



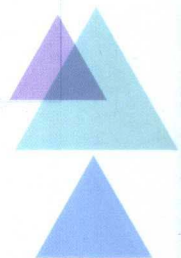
应用统计学：

学习指导、软件介绍 及习题

耿修林 编著



科学出版社
www.sciencep.com



应 用 统 计 学:

——学习指导、软件介绍及习题

耿修林 编著

科 学 出 版 社

北 京

内 容 简 介

本书是学习应用统计学的辅助教材。编写该书的目的,主要是帮助读者理清应用统计学各章的学习要点,并通过做习题加深和巩固所学的内容。

全书共十五章,每章的结构安排是:学习要点,内容提要,计算机软件应用介绍,复习思考题,以及案例分析等。

本书可作为经济、管理、财经、工商专业大学本科生学习应用统计学的辅助材料,也可供其他读者自学用书。

图书在版编目(CIP)数据

应用统计学:学习指导、软件介绍及习题/耿修林编著. —北京:科学出版社,2004

ISBN 7-03-014189-X

I. 应… II. 耿… III. 应用统计学-高等学校-教材 IV. C8

中国版本图书馆 CIP 数据核字(2004)第 083881 号

责任编辑:卢秀娟 王伟娟/责任校对:张怡君

责任印制:安春生/封面设计:耕者设计工作室

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

源海印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2004年8月第 一 版 开本:BS(720×1000)

2004年8月第一次印刷 印张:13 3/4

印数:1—3 000 字数:258 000

定价:20.00 元

(如有印装质量问题,我社负责调换〈路通〉)

前 言

《应用统计学:学习指导、软件介绍及习题》是学习《应用统计学》(耿修林、谢兆茹编著,科学出版社,2004)的配套材料,编写这样一本书,主要是帮助大家理清各章的学习要领,进一步熟悉和掌握相应的统计方法,达到“温故而知新”的目的。

本书每章的结构安排是:

1. 学习要点。在这一部分,我们提纲挈领地指出了每一章需要重点掌握的内容,希望大家在学习的时候,能够做到心中有数,认真领会,并尽可能达到熟练地掌握的目的。

2. 内容提要。在这一部分,我们把各章所介绍的知识做了一个大致的总结,尤其是其中的主要知识点。

3. 统计软件应用。在这一部分,我们将结合例题和案例,对一些比较常用的统计数学软件进行了介绍,目的是希望大家在学习统计学原理和统计方法的时候,能够学会利用计算机的帮助进行数据资料的处理和分析。统计软件非常多,在这里,我们主要介绍:EXCEL、SPSS、MINITAB 等。

4. 复习思考题。在这一部分,我们主要针对各章的内容,编写了相应的练习和复习思考题,希望能通过练习检查学生对内容的掌握和理解的程度有所帮助。

5. 案例讨论。各章几乎都编写了案例讨论题,以培养大家综合运用统计方法解决问题的能力。

对于学习统计学的学生来说,在课后适当地做些练习是必要的,但愿这本辅助性用书,能够给大家的学习提供方便。

耿修林

2003 年 12 月于南京大学

目 录

前言	
第一章 绪论	1
第二章 统计数据的来源与数据质量	6
第三章 数据分布与图示法	13
第四章 频数分布的特征数字	28
第五章 概率原理	50
第六章 随机变量与概率分布	61
第七章 抽样分布	76
第八章 参数估计	91
第九章 参数假设检验	108
第十章 非参数检验	126
第十一章 相关与回归分析	141
第十二章 方差分析与正交试验设计	161
第十三章 时间序列分析	176
第十四章 抽样调查	192
第十五章 统计决策	203
参考文献	211

第一章 绪 论

一、学习要点

在这一章里,我们介绍了统计学的发展历程,现代统计学的学科体系,统计学的概念,统计学的性质与特征,统计方法在相关学科应用中的一些课题,以及开展统计应用需要注意的几个问题。

通过本章的学习,要求大家能够了解以下几方面的知识:

