



全无敌·经典教材配套丛书

配套·人大社·贾俊平《统计学（第六版）》

TONGJIXUE  
TONGBU FUDAO YU XITI QUANJIE

人大社·贾俊平·第六版

# 统计学

## 同步辅导与习题全解

龚秀芳 宁同科 朱 敏 / 编著



华东理工大学出版社  
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS

全无敌·经典教材配套丛书

数据(CIP)目录题录并图

原·平刘贾·社大社·全圖区計导輔全同学书第  
照中·生一·香余類求·符同宁·农秀芳\第  
# 3102·社理出学大工

(第法分·計从分類目錄典登)

ISBN 978-7-311-32207-1 · 079 28821

# 统计学同步辅导与习题全解 (人大社·贾俊平·第六版)

龚秀芳 宁同科 朱敏 编著



华东理工大学出版社  
EAST CHINA UNIVERSITY OF SCIENCE AND TECHNOLOGY PRESS

· 上海 ·

## 图书在版编目(CIP)数据

统计学同步辅导与习题全解:人大社·贾俊平·第六版 / 龚秀芳,宁同科,朱敏编著. —上海:华东理工大学出版社,2015.6

(经典教材配套丛书:全无敌)

ISBN 978-7-5628-4280-4

I. ①统… II. ①龚… ②宁… ③朱… III. ①统计学—高等学校—教学参考资料 IV. ①C8

中国版本图书馆 CIP 数据核字(2015)第 108453 号

全无敌·经典教材配套丛书

## 统计学同步辅导与习题全解(人大社·贾俊平·第六版)

编 著 / 龚秀芳 宁同科 朱 敏

责任编辑 / 焦婧茹

责任校对 / 金慧娟

封面设计 / 裘幼华

出版发行 / 华东理工大学出版社有限公司

地 址:上海市梅陇路 130 号,200237

电 话:(021)64250306(营销部)

(021)64252344(编辑室)

传 真:(021)64252707

网 址:press.ecust.edu.cn

印 刷 / 上海展强印刷有限公司

开 本 / 787 mm×1092 mm 1/16

印 张 / 16.5

字 数 / 390 千字

版 次 / 2015 年 6 月第 1 版

印 次 / 2015 年 6 月第 1 次

书 号 / ISBN 978-7-5628-4280-4

定 价 / 39.00 元

联系我们:电子邮箱 [press@ecust.edu.cn](mailto:press@ecust.edu.cn)

官方微博 [e.weibo.com/ecustpress](http://e.weibo.com/ecustpress)

天猫旗舰店 <http://hdlgdxpbs.tmall.com>



# 前 言

统计学是关于数据的科学,它所提供的是一套有关数据收集、处理、分析、解释并从数据中得出结论的方法,统计研究的是来自各个领域的数据。统计学的分析方法主要分为描述统计方法和推断统计方法,其应用领域几乎覆盖了社会科学和自然科学的各个领域。现今社会正处于大数据时代,各种数据正在迅速膨胀,人们越发重视对数据信息的处理。通过对大量数据的研究,我们可以发现数据背后的规律以及存在的价值。因此,现代社会离不开统计学,统计学的分析方法已经成了各学科领域研究者和实践工作者的必备知识。

本书是为了配合贾俊平、何晓群、金勇进编著的《统计学》(第六版)而编著的学习辅导书。全书每一章都提出了学习目标,对基本知识点进行了梳理,对重难点进行了解析,并对每一章后面的思考题和练习题给出了详细的解答,最后编制了三套自测题,使学习者能理解并掌握相关的统计学分析方法,并能熟练地应用到实践中去。

本书既可作为高等院校经济、管理类各专业本科生学习统计学课程的辅导用书,又可作为成人教育、函授大学以及其他管理学科相关专业和干部培训的参考书,对广大实践工作者也具有一定的参考价值。

