则又称为多分变量（multichotomous variable）。而等距或比率变量则应归类为连续变量（continuous variable），在t检验或F统计分析法、回归分析、协方差分析等参数统计法中，因变量都必须为连续变量。

在变量属性的转换上，等距/比率变量可转化为次序变量或名义变量，如按某一分量表示得分高低，将样本分成高分组、中分组、低分组，但次序变量或名义变量不能转换为等距变量/比率变量，如要转换，常以虚拟变量方式出现。有关虚拟变量的转换与应用，请参考多元回归分析章节的内容。

### 1.1.3 统计结论效度

所谓统计结论效度（statistical conclusion validity）是指能正确运用统计方法解释研究结果的程度（周文钦，2004）。定量研究或推断检验常被人批评的一点是玩数字游戏，统计结论效度不高。然而，如果要解释假设性的总体的参数，用其所对应的样本统计值来进行推断检验，在实际的情况中仍有其必要性。因为人力、物力、财力以及实际情况的限制，要进行普查并不容易，也不是多数研究者能力之所及，因而定量研究中的推断检验仍是无可替代的方法（Hinkle & Oliver, 1983）。在定量研究的统计分析中，根据之前的经验法则，不当的统计检验或误用统计方法包括以下几种情况：

1. 在准实验设计中，直接用t检验或方差分析检验实验组与控制组的实验处理效果差异，而没有采用统计控制的协方差分析。

2. 在线性回归分析中，投入过多的背景变量（均属离散变量），且未将背景变量转化为虚拟变量，而是将背景变量照原先数据文件的水平编码投入回归模型中，造成结果分析的偏误。

3. 小样本的调查研究中，可能违反参数统计中总体的基本假定，或因变量为名义变量或次序变量时，没有采用非参数统计而直接使用参数统计。

4. 统计逻辑顺序颠倒，如第一部分先用单因素方差分析，第二部分采用双因素方差分析，而非直接使用双因素方差分析进行检验。

5. 统计检验过分重视其是否达到显著性，对于未达到显著性的部分没有进一步

多重比较 因变量：工作压力 Scheffe 法

(I) 年龄	(J) 年龄	平均差异 (I-J)	标准误	显著性	95% 置信区间	
					下界	上界
1	2	0.23	0.29	0.888	-0.58	1.05
	3	0.62	0.30	0.231	-0.22	1.45
	4	0.93	0.36	0.091	-9.23E-0.02	1.95
2	1	-0.23	0.29	0.888	-1.05	0.58
	3	0.39	0.26	0.543	-0.35	1.12
	4	0.70	0.34	0.233	-0.25	1.64
3	1	-0.62	0.30	0.231	-1.45	0.22
	2	-0.39	0.26	0.543	-1.12	0.35
	4	0.31	0.34	0.541	-0.65	1.27
4	1	-0.93	0.36	0.091	-1.95	9.23E-0.02
	2	-0.70	0.34	0.233	-1.64	0.25
	3	-0.31	0.34	0.841	-1.27	0.64

上述所介绍的统计方法的误用及其正确的统计程序在书中后面的章节均会详细介绍，本节旨在厘清一些统计应用的概念，以便研究者能采用适当的应用统计方法以分析搜集到的数据，提高统计结论的效度。

## 1.2 量表或测验试题质量分析的步骤

在问卷调查法 (questionnaire survey) 或实验研究法 (experimental method) 中，研究或调查工具的编制非常重要。问卷如果编制或选用得当，则研究会更具可靠性与价值性。本章先说明问卷信效度分析的简要步骤，各步骤的操作方法则分别于以后各章节详细说明。

问卷信效度分析的步骤，可简要分述如下：

### 1.2.1 编制预试问卷

在预试问卷的编制或修订上，应根据研究目的、相关文献资料与研究架构等方面加以考虑，如果有类似的研究工具，可根据当时研究的实际情形，加以修订、增删；如果是自己重新编制问卷，则问卷内容应依据研究架构的层面，加以编制。在编制问卷时，应注意：

#### 1.2.1.1 视问卷题目性质增加测谎题或反向题

如果问卷的题目内容过于敏感（如有关当事者婚前性行为的调查研究、有关学生偏差与不当行为的研究等），应在问卷中穿插数个测谎题，以探知填答者是否

据实填答。测谎题旨在判定受试者回答的真实性，在统计分析时不将其纳入分析。另外，在问卷编制中除了正向题目外，还可根据各构念的涵义编制反向题，以测知受测者对问卷题目回答的可靠性。反向题在之后的统计分析中只要反向计分即可（反向计分即重新编码），至于反向题题目的多少，学者们并未提出相关的论点或确切的看法。研究者编制的题目内容重点在于能否真正测量出所要测量的心理特质，或真正有效以达到研究所列的研究目的（反向题重新编码的操作程序，在第5章中有详细介绍）。