- 1) 现代统计学的总体体系。
- 2) 统计学的含义。
- 3) 统计学的性质。
- 4) 开展统计应用需要注意的几个问题。

二、内容提要

(一) 统计学的发展回顾

1. 创立与形成阶段

统计是适应社会经济管理的需要而产生和发展起来的,作为一项重要的社会实践活动,统计的历史非常久远。但作为一门学科的统计学,它的出现却晚得多。如图 1.1 所示。从事统计学史研究的人多数认为,统计学大概兴起于三个半世纪以前。17 世纪中叶,在英国首次出现了有意识地用数字语言说明问题,注重从数量角度探索客观事物变化规律的研究活动。W. 配第《政治算术》(1676)一书的问世,标志着统计学的开始。J. 格朗特是“政治算术”统计学另一位重要的创立者,他的代表性著作《关于死亡表的自然观察与政治观察》(1662),通过对人口变动数据的分析,揭示了一系列人口变化的统计规律。

与“政治算术”产生的时期差不多,在德国也出现了一门统计学——“国势学”。“国势学”主要研究“国家的有关显著事项”。H. 康令是“国势学”的奠基人,他对“国势学”的研究目的、研究对象、研究方法等基本问题,作了深入细致的阐述。G. 阿肯瓦尔在继承康令开创的研究体系和研究方法的基础上,全面发展了“国势学”,他的《欧洲主要国家国势学纲要》(1749)与康令的《国势学讲义》(1730),被奉为“国势学”的经典文献。“国势学”最大的特点,是它较少使用数字工具,而主要诉诸于文字记述和逻辑比较。

统计学另一个重要的起源是概率论,14 世纪,在工商业比较繁荣的意大利以及地中海沿岸其他地区,由于赌博游戏盛行和保险活动的萌起,人们已经对“机会”问题发生了兴趣。但真正意义上的概率论,是从 17 世纪开始的。B. 帕斯卡和 P. 费马关于“得点问题”的讨论,奠定了概率论的基础。

“政治算术”研究的是简单的、确定的数量关系,而概率统计则研究复杂的、随机性现象,这极大地充实和深化了数量问题研究的内容。以概率论为基础,统计学进入了一个新的发展时期。

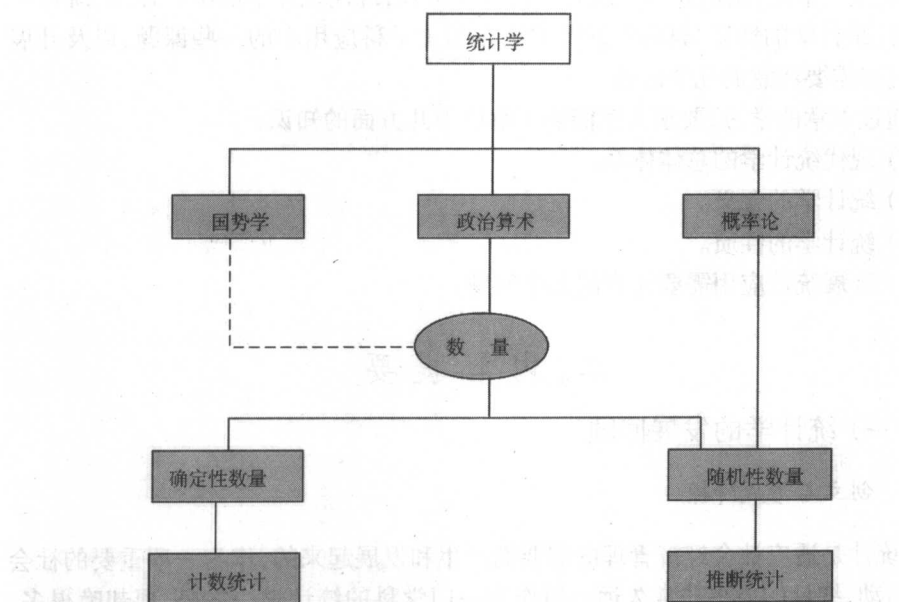


图 1.1 统计学的来源

2. 统计学的“凯特勒时代”

A·凯特勒是统计学发展史上承前启后的重要人物,他一生撰写了大量的有关统计学方面的著作,代表性的就有《社会物理学》(1869)、《概率论书简》(1846)、《犯罪学》(1831)、《比利时统计研究》(1829)等。凯特勒在统计学上的突出贡献,是他把概率论全面引进“政治算术”、“国势学”以及其他社会问题的研究,从而大大推动了概率论和数学方法在社会科学领域的应用,促进了数量研究由“算术”水平向“数理”阶段的迅速转化。