参加本书编撰的三位教师都长期从事统计学教学与科研工作。本书的第1~4章和第14章由龚秀芳执笔;第5~8章由宁同科执笔;第9~13章由朱敏执笔。全书由龚秀芳统稿。

上海师范大学商学院学生朱晓隽、何秋辰、潘奔、徐静婷、贾光华等也参与了部分习题答案的编辑工作,为本书的完成付出了辛勤的劳动,编者在此深表感谢。同时,本书的编著参考了大量的文献资料,在此,编者对相关作者也深表感谢。

由于作者水平有限,书中的不足之处在所难免,敬请读者提出宝贵意见。

编 者

2015年3月

# 目 录

<b>第 1 章</b>		<b>章 3 美</b>
	导论 .....	1
1.1	学习目标 .....	1
1.2	基本知识点 .....	1
1.3	重难点解析 .....	1
1.4	思考题全解 .....	4
1.5	练习题全解 .....	6
<b>第 2 章</b>		
	数据的搜集 .....	8
2.1	学习目标 .....	8
2.2	基本知识点 .....	8
2.3	重难点解析 .....	9
2.4	思考题全解 .....	11
<b>第 3 章</b>		
	数据的图表展示 .....	14
3.1	学习目标 .....	14
3.2	基本知识点 .....	14
3.3	重难点解析 .....	15
3.4	思考题全解 .....	17
3.5	练习题全解 .....	19
<b>第 4 章</b>		
	数据的概括性度量 .....	36
4.1	学习目标 .....	36

4.2	基本知识点	36
4.3	重难点解析	37
4.4	思考题全解	40
4.5	练习题全解	42
<b>第5章</b>		
	概率与概率分布	52
5.1	学习目标	52
5.2	基本知识点	52
5.3	重难点解析	53
5.4	思考题全解	62
5.5	练习题全解	62
<b>第6章</b>		
	统计量及其抽样分布	68
6.1	学习目标	68
6.2	基本知识点	68
6.3	重难点解析	69
6.4	思考题全解	71
6.5	练习题全解	73
<b>第7章</b>		
	参数估计	75
7.1	学习目标	75
7.2	基本知识点	75
7.3	重难点解析	75
7.4	思考题全解	80
7.5	练习题全解	81
<b>第8章</b>		
	假设检验	97
8.1	学习目标	97
8.2	基本知识点	97

8.3	重难点解析	98
8.4	思考题全解	103
8.5	练习题全解	105

<b>第9章</b>		
	分类数据分析	111
9.1	学习目标	111
9.2	基本知识点	111
9.3	重难点解析	112
9.4	思考题全解	114
9.5	练习题全解	115

<b>第10章</b>		
	方差分析	120
10.1	学习目标	120
10.2	基本知识点	120
10.3	重难点解析	121
10.4	思考题全解	130
10.5	练习题全解	131

<b>第11章</b>		
	一元线性回归	142
11.1	学习目标	142
11.2	基本知识点	142
11.3	重难点解析	143
11.4	思考题全解	149
11.5	练习题全解	152