### 1.2.1.2 李克特量表以采用四点量表至六点量表法为最佳

态度或心理特质行为量表测量通常采用的是李克特式量表（Likert-type Scale）法，量表填答方式以四点量表法至六点量表法最为研究者所采用。对于应采用几点量表法，学者 Berdie（1994）根据研究经验，提出以下看法，可供研究者参考：

1. 在大多数的情况下，五点量表（points）是最可靠的，选项超过五点，一般人难有足够的辨别力。

2. 三点量表限制了温和意见与强烈意见的表达，五点量表则正好可以表明温和意见与强烈意见之间的区别。

3. 由于人口变量的异质性关系，对于没有足够辨别力的人而言，使用七点量表法，会导致信度的丧失；对于具有足够辨别力的人而言，使用五点量表，又令人有受限的不适感。以上问题至目前还没有一个很好的解释理由，然而通过预试（pretesting），可以发现这些问题的存在。

4. 量表的点数越多，选答分布的情形就越广，方差也会变得越大。有些人认为这种情形在统计检验上会具有很好的区别力，然而这种选答很广的分布缺乏可信度，因此完全没有意义。而且，较大的选答方差，表示也会有较大的抽样误差，就统计资料分析的意义而言，是不该有太大的方差的（彭仁信，1994）。

李克特式量表法，重视其内在一致性程度，这是量表题目两两之间关系强度的函数，也是题目与潜在变量间的关系指标。函数值大小与题目数多少有密切关系，题目数越多，越有可能涵盖所要测量的潜在变量。不过，题目数过多，在实际研究情形中多数会有实际的困难，如受试者时间不允许或造成填答者不用心作答等。学者 DeVellis（1991）对于预试问卷题目数提出以下两点看法，可作为研究者的参考：

1. 如果研究者编制或发展一个正式的测验或量表，以用于其他心理测量，则预试题目数最好是将来所需正式题目总数的3~4倍。

2. 在某些特定内容范围中，有些量表题目的发展不容易。或先前相关的研究显示，这些理念不需要过多的题目即可获得良好的内在一致性，预试量表的题目数约为正式量表题目数的1.5倍即可。如正式量表题目数预计为20，则预试问卷时，其预试题目数大约为30。上述是分量表的题目数，非问卷的总题目数。如果某研究者想研究企业员工组织承诺与组织气氛的关系，则可编制一份工作态度知觉问卷，该问卷包括员工组织承诺与组织气氛知觉两种量表。研究者计划正式问卷题目数以70为限，其中员工组织承诺量表约有30题、组织气氛知觉量表约有40题，则问卷编拟时员工组织承诺量表约有36题（1.2倍）至45题（1.5倍），组织气氛知觉量表约有48题

(1.2 倍) 至 60 题 (1.5 倍), 这是比较恰当的。

## 1.2.2 建构专家效度

问卷或测验的效度是指一份测验或量表能正确测量所要测量的特质的程度, 效度也就是测验内容的正确性、可靠性或有效性。在行为科学研究中, 效度通常分为内容效度 (content validity)、效标关联效度 (criterion-related validity)、建构效度 (construct validity) 与专家效度。在成绩测验中, 大都使用双向细目表来描述测验所要测量特质的内容领域。在态度量表的编制中, 会使用内容效度比 (content validity ratio) 来代表专家判断内容效度程度的量数, 编制好问卷后敦请学者、专家及该领域的实务工作者审核, 以判断构念 (construct) 及所包含的题目内容是否适当, 这种经专家、学者审核判断题目的适当性, 作为初步题目筛选及题目词句语义修饰修改的参考, 称为专家效度。虽然专家效度只是专家进行判断的一致性指标, 而非内容效度本身 (Murphy & Davidshofer, 1994), 但却可作为构念与题目适当性的参考。近年来, 将专家效度作为问卷效度所进行的研究, 越来越普遍。至于专家效度一致性判断指标, 并没有确切的标准, 但总结起来有两个判断依据: 一是根据题目适合被勾选的百分比, 如题目被判定适合的百分比为 70% 或 80% 以上; 二是根据题目适合及题目修正后适合的累积百分比, 如题目被判定适合或修正后适合的累积百分比为 80% 或 90% 以上。