3. 生物统计学

从 19 世纪后半期起,兴起于生物遗传学、农业田间试验领域的统计数学方法

的发展取得了引人瞩目的成就。F. 戈尔登是生物统计学的主创者,他先后提出了“百分位数”、“中位数”、“四分位数差”、相关与回归等重要的统计概念。K. 皮尔逊是生物统计学的集大成者,他系统发展了戈尔登的相关与回归理论,研究了复相关和偏相关,把物理学上“矩”的概念移植到统计学中,给出了极大似然估计方法,导出了重要的 χ^2 -分布。以 K. 皮尔逊为代表,以大量观察和正态分布为基础的关于总体分布曲线的研究,确立了“大样本”统计理论,奠定了“描述统计学”的基本框架。

4. 统计学的“费暄时代”

进入 20 世纪,统计学的发展出现了新一轮飞跃性成就。一方面新的统计思想和统计方法大量涌现,另一方面带有归纳性质的统计推断逐渐占据了主流地位。1908 年, W. 戈塞特导出了 t -分布,以此为标志,统计学逐渐实现了由以前的“描述统计”阶段向“推断统计”阶段过渡,由大样本统计向小样本理论的转变。R. 费暄是推断统计学的建立者,他开辟了方差分析、试验设计等统计分支,给出了戈塞特 t -分布的简洁证明,论证了相关系数的抽样分布,提出了 t -检验、 F -检验、相关系数检验,并编制出了相应的检验概率表。费暄的代表著作包括《供研究人员使用的统计方法》(1925)、《试验设计》(1955)、《统计方法与统计推断》(1956)等。费暄在统计学发展史上有着辉煌崇高的地位,对现代统计学的发展几乎作出了全方位的贡献,正因为如此,有人把 20 世纪前半期称为统计学的费暄时代。

5. 现代主流统计学的发展

J. 内曼和 E. 皮尔逊是费暄之后的杰出统计学家,他们两人经过多年的合作,共同完善了现代统计学的核心内容——区间估计和假设检验的理论。20 世纪 50 年代, A. 瓦尔德提出了“统计决策理论”和质量检验的“序贯分析”,进一步推动了统计学研究和应用的范围。随着计算机在统计中的应用,由 J. 威夏特、H. 赫特林等人发展起来的多变量统计,又重新活跃起来。50、60 年代以来,稳健统计、时间序列、抽样理论、统计诊断、探索性分析、贝叶斯统计等,皆取得了重要的进展。

总之,统计学的发展,经历了从意义和概念不甚明确的阶段,到作为一门独立学科的转化,从数量研究的“算术”水平,到需要较高数学知识的“数理”阶段的转化,从确定性问题的研究,到随机现象研究的转化,从大量观察消除误差干扰以达到对客观现象规律认识的大样本理论,到控制试验次数提高数据质量的小样本推断的转化。统计学就是这样一步步发展起来的,现在它仍然充满着无限的活力。

(二) 统计学的学科体系

统计学的产生虽然可以追溯到 17 世纪,但它的快速发展却是自凯特勒之后出

现的。目前,统计学已经形成了由若干个分支组成的庞大的学科体系。从理论和实践的角度看,统计学可以划分为理论统计学与应用统计学。其中,理论统计学是指,侧重于从数学学科中吸取营养,研究统计学的数学方法和基础原理,以解决统计学自身发展中的重大问题为目标的统计学研究分支。应用统计学就是运用统计思想和方法,处理实际中属于统计方面的问题。在统计学的学科体系中,理论统计学是核心。只有理论统计研究的扩大和深入,才能更好地促进统计学的应用。当然,广泛的统计实践,必将为统计理论准备更多的素材,从而推动理论统计学的进一步发展。

(三) 统计学的概念

统计学是关于资料搜集、整理、分析和解释的科学。

统计学是研究客观现象数量方面的,要达到这个认识目的,首先要搜集能够反映或说明客观现象的数字资料,这是统计活动首要的、基本的环节。统计整理在统计活动过程中,处于承上启下的位置。一方面它是统计搜集资料工作的继续,另一方面又是统计分析的前提,能够为统计分析准备系统的资料形式。统计整理,就是运用各种合适的形式展示和表述统计资料。统计分析是从已有的数据资料出发,提取有关综合信息的过程,其目的是要使一个概括的、全面的数量描述能够形成。在统计分析的基础上,导出一般性结论的过程,就是统计解释,它往往涉及到如何依据样本观察,对同类大量现象作出估计、检验、预测等问题。完整的统计活动过程,一般地就是由资料搜集、整理、分析和解释这四个部分组成的。