<b>第12章</b>		
	多元回归分析	172
12.1	学习目标	172
12.2	基本知识点	172
12.3	重难点解析	173



12.4	思考题全解	180
12.5	练习题全解	182

**第 13 章**

<b>时间序列分析和预测</b>		193
------------------	--	-----

13.1	学习目标	193
13.2	基本知识点	193
13.3	重难点解析	194
13.4	思考题全解	198
13.5	练习题全解	199

**第 14 章**

<b>指数</b>		215
-----------	--	-----

14.1	学习目标	215
14.2	基本知识点	215
14.3	重难点解析	216
14.4	思考题全解	223
14.5	练习题全解	225

<b>自测题及解答</b>		233
---------------	--	-----

<b>参考文献</b>		255
-------------	--	-----



# 第1章

# 导论

## 1.1 学习目标

- 理解统计的含义及统计的应用领域；
- 掌握统计数据的类型；
- 学习并掌握统计学中的几个基本概念。

## 1.2 基本知识点

表 1.1 基本知识点梳理

章节	主要内容	学习要点
1.1 统计及其应用领域	统计学的含义 统计分析方法 统计的应用领域	<ul style="list-style-type: none"> <li>◆ 统计学的概念</li> <li>◆ 描述统计和推断统计的含义</li> <li>◆ 统计在工商管理中的应用</li> </ul>
1.2 统计数据的类型	按计量尺度分类 按收集方法分类 按时间状况分类	<ul style="list-style-type: none"> <li>◆ 分类数据、顺序数据和数值型数据的概念</li> <li>◆ 观测数据和实验数据的概念</li> <li>◆ 截面数据和时间序列数据的概念</li> </ul>
1.3 统计中的几个基本概念	总体和样本 参数和统计量 变量	<ul style="list-style-type: none"> <li>◆ 总体的概念；样本和样本量的概念</li> <li>◆ 几种常用的总体参数和样本统计量</li> <li>◆ 变量、变量值的概念；变量的分类</li> </ul>

## 1.3 重难点解析

### 1. 统计数据的计量尺度

统计数据是对客观现象进行计量的结果。按照由低级到高级、由粗略到精确的标准，统计数据的计量尺度分为分类尺度、顺序尺度、间隔尺度和比率尺度。

(1) 分类尺度。将数据作为现象总体中不同类别或不同组别的代码，这是最低层次的尺度。分类尺度是按照事物的某种属性对其进行平行的分类，数据表现为类别，即不同的数据仅表示不同类(组)别的品质差别，而不表示它们之间量的顺序或量的大小。分类尺度的主要数学特征是“=”或“≠”。

例如,将国民经济按其经济类型,可以分为国有经济、集体经济、私营经济、个体经济等类,并用(01)代码表示国有经济、(02)代码表示集体经济、(03)代码表示私营经济、(04)代码表示个体经济等。不同代码反映的是不同的类(组)别,并不反映其大小顺序。

(2) 顺序尺度。顺序尺度是对事物类别顺序的尺度,数据表现为有序的类别。顺序尺度可以用数据表示不同类(组)别,也可以反映大小顺序关系,从而可以列出各单位、各类(组)的次序。顺序尺度的主要数学特征是“ $>$ ”或“ $<$ ”,也就是能比较大小顺序关系。

例如,对合格产品按其性能和好坏,可分成一等品、二等品、三等品等。这种尺度虽然不能表明一个单位一等品等于几个单位二等品,但却明确表示一等品性能高于二等品,而二等品性能又高于三等品等。顺序尺度除了用于分类(组)外,在统计分析中还可以确定中位数、四分位数、众数等指标的位置。

(3) 间隔尺度。间隔尺度是对事物类别或次序之间间距的测度,数据表现为数字,它通常使用自然或度量衡单位作为计量尺度。间隔尺度是比顺序尺度高一层次的计量尺度。它不仅能将事物区分为不同类型并进行排序,而且可以准确地指出类别之间的差距是多少。间隔尺度除了比较大小外,还可以做加或减的运算。

例如,学生某门课程的考分,可以从高到低分类排序,形成90分、80分、70分,直到0分的序列。它们不仅有明确的高低之分,而且可以计算差距,90分比80分高10分,比70分高20分等。间隔尺度的计量结果表现为数值,可以进行加或减的运算,但不能进行乘或除的运算,其原因是在等级序列中没有固定的、有确定意义的“零”位。例如,学生甲得分为90分,学生乙得0分,可以说甲比乙多得90分,却不能说甲的成绩是乙的90倍或无穷大。因为“0分”在这里不是一个绝对的标准,并不意味着乙学生毫无知识。恰如我们不能说 $40^{\circ}\text{C}$ 比 $20^{\circ}\text{C}$ 暖和2倍一样。没有确定的标准的“零”位,但有基本的确定的测量单位,如学生成绩的测量单位是1分,质量价差的测量单位是1元,温度的测量单位是 $1^{\circ}\text{C}$ 等,这是间隔尺度的显著特点。