## 1.2.3 实施预试

预试问卷编制完后, 应实施预试, 预试对象的性质应与将来正式问卷要抽取的对象性质相同, 如研究对象是初中生, 则预试的受试者也应是初中生。预试对象人数以问卷中包括最多题目的分量表的 3~5 倍人数为原则, 如调查预试问卷中, 包括 3 种分量表, 每种分量表包含的题目数分别为 40、35、25, 则预试对象数最好介于 120 (40×3) ~ 200 (40×5) 之间。如果样本比较特殊, 则在预试人数的选取上可再酌减一些。

预试时, 抽选样本数应为多大最适宜? 应考虑问卷量表是否进行因子分析。因为进行因子分析时, 会用较大样本进行分析, 比只用较小样本稳定。进行因子分析时, 量表的题目数越多, 及预期要有较多的因子层面的话, 应包括越多的受试者 (DeVellis, 1991)。学者 Comrey (1973) 认为, 如果预试问卷要进行因子分析, 以求其建构效度, 则样本数最好在 300 以上; 如果抽样观察值少于 100, 则不宜进行因子分析。学者 Tinsley 和 Tinsley (1987) 建议: 进行因子分析时, 每个题目数与预试样本数的比例大约为 1:5 至 1:10, 如果受试者总数在 300 以上, 这个比例就不太重要。而学者 Gorsuch (1983) 则建议, 进行因子分析时, 有效样本数最少为量表题目数的 5 倍, 且有效样本数要大于 100。最近, 另一学者 Comrey (1988) 也提出一个观点, 如果量表的题目数少于 40, 中等样本数约为 150, 则较佳的样本数为 200。其观点与 Tinsley 的观点接近, 即量表题目数与预试人数的比例约为 1:5 最适合。如有一份成人生活满意度预试问卷, 内有 3 份量表, 量表一为社会参与量表, 包含 25 题, 量表二为健康状态知觉



量表, 包含 20 题, 量表三为生活满意度量表, 包含 30 题, 则预试人数应在 30 至 150 间, 最佳的预试人数约为 150 (30×5), 次佳的预试人数约为 90 (30×3)。

#### 1.2.4 整理问卷、编号与数据录入

问卷回收后, 应一份一份检查筛选, 对于数据不全或不诚实填答的问卷, 应考虑将其删除; 对于填答时都填同一种答案的, 是否删除, 研究者应考虑问卷题目本身的内容与描述, 自行审慎判断。上述所提的测谎题与反向题可作为判断问卷是否有效的参考指标, 如测谎题中当事人回答背离实际的经验法则, 则该问卷可视为无效问卷, 反向题与正向题的填答相互矛盾, 也可视为无效问卷。当事者未据实回答问卷题目, 会得到“垃圾进、垃圾出”的资料, 即使采用多变量与高等统计, 分析结果的可信度也很低。

筛选完的问卷应加以编号, 以便将来用它们核对数据, 问卷的编码如 0001、0002、…、0989 或 950001、950002、…、950003。之后, 再给予各变量、各题目一个不同的代码, 并按问卷内容, 按顺序输入计算机。其中, 单选题与多选题的编码不同, 如果是单选题, 每个题目应给予一个变量名称; 如果是多选题或排序勾选等级的题目, 则每个选项应给予一个变量名称, 如在高中职业学校行政主管时间管理问卷中的第 5 题为主管时间分配 (陈明华, 2004), 其内容如下:

排列下列项目, 使其最能反映出您平日工作时间的分配情况, 请将数字按顺序填入□内, 时间花费最多的填 1, 其次填 2, 以此类推……排至 8。

- 组织发展 (包括: 拟定校务发展方针、规划学校业务进度、评鉴同事工作绩效、展望教育发展  
与趋势、实现教育目标与政策……)

---

- 行政领导 (包括: 激励员工、处理或避免员工冲突、适当授权及合理决定、分配并监督同事  
工作、主持或参加会议……)

---

- 事务管理 (包括: 经费编列与执行、文书处理与审阅、校园规划与执行、校舍兴建与维护、  
设备购置与管理……)

---

- 教学视导 (包括: 推动教师研究与进修、监督教学计划与过程、沟通教学理念与做法、查阅  
学生作业与试题……)

---

- 学生辅导 (包括: 增进教师辅导意识、督办学生自治活动、规划学生生涯辅导、做好学生咨  
询工作……)

---

- 公共关系 (包括: 拟定公开计划, 与领导、社区人士、家长等保持密切联系, 提供社区参与  
学校活动的机会, 建立学校良好形象……)

---

- 研习进修 (包括: 参与教学有关的进修活动, 参加校内、外进阶成长研习以及自学  
辅导……)

---

- 偶然事件 (包括: 意外事件、非正式访问、临时交办事项、家庭问题、不可抗力事故……)