

21世纪高等学校经济管理系列教材

统计学

●●● TONGJIXUE

曾霞 王磊 朱冬辉 主编



课程思政 挖掘思政资源 践行“三全育人”
教学视频 “看”书“听”书 直观学习
在线测试 “码”上做“码”上答



中国财经出版传媒集团
中国财政经济出版社

图书在版编目 (CIP) 数据

统计学 / 曾霞, 王磊, 朱冬辉主编. — 北京: 中国财政经济出版社, 2022. 8 (2024. 1 重印)

21 世纪高等学校经济管理系列教材

ISBN 978 - 7 - 5223 - 1439 - 6

I. ①统… II. ①曾… ②王… ③朱… III. ①统计学 - 高等学校 - 教材 IV. ①C8

中国版本图书馆 CIP 数据核字 (2022) 第 080925 号

责任编辑: 张怡然

责任印制: 张 健

封面设计: 易 帅

责任校对: 徐艳丽

本书主审: 朱建平

统计学

TONGJIXUE

中国财政经济出版社 出版

URL: <http://www.cfeph.cn>

E-mail: cfeph@cfeph.cn

(版权所有 翻印必究)

社址: 北京市海淀区阜成路甲 28 号 邮政编码: 100142

营销中心电话: 010 - 88191522

天猫网店: 中国财政经济出版社旗舰店

网址: <https://zgczjjcbs.tmall.com>

三河市宏图印务有限公司印刷 各地新华书店经销

成品尺寸: 185mm × 260mm 16 开 21 印张 423 000 字

2022 年 8 月第 1 版 2024 年 1 月河北第 2 次印刷

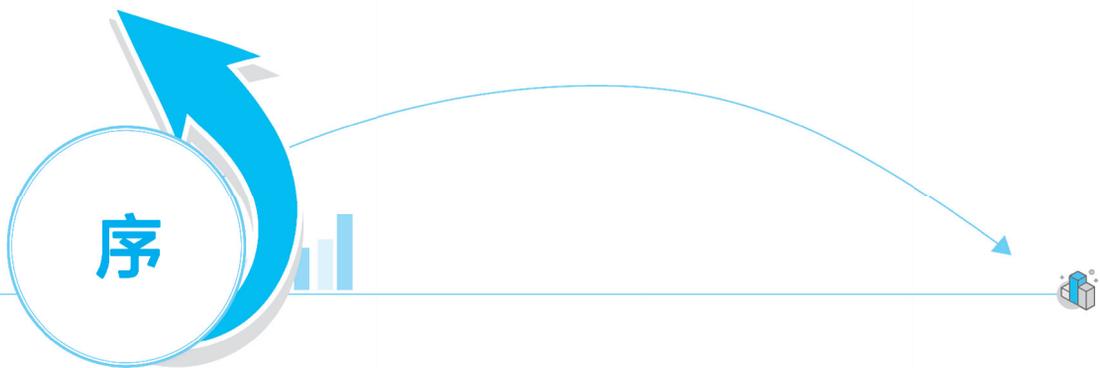
定价: 58.00 元

ISBN 978 - 7 - 5223 - 1439 - 6

(图书出现印装问题, 本社负责调换, 电话: 010 - 88190548)

本社质量投诉电话: 010 - 88190744

打击盗版举报热线: 010 - 88191661 QQ: 2242791300



中国财政经济出版社的《统计学》教材已经出版，该书的主编之一王磊老师邀请我写一篇序。湖北经济学院统计人长期以来，致力于统计学教材的编写与修订，《统计学》一书就是他们在这方面工作的重要成果。

教育是国之大计、党之大计，办好教育，离不开高质量教材建设。近年来，我国高校经济管理类专业统计学教材建设取得了一系列积极的成果。这些成果不仅丰富了我国统计学教育的资源，也为培养优秀的统计学人才提供了坚实的基础。但随着我国社会经济的快速发展和数字经济的全面推进，现有教材的实例往往难以体现我国社会经济发展取得的成绩，难以体现新时代的奋斗目标。

为了进一步深化教育领域的综合改革，深入贯彻党的二十大精神，在习近平新时代中国特色社会主义思想的指导下，曾霞、王磊与朱冬辉老师主编的《统计学》尝试将新时代明确的奋斗目标融入教材编写过程。该教材的编写秉持教育为基础、文化为魂的原则，引导学生进行创新性思考，努力推进党的二十大精神在教材、课堂和学生中的深入传播与实践。统计学是一门关于数据的科学，数据中蕴含着大量的辩证唯物主义观点，为统计学与思政教育结合提供了可能，从而实现全员、全程、全方位育人的要求。

本书将社会经济统计学和数理统计学有机融合起来，并注重思想政治教育的融入，主要特点如下：

一是教材理论体系框架结构合理，内容丰富且完整。按照统计学的定义，教材涵盖了收集、处理、分析和解释数据等内容，并详细介绍了描述统计分析方法和推断统计分析方法。

二是教材采用了与时俱进和生动的案例与例题，这些来源于现实的例题既能提高学生的学习兴趣，又能帮助学生培养将实际问题抽象为统计学问题的能力，同时也实现了教学内容和思政教育的有机结合，强化培养学生知、行、思、辩的全方位学习能力。