(4) 比率尺度。比率尺度是在间隔尺度的基础上,确定可以作为比较的基数,将两种相关的数加以对比,而形成新的相对数,用以反映现象的构成、相对密度、速度、密度等数量关系,数据表现为数字。由于它是在比较基数上形成的尺度,所以能够显示更加深刻的意义。除了可做比较大小、加减运算外,比率尺度还可以做乘或除的运算。

例如,将某地区人口数和土地面积对比计算人口密度指标,说明人口相对的密集程度。甲地区人口数可能比乙地区的多,但甲地区的土地更广阔,用人口密度指标就可以说明相对来说甲地区人口数不是多了,而是少了。在统计的对比分析中,广泛地运用比率尺度进行计量。

上述四种计量尺度对事物的计量层次是由低级到高级、由粗略到精确逐步递进的。高层次的计量尺度具有低层次计量尺度的全部特性,但不能反过来。显然,可以很容易地将高层次计量尺度的测量结果转化为低层次计量尺度的测量结果,比如将考试成绩的百分制转化为五级计分制。在统计分析中,一般要求测量的层次越高越好,因为高层次的计量尺度包含更多的数学特性,所运用的统计分析方法越多,分析时也就越方便,因此应尽可能使用高层次的计量尺度。上述四种计量尺度的比较见表1.2。

表 1.2 不同计量尺度的比较

类别	含义	特点
分类尺度	分类尺度是按照事物的某种属性对其进行平行的分类,数据表现为类别	(1) 分类尺度是最粗略、计量层次最低的计量尺度; (2) 计量结果不可以区分大小或进行任何数学运算,主要数学特征是“=”或“≠”
顺序尺度	顺序尺度是对事物类别顺序的尺度,数据表现为有序类别	(1) 可以将研究对象分为不同的类别,而且可以反映各类的优劣、量的大小或顺序; (2) 只是测度了类别之间的顺序,而未测量出类别之间的精确差值; (3) 计量结果只能比较大小,不能进行加、减、乘、除等数学运算
间隔尺度	间隔尺度是对事物类别或次序之间间距的测度,数据表现为数字,它通常使用自然或度量衡单位作为计量尺度	(1) 可以用数字表示现象各类别之间的不同和顺序大小的差异,可以用确切的数值反映现象之间在量方面的差异; (2) 反映现象规模水平的数据必须以间隔尺度计量; (3) 计量结果可以进行加、减运算
比率尺度	比率尺度是将两种相关的数加以对比,而形成新的相对数,用以反映现象的构成、相对密度、速度、密度等数量关系,数据表现为数字	比率尺度的计量结果可以进行加、减、乘、除等数学运算

## 2. 统计数据类型

统计数据是对现象测量或计算的结果,也就是在统计工作活动过程中所取得的反映国民经济和社会现象的数字资料以及与之相联系的其他资料的总称。统计数据有三种分类方法,按计量尺度分类,可以分为分类数据、顺序数据和数值型数据;按收集方法分类,可以分为观测数据和实验数据;按时间状况分类,可以分为截面数据和时间序列数据。

按计量尺度分类的统计数据是最常用的。在统计数据的分类中,采用分类尺度计量的数据是分类数据,采用顺序尺度计量的数据是顺序数据,采用间隔尺度和比率尺度计量的数据是数值型数据。

分类数据是只能归于某一类别的非数值型数据,它是对事物分类的结果,数据表现为类别。如性别可分为男性和女性两类,民族可分为汉族、少数民族等。虽然也可以用1或者0表示男性与女性、汉族与少数民族,但是这些数字没有大小比较之分,只是不同类别现象的一个代码,并不代表真正的值,不能进行数学计算。用分类数据对现象进行分析时,由于不同类别间地位平等,没有高低和大小之分,因此各类之间的顺序是可以改变的。分类数据是最粗略、计量层次最低的数据。