(四) 统计学的性质与特点

现代统计学的性质和特征,主要表现在以下几个方面:

- 1) 统计学是研究数量问题的学问。
- 2) 统计学研究的是客观现象总体的数量。
- 3) 统计学主要研究不确定性现象。
- 4) 统计方法带有归纳推断的特点。

(五) 统计学的应用

统计应用的领域极其广泛,在运用统计方法解决实际问题的时候,需要注意以下几个问题:

- 1) 搞统计应用需要具备两方面的知识,一方面要掌握统计学的基本方法,另一方面要有实际问题的知识背景。
- 2) 统计应用不能被简单地当成对现成的统计方法的生搬硬套,统计应用也是科学研究的过程,同样需要创新性思维。

3) 在用样本资料进行推断的时候,统计方法并不总是能保证不犯错误。因此,对统计估计和检验误差应有一个正确的认识。

4) 每一种统计方法,都有各自特定的用途,在使用它们之前,应先搞清楚它们适用的对象和条件。要针对不同的问题,不同的资料,有选择地运用不同的处理方法。

5) 统计资料是统计分析的原料,资料准确可靠,方法运用得当,才有可能得到正确的统计结论。

三、复习思考题

1. 统计学的三个来源是什么?
2. 统计学的含义是什么?
3. 现代统计学的性质和特征是什么?你是怎样理解的?
4. 运用统计方法解决问题需要注意哪些方面的要求?

第二章 统计数据的来源与数据质量

一、学习要点

统计是依据数字资料进行反映和认识的。从统计应用的角度来看,不掌握数字资料,统计活动便无从下手。因此,怎样从研究对象中获得必要的数字资料,是统计活动一项非常重要的任务。取得统计数据有多种途径,但概括起来不外乎是直接方式和间接方式,其中属于直接方式的有统计调查(普查、抽样调查等)和统计试验。在这一章中,我们首先大致介绍了统计数据搜集的几种比较常用的方法。

仅有数字资料还不够,进一步地,还必须关心我们所拥有的数字资料的质量,因为数字资料的准确性和有效性是统计工作的生命。有鉴于此,在这一章里,我们就统计数据质量范畴,统计数据质量的特点,影响统计数据质量的因素,以及统计数据质量检查办法等问题,进行了比较系统的介绍和讨论。

另外,在这一章中,我们还介绍了统计数据的测量尺度,以及统计数据的分类等问题。总之,通过本章的学习,希望大家能重点掌握和了解如下几方面的内容:

1) 常用的统计调查方法。包括:每一种调查方法的含义、特征,各种调查方法适用的场合,如何科学地使用各种调查方法等。

2) 数据资料间接来源的主要渠道。最好能集合专业课程的学习,熟悉一下与本专业有关的大型统计数据的主要载体,尤其是,要懂得如何利用 INTERNET 获得研究数据资料。此外,还需要注意怎样评估和识别间接来源资料的可利用价值。

3) 统计数据的测量尺度及其基本分类。统计数据共有四类测量尺度,即定类尺度、定序尺度、定距尺度和定比尺度。在这一节的学习中,要注意掌握不同测量尺度的统计数据的特征,以及它们适用的统计处理方法等。

4) 统计数据的质量。认识统计数据质量的重要性,了解统计数据质量的基本内涵,识别影响统计数据质量的基本因素,掌握统计数据质量的评估方法。

二、内容提要

(一) 统计资料的直接来源

直接获取第一手统计资料的主要手段包括:统计调查和试验设计。

1. 统计调查

统计调查是获取数据资料的根本手段之一,常用的统计调查方法主要有:普查、抽样调查等。

(1) 普查

普查是专门组织的、一次性全面调查。它主要用来搜集某一时点或一定时期内现象总量资料的。普查是一种重要的搜集全面调查资料的统计手段,通过普查所得到的资料,在了解国情国力,制定社会发展规划,确定重大决策等方面,发挥着重要的作用。