顺序数据是只能归于某一有序类别的非数值型数据。如高校教师的职称有助教、讲师、副教授和教授,业主对住房的满意度有很满意、满意、一般、不满意、很不满意等,属于顺序数据。可以用数字1、2、3、4来表示职称,用5、4、3、2、1来表示满意程度,但这些数字代码只能体现一种顺序或者程度,不能体现事物之间的具体数量差别。由于客观现象的不同类别间存在顺序性差异,因此用顺序数据在对现象进行分析时其顺序是不能随意排列的。

数值型数据是按间隔尺度和比率尺度计量的,其结果表现为具体的数值。现实中处理的大多数数据都是数值型数据。

分类数据和顺序数据说明事物的品质特征,通常是用文字来表达的,因此也称为定性数

据或品质数据;数值型数据说明现象的数量特征,通常用数值来表现的,因此也称为定量数据或数量数据。在后续的计算中,不同的统计数据运用的统计计算方法也是不同的。比如,在第4章中,对集中趋势和离散趋势度量时,对分类数据可以计算众数和异众比率;对顺序数据可以计算中位数和四分位差;对数值型数据可以计算平均数和标准差等。

不同类型的数据选用的统计方法也不同。例如,研究分类数据和顺序数据时,可以进行频数分析、列联分析等;研究分类数据与数值型数据的关系时,可以用方差分析;研究数值型数据之间的关系时,可以进行相关分析与回归分析等。如果是时间序列数据,可以进行时间序列分析等。

### 3. 统计数据和变量的区别

统计数据是对现象测量或计算的结果,也就是在统计工作活动过程中所取得的反映社会现象和自然现象的数字资料以及与之相联系的其他资料的总称。

变量是说明现象的某种特征的概念,其特点是从一次观察到下一次观察结果会呈现出差别或变化。例如,人的年龄是变量,人的身高也是变量。变量的具体取值称为变量值。统计数据就是某些统计变量的取值。例如,性别是个分类变量,其变量值“男”或“女”是统计数据,且是分类数据;又如,年龄是个数值型变量,其变量值20岁、30岁、40岁等是统计数据,且是数值型数据。

除了那些变量值是统计数据外,经过变量运算得到的数值也是统计数据。因此,统计数据的含义比变量值的含义更加广泛。

## 1.4 思考题全解(见原书 P10)

### 思考题 1.1 什么是统计学?

**答:** 统计学是收集、处理、分析、解释数据并从数据中得出结论的科学。统计研究的是来自各个领域的数据。数据收集也就是取得统计数据;数据处理是将数据用图表等形式展示出来;数据分析则是选择适当的统计方法研究数据,并从数据中提取有用的信息进而得出结论。

### 思考题 1.2 解释描述统计和推断统计。

**答:** 描述统计研究的是数据的收集、数据的加工处理、数据的汇总、数据的显示、数据分布特征的概括与分析等统计方法。推断统计是研究如何根据样本数据去推断总体数量特征的方法,它是在对样本数据进行描述的基础上,对统计总体的未知数量特征做出推断,主要包括假设检验和参数估计等统计方法。

描述统计和推断统计是现代统计学的两个组成部分,相辅相成、缺一不可。描述统计是现代统计学的基础和前提;推断统计是现代统计学的核心和关键。

### 思考题 1.3 统计数据可分为哪几种类型? 不同类型的数据各有什么特点?

**答:** (1) 统计数据按计量尺度分类,可以分为分类数据、顺序数据和数值型数据;按收集方法分类,可以分为观测数据和实验数据;按时间状况分类,可以分为截面数据和时间序列数据。其中,按计量尺度分类的统计数据是最常用的。

(2) 按计量尺度分类的统计数据特点如下。

分类数据是只能归于某一类别的非数值型数据,它是对事物分类的结果,数据表现为类

别,用文字表述。分类数据可以用于计算各组的频数或频率,计算其众数和异众比率,进行列联表分析和 $\chi^2$ 检验等。

顺序数据是只能归于某一有序类别的非数值型数据。顺序数据虽然也有类别,但这些类别是有序的,可以计算其中位数和四分位差,计算等级相关系数等。

数值型数据是按间隔尺度和比率尺度计量的观测值,其结果表现为具体的数值。数值型数据可以用更多的统计方法进行分析,如计算各种统计量、进行参数估计和检验等。

**思考题 1.4** 解释分类数据、顺序数据和数值型数据的含义。

**答:** 分类数据只能归于某一类别的非数值型数据,它是对事物进行分类的结果,数据表现为类别,是用文字来表述的。顺序数据只能归于某一有序类别的非数值型数据。顺序数据虽然也有类别,但这些类别都是有序的。数值型数据是按间隔尺度和比率尺度计量的观测值,其结果表现为具体的数值。现实中处理的大多数数据都是数值型数据。

**思考题 1.5** 举例说明总体、样本、参数、统计量、变量这几个概念。

**答:** 总体是包含所研究的全部个体(数据)的集合,它通常由所研究的一些个体组成。例如某地区 5 000 个企业构成的集合是总体;某居民小区 500 户居民住户构成的集合是总体;某学校 2 万学生构成的集合是总体。

样本是从总体中抽取的一部分元素的集合,构成样本的元素的数目称为样本量。如从 5 000 个企业中随机抽出 100 个企业,这 100 个企业就构成了一个样本,100 是这个样本的样本量。

参数是用来描述总体特征的概括性数字度量,它是研究者想要了解的总体的某种特征值。总体的参数主要有总体平均数、总体标准差、总体比例等,分别用希腊字母 $\mu$ 、 $\sigma$ 、 $\pi$ 等表示。例如,2014 年某地区 5 000 个企业产值的平均数为 500 万元,总体标准差为 45 万元。

统计量是用来描述样本特征的概括性数字度量。它是根据样本数据计算出来的一个量,由于抽样是随机的,因此统计量是样本的函数。样本统计量主要有样本平均数、样本标准差、样本比例等,分别用英文字母 $\bar{x}$ 、 $s$ 、 $p$ 等表示。抽样的目的就是用样本的统计量去估计总体的参数,如用样本平均数 $\bar{x}$ 去估计总体平均数 $\mu$ ,用样本标准差 $s$ 去估计总体标准差 $\sigma$ ,用样本比例 $p$ 去估计总体比例 $\pi$ 。例如,100 个企业的样本平均产值为 485 万元,可用 485 万元去推断 5 000 个企业总体的平均产值。

变量是说明现象某种特征的概念,其特点是从一次观察到下一次观察结果会呈现出差别或变化,比如“商品销售价格”“人的年龄”“产品的数量”等都是变量。

**思考题 1.6** 变量可分为哪几类?

**答:** 变量可以分为分类变量、顺序变量和数值型变量这三类。

分类变量是说明事物类别的一个名称,其取值是分类数据。如“民族”是一个分类变量,其变量值“汉族”或“彝族”等属于分类数据。

顺序变量是说明事物有序类别的一个名称,其取值是顺序数据。如“高校职称”是一个顺序变量,其变量值“教授”“副教授”或“讲师”等属于顺序数据。

数值型变量是说明事物数字特征的一个名称,其取值是数值型数据。如“企业产值”是一个数值型变量,其变量值“500 万元”“600 万元”等属于数值型数据。

**思考题 1.7** 举例说明离散型变量和连续型变量。

**答:** 数值型变量可以分为离散型变量和连续型变量两种。离散型变量是指其数值只能



用自然数或整数单位计算的变量。例如,企业个数、职工人数、设备台数等,只能取整数。离散型变量的数值一般用计数方法取得。连续型变量是可以在一个或多个区间中取任何值的变量,它的取值是连续不断的,不能一一列举。例如,生产零件的规格尺寸、人体测量的身高、体重、胸围等为连续型变量,它们的取值可以取到小数。连续型变量的数值只能用测量或计算的方法取得。