现代普查活动一般具有下列几个特征:定期举行,依法进行,应用广泛,质量保障,内容系统,计算机介入。

对一些重大的社会问题的调查,普查依然发挥着不可替代的作用,但是,我们也要注意普查也存在着许多局限性,表现在:人力、物力和财力耗费过大,时间持续较长,不易组织实施,适用的场合非常有限等。

(2) 随机抽样调查

随机性抽样调查,习惯上直接称之为抽样调查,它是一种最常用、最重要的获取统计资料的手段。所谓抽样调查是指,按照随机性原则,从统计总体中抽取一部分单位组成样本,然后根据样本调查的结果对总体情况进行推断。

抽样调查的特征:按照随机性原则确定观察单位,根据部分单位的调查结果对总体进行科学推断,抽样误差可以计算。

2. 试验设计

统计中的试验设计是科学试验研究的组成部分之一,但与通常意义上的科学试验又有一定的差别。统计试验不涉及试验室设备,也不是去测量某一物理量的绝对值,或验证某一现象是否绝对存在。统计试验,首先是产生数据的计划,然后使用统计方法对这些数据进行分析,目的是要从数据的考察中发现人们感兴趣的因素对试验结果的影响程度,以便采取相应的行动方案。

作为统计范畴的试验即试验设计,包括五个相互关联的环节,分别是:方案设计,方案实施,数据采集,数据分析,优化生产。

(二) 统计资料的间接来源

凡不是通过直接的统计调查和试验,而是从其他各种渠道搜集第二手资料,我们把它总称为统计资料的间接来源。

间接资料的来源大体包括:统计年鉴,统计摘要,统计资料汇编,统计台账,统计公告,报纸,杂志,网上资料等。

利用间接资料需要注意:

- 1) 结合研究的目的,有针对性地获取资料。
- 2) 估计间接资料的可用价值。
- 3) 注意指标的含义、口径、统计方法是否具有可比性。
- 4) 注意弥补缺少资料。
- 5) 进行质量检查。

(三) 统计数据的测量尺度

在统计调查和试验中,测量是获得数据的基础性工作。所谓测量就是对客观现象的研究特征指派数字的活动,指派的规则不同便会产生不同种类的测量结果。从统计的观点看,所有指派规则将会产生四类测量数据,即定类数据、定序数据、定距数据和定比数据,它们对应的测量水平,分别称为定类尺度、定序尺度、定距尺度以及定比尺度。

1. 定类尺度

这是一种最低级别的测量水平,它的主要特征是采用文字、数字代码和其他符号对事物进行简单的分类或分组。如果统计测量仅建立在对事物进行分类分组的基础上,则这样的测量尺度称为定类尺度,也可称为列名尺度。

需要注意:① 定类尺度资料对各个类别之间的关系,不作任何假定;② 分类或分组应做到穷尽,并且类别间要互不重叠。

定类尺度统计资料,形式上具有对称性和传递性。

2. 定序尺度

语义上表现出明显的等级或顺序关系的定类尺度,称为定序尺度。

定序资料的特征:① 仍然属于定性资料,但与定类资料相比含带的信息量有所增加;② 定序尺度的具体意义,可用带有比较级意味的词语来表示;③ 在定序尺度之间,能够进行差别和好坏的比较,但差别或好坏的程度,仍然无法给出具体的概念;④ 定序尺度保留了定类尺度的传递性,但不再具有对称性;⑤ 在定序尺度中,保序变换不会改变定序资料的信息。

3. 定距尺度

具有定序尺度所有特征,并且任意两个数之间可以比较差距大小的测量尺度,就是定距尺度。

定距尺度是最高级别的测量水平,所有常用的统计方法包括参数方法和非参数方法,都可以用于定距资料的分析研究。

4. 定比尺度

具有定距尺度所有性质特征,并且有一个绝对原点的测量尺度,就是定比尺度。

定比尺度具有下列四个特征:(1) 对称性;(2) 大于或小于关系;(3) 任两个数据间距离相比有意义;(4) 任两个数值相比也有意义。

一般地讲,适用于定距和定比尺度的统计方法,主要是参数统计方法,而定类和定序尺度的资料主要使用非参数统计方法。

(四) 统计数据质量问题

1. 数据质量的一般性解释

数据质量可以从误差的角度进行评价,在此,我们把统计数据质量解释为,获得的观察值与客观现象实际数量水平之间的离差。其描述性的计算公式为:

$$\text{误差} = \text{观察值} - \text{相对真值}$$

2. 数据质量的相对性及影响因素

统计数据误差是无处不在的,对此,我们也无须感到悲观。统计数据质量的最大特征在于它的相对性。例如,一个地区有 500 万人口,少计几千、几万,甚至十几万,没有必要过多地计较,全世界陆地总面积 14 900 万平方公里,恐怕谁也不会要求精确到几位小数。一块田地实测到的面积是 4.8 亩,假定它的准确面积应为 5 亩,但方案只要求四舍五入取整,那么就没有误差可言了。统计数据质量的相对性表明,为了减少或降低数据误差,在不影响对现象数量反映的前提下,可以通过修改统计活动方案中的某些要求达到目的。

统计数据质量误差普遍存在,导致误差产生的原因可分为三大类:① 由于统计调查准备工作不充分而引起的误差,在这一阶段,可能容易产生的问题是:对正确了解调查对象有帮助的某些重要的特征被忽视,概念的提法和定义不妥当,调查项目和统计指标界定不清晰,调查表的格式可能难以填写等。② 数据搜集阶段发生的误差,在这一阶段可能产生误差的情况有:调查员不胜任工作,对调查意图和发出的指令理解不透不准,在工作中掺进个人的想法和意见,误导被调查人偏离方案要求作出回答。访问和沟通方式生硬,从被调查者角度看,可能担心个人的资料被用于统计以外的目的,由于后怕心理作祟故意给出不实回答,此外,还有被调查人员的知识背景、社会背景、情感背景等问题。在实际搜集资料阶段,如果使用的搜集资料的具体方式不同,也会产生差别很大的结果,如果使用的测量工具有问题,要获得准确的数据也是不可能的。③ 资料整理过程中产生的误差,数据处理

阶段同样还会发生许多新的差错,比如编码、打孔、录入、分类、汇总、计算等过程中出现的错误。

3. 统计数据质量检查

如何评估统计数据误差的严重性,是什么原因引起了误差,有无修正的必要和可能,这一系列的活动,统称为数据质量检查。

统计数据质量检查,主要有两大类方法,即后验技术和抽样方法。

(1) 数据质量检查的后验技术

后验技术是在调查工作已经完成,进入数据编辑和整理时所用的评估数据质量的方法。这种方法的特征是,不再亲临调查现场,而是通过逻辑关系分析、计算,以及将调查数据与独立来源的资料对比,以确定调查数据的质量。

常用的后验检查技术包括:逻辑关系分析法,计算比较法,设置疑问框法,与独立来源数据对比法。

运用后验技术检查数据质量时需要注意:后验技术的使用范围比较有限;所有的后验技术都仅适用于对最后调查结果的检查,不能用于单项数据误差的评估;使用后验技术一般皆要求要有关于同一研究主题大量的统计资料;作为对比、推算基础的资料必须准确;后验技术检查的结果,只反映在质量上,不可能提供新的统计数据。

(2) 数据质量检查的抽样方法

所谓抽样检查方法是指,在一次调查之后,紧接着再从这些被调查单位中抽取一定数量的单位组成样本,经过重新登记,最后将二者的结果进行对比,以检查先前调查数据的质量,并进行适当的调整。

抽样方法有许多优点,主要表现在:数据质量检查完全依据样本资料,因此不管有没有相关的统计资料可依,无论过去是否作过类似的调查,都不会影响现在的数据质量评估;抽样检查法运用起来十分灵活,适用于各种场合的调查数据检查,也适用于调查数据各部分的检查,不仅可以帮助我们了解数据的质量水平,还可以帮助我们寻找发生误差的原因;抽样检查法有助于提高数据的质量;抽样法不受调查项目之间关系的任何假设限制;运用抽样检查法,可以帮助统计人员获得更多的资料;借助抽样检查方法可以估算数据误差的大小,从而对原调查资料进行适当的修正和调整等。

三、复习思考题

1. 什么是普查,组织统计普查活动需要注意哪些问题?
2. 什么是抽样调查,它有什么特征?