**思考题 1.8** 请举出统计应用的几个例子。

**答:** 统计的应用无处不在。比如,在企业中,对职工按学历或者按年龄段进行频数分析,对企业历年来的产值进行时间序列分析,对企业的产品质量进行抽样检验并进行参数估计等;在高校中,对教职工按职称进行频数分析,对历年来学生的就业情况进行动态分析,对每年的新生来源进行结构分析等;在医学方面,在新药投入生产前,对药品进行假设检验等。

**思考题 1.9** 请举出应用统计的几个领域。

**答:** 统计的应用领域非常广泛,已经应用到自然科学和社会科学的众多领域中。可以说,几乎所有的研究领域都要用到统计方法,例如,经济与管理、社会学、生态学、教育学、市场营销学、军事科学、心理学、气象学、医学诊断、分类学、质量控制等。

## 1.5 练习题全解(见原书 P10)

**练习题 1.1** 指出下面变量的类型。

- (1) 年龄。
- (2) 性别。
- (3) 汽车产量。
- (4) 员工对企业某项改革措施的态度(赞成、中立、反对)。
- (5) 购买商品时的支付方式(现金、信用卡、支票)。

**解:** (1) 数值型变量。

(2) 分类变量。

(3) 数值型变量。

(4) 顺序变量。

(5) 分类变量。

**练习题 1.2** 某研究部门准备抽取 2 000 个职工家庭推断该城市所有职工家庭的年人均收入。

**要求:**

- (1) 描述总体和样本。
- (2) 指出参数和统计量。

**解:** (1) 描述的总体是“该城市所有的职工家庭”,样本是“抽取的 2 000 个职工家庭”。

(2) 参数是“该城市所有职工家庭的年人均收入”,统计量是“抽取的 2 000 个职工家庭计算出的年人均收入”。

**练习题 1.3** 一家研究机构从 IT 从业者中随机抽取 1 000 人作为样本进行调查,其中 60%的人回答他们的月收入在 5 000 元以上,50%的人回答他们的消费支付方式是用信用卡。

回答以下问题：

- (1) 这一研究的总体是什么？
- (2) 月收入是分类变量、顺序变量还是数值型变量？
- (3) 消费支付方式是分类变量、顺序变量还是数值型变量？
- (4) 这一研究涉及截面数据还是时间序列数据？

解：(1) 这一研究的总体是“所有 IT 从业者”。

- (2) 月收入是数值型变量。
- (3) 消费支付方式是分类变量。
- (4) 这一研究涉及的数据是截面数据。

**练习题 1.4** 一项调查表明，消费者每月在网上购物的平均花费是 200 元，他们选择在  
网上购物的主要原因是“价格便宜”。

回答以下问题：

- (1) 这一研究的总体是什么？
- (2) “消费者在网上购物的原因”是分类变量、顺序变量还是数值型变量？
- (3) 研究者所关心的参数是什么？
- (4) “消费者每月在网上购物的平均花费是 200 元”是参数还是统计量？
- (5) 研究者所使用的主要是描述统计方法还是推断统计方法？

解：(1) 这一研究的总体是“所有的网上购物者”。

- (2) “消费者在网上购物的原因”是分类变量。
- (3) 研究者所关心的参数是“所有的网上购物者的月平均花费”。
- (4) “消费者每月在网上购物的平均花费是 200 元”是统计量。
- (5) 研究者所使用的统计方法是推断统计方法。



## 第2章

# 数据的搜集

### 2.1 学习目标

- 了解统计数据的来源；
- 了解统计数据搜集的具体方式和方法；
- 了解数据误差的概念和种类。

### 2.2 基本知识点

表 2.1 基本知识点梳理

章节	主要内容	学习要点
2.1 数据的来源	数据的间接来源	◆ 二手资料获取的主要渠道 ◆ 二手资料的特点与作用 ◆ 二手资料的评估
	数据的直接来源	◆ 一手资料获取的方法 ◆ 调查数据和实验数据的概念
2.2 调查数据	搜集数据的具体方式	◆ 概率抽样的概念和特点 ◆ 概率抽样的方式 ◆ 非概率抽样的概念和特点 ◆ 非概率抽样的方式
	搜集数据的基本方法	◆ 自填式、面访式、电话式的概念与特点 ◆ 数据搜集方法的选择
2.3 实验数据	实验组和对照组	◆ 实验组和对照组的概念 ◆ 实验组和对照组的选择
	实验中的若干问题	◆ 注意人的意愿、心理问题、道德问题对实验的影响
2.4 数据的误差	抽样误差	◆ 抽样误差的概念及影响因素
	非抽样误差	◆ 非抽样误差的概念及其类型 ◆ 抽样框误差、回答误差、无回答误差、调查员误差、测量误差的概念
	误差的控制	◆ 误差控制的方法

## 2.3 重难点解析

### 1. 统计数据的来源

统计数据的来源主要有两条渠道,一是数据的间接来源,二是数据的直接来源。

数据的间接来源亦称二手资料,是指由别人通过调查或者实验的方式搜集的,使用者只是找到它们并对这些信息重新加工、整理,使之成为统计分析可以使用的数据。从搜集的范围看,间接数据可以取自系统外部,也可以取自系统内部。系统外部数据主要来源于各种出版物、各类机构和各级政府统计网站所公布的统计公报、统计分析报告和统计数据资料。系统内部数据主要来源于业务资料,即与业务经营活动有关的各种单据、记录,经营活动中产生的各种统计报表,以及各种财务、会计核算和分析的资料等。二手数据搜集的成本低,搜集方便、快捷,比较容易,但有时针对性不够。

数据的直接来源亦称一手资料,是指通过自己的调查或者实验活动,直接获取的第一手数据,包括调查数据和实验数据。通过调查方法获得的数据称为调查数据,通过实验方法得到的数据称为实验数据。一手数据通过调查或者实验取得,因此,搜集数据的成本比较高,耗费的人力、物力也比较大,但数据针对性比较强。

是采用一手数据还是采用二手数据,要根据具体的时间、人力、物力、调查费用以及搜集资料的难易程度等情况而定。

### 2. 概率抽样与非概率抽样

表 2.2 概率抽样与非概率抽样比较表

抽样类型	特 点	常见抽样方式	含 义
概率抽样	1. 按随机原则抽取样本 2. 能对总体进行参数估计 3. 技术含量高、抽样成本高 4. 适合掌握总体的数量特征	简单随机抽样	从总体中随机抽取个体单位组成样本进行调查
		分层抽样	将抽样单位按照某种特征或某种规划进行分类(得到不同的层),再从不同的层中随机抽取样本进行调查
		整群抽样	从总体中随机抽取群,然后对中选群中的所有单位进行全面调查
		系统抽样	将总体中各单位按一定顺序排列,根据抽选间隔,随机确定一个单位作为起始单位,然后按照一定的规则,每隔一定的间隔抽取一个单位的一种抽样方式
		多阶段抽样	随机抽选群,然后再从中选群中随机抽取抽样单位,若此时抽样结束,则是两阶段抽样。若抽样的段数增多,就是多阶段抽样
非概率抽样	1. 不按随机原则抽取样本 2. 不能对总体进行参数估计 3. 操作简便、时效快、成本低 4. 适合做探索性研究	方便抽样	依据方便的原则,自行确定入抽样本的单位。例如,街头拦截式调查
		判断抽样	根据经验、判断和对研究对象的了解,有目的地选择一些单位作为样本。有重点抽样、典型抽样和代表抽样等
		自愿样本	被调查者自愿参加,成为样本之一,向调查人员提供信息