3. 说明抽样调查和统计普查之间的异同点,并指出为什么在某些情况下抽样调查比普查更可取。

4. 利用间接来源的统计资料,需要注意哪些问题?

5. 你是怎样理解统计数据质量的? 如何检查统计数据的质量?

6. 为了解校园文化生活的开展情况,学校学工处决定进行一次摸底调查,请你想一想,可能会影响调查结果客观性的因素有哪些?

7. 假定让你组织一次某地区的居民家庭收支情况调查,你会采取哪些措施来保证调查数据的质量?

8. 统计调查和统计试验设计都是直接取得数据资料的重要手段,试比较一下它们二者之间的差别?

9. 下列资料是 10 个公司的 CEO 报酬、所在行业、年销售额以及 CEO 报酬对股票持有人回报的排名,其中 CEO 报酬对股票持有人回报排名 1 表明,该公司位列 CEO 报酬对股票持有人回报比率最好的公司之中,排名越差表明 CEO 报酬对股票持有人回报率越糟糕。

CEO 报酬及有关资料

公司名称	CEO 报酬	所在行业	销售额	排名
银行信托	8 925	银行	9 565	3
可口可乐	2 437	饮料	18 546	5
大众磨房	1 410	食品	5 567	1
LSI 逻辑	696	电力	1 239	2
摩托罗拉	1 847	电子	27 973	4
读者文摘	1 490	出版	2 968	3
塞 尔 思	3 414	零售	38 236	4
雪 碧	3 344	通信	14 045	4
瓦尔格林	1 490	零售	12 140	2
威 法 果	2 861	银行	8 723	3

资料来源:商务周刊,1997 年 4 月 21 日。

试根据上述资料回答下列问题:

(1) 这个数据集中共使用了哪几种测量尺度? 分别是什么?

(2) 哪些是定性资料? 哪些是定量资料?

10. 指出下列变量可能需要采用的测量尺度:

(1) 年龄 (2) 学历 (3) 社会等级 (4) 汽车牌号

(5) 利润率 (6) 就业率 (7) 生产规模 (8) 雇员所在的岗位

(9) 每股收益 (10) 支付方式 (11) 总成本 (12) 交易场所 (13) 出生地

(14) 个人社会保险代码 (15) 上班途中花费的时间

11. 《商务周刊》北美订阅者研究会搜集了 3200 个订阅者的样本资料,其中

59%的人反映他们的年收入在 75 000 美元以上,另有 50%的人声称他们拥有专门的信贷卡。试根据这些资料回答:

- (1) 年收入是何种性质的变量?
- (2) 持有专门信贷卡是属性变量还是数量变量?

12. 取得下列调查项目的资料,需要采用何种测量尺度:

- (1) 钢丝的抗断强度
- (2) 城市中停车标志的数目
- (3) 办公室秘书接答一个电话所需要的时间
- (4) 产品是否存在缺陷
- (5) 一场高尔夫球赛中平均击球码
- (6) 人的身材
- (7) 疾病严重程度
- (8) 药品的疗效

13. 某公司欲检验一种新型电视商品广告的效果,先在地方电视台的新闻节目中进行播放,两天之后委托一个市场研究所通过电话询问,了解人们对该商品广告的印象。对于这项调查,为什么要使用抽样调查方法?具体谈谈你的理由。

14. 一家公司的经理为挽留某位员工,决定给这位员工进行加薪,初步定为 10000 元,如果要评估这样的加薪幅度是否合适,可能需要用到哪些统计方面的资料。

15. 据说,A 型血和 B 型血的人,在工作业绩方面会存在一定的差异。原因是 A 型血的人,在工作中有时会显得心情暴躁、说话很快、雄心勃勃、缺乏耐心和对人常怀敌意等,而 B 型血的人则刚好相反。为检验这个说法是否正确,某研究机构特地做了一次调查,以下资料就是该项调查的一小部分:

编号	血型	年龄	职务	下属人数	业绩得分
1	A	45	高层	202	3
2	B	37	中层	34	5
3	B	42	高层	105	2
4	A	51	低层	5	1

试指出这一调查涉及的调查变量,哪些是定性的哪些是定量的。

16. S 城市的房地产开发协会定期发布“住宅建设商情指南”,其中包含了 5 个变量,分别是:城区、县郊区、住宅类型、销售价格、建筑面积。说一说这些变量可能产生的数据类型。