三是教材借助于 Excel、SPSS 等软件，将统计数据的分类、汇总和结果用图形展示，并应用于假设检验、方差分析等推断统计方法，提升学生将统计学理论方法应用



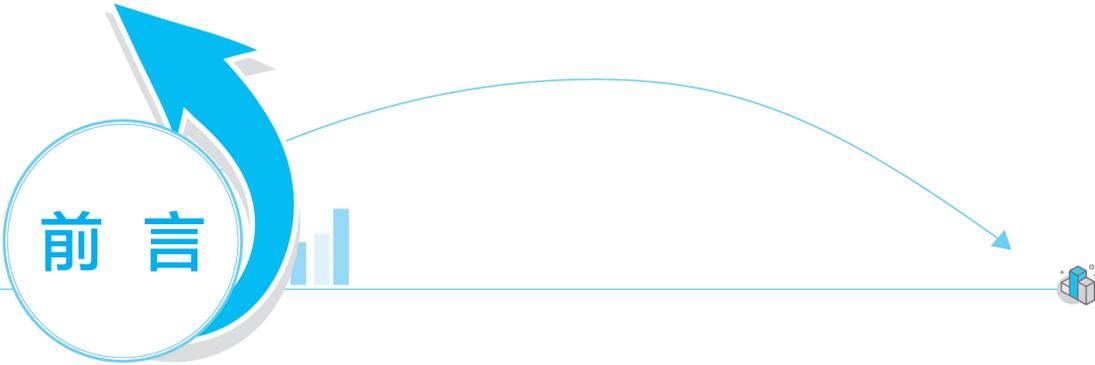
于实际问题的能力。

四是教材构建立体化的教学资源，考察重要知识点的练习和软件操作均配有二维码，扫码即可观看相关视频资源，除此之外，还配备了包括教学课件等丰富的增值辅助教学资源，可免费下载使用。

五是教材积极响应思想政治教育贯穿人才培养体系指导精神，在每章结束时讲述了中国统计学家的故事，通过统计学家的生平和学术贡献，实现知识传授和价值引领的统一，切实提高立德树人成效。

笔者相信，本书的出版有助于提高经济管理类专业学生的专业水平和技能，有助于培养他们的思想道德素质、社会责任感和创新精神，为国家和社会的发展做出更大的贡献。在新时代背景下，我们要始终坚持党的领导，以习近平新时代中国特色社会主义思想为指导，推动统计学教育不断创新和发展。

厦门大学教授



前言

教育是国之大计、党之大计。党的十八大以来，以习近平同志为核心的党中央高度重视教育问题，始终把“为谁培养人、培养什么人、怎样培养人”作为根本问题，强调把立德树人作为根本任务。党的二十大报告明确指出：“教育、科技、人才是全面建设社会主义现代化国家的基础性、战略性支撑。”这为我国教育、科技和人才事业未来发展提供了重要遵循。我们教育工作者学习贯彻党的二十大精神，就是要结合自身工作实际，不忘立德树人初心，牢记为党育人、为国育才使命，以课程建设为抓手，培养堪当民族复兴大任的时代新人，统计学课程建设亦是如此。

统计学是一门收集、处理、分析、解释数据并从数据中得出结论的科学，揭示的是现象总体内在的数量规律，主要包括收集数据、处理数据、分析数据、解释数据和得到结论五个部分。随着社会经济的发展和计算机技术的进步，统计学的应用越来越广泛。无论是政府部门为了进行有效的管理与指导，还是公司或企业的生产管理以及居民的日常生活等，都离不开统计学。近年来，我国高校经济管理类专业统计学教材在内容和体系上都有了较大改进。然而，现今社会知识更新的速度正逐步加快，与此同时，习近平总书记在2019年3月提出了“挖掘其他课程和教学方式中蕴含的思想政治教育资源，实现全员全程全方位育人”的要求。党的二十大报告进一步指出，青年强，则国家强。而统计学是一门关于数据的科学，数据中蕴含着大量的辩证唯物主义观点，为统计学与思政教育结合提供了可能，为引导青年学生树立为国奋斗的理想提供了载体，也为我们出版这本融合了思政元素内容的教材提供了历史机遇。我们统计学教师要用自己的学识、阅历、经验去点燃学生对真善美的向往，通过统计分析方法讲好中国故事，使社会主义核心价值观润物无声地滋养青年学生的精神世界，引导学生把人生理想融入中华民族伟大复兴的历史进程。

依据统计学的定义，本书的主要内容分为四大部分。第一部分包含第1章，主要学习和了解统计的产生和发展、统计学的定义、统计学的应用以及一些基本定义。第二部分主要介绍了描述性统计分析方法，包含第2章、第3章和第4章，具体内容为统计数据的搜集、数据的整理与展示和数据分布特征的度量。第三部分主要介绍了推断统计分析方法，包含第5章至第11章，具体内容为抽样分布理论、参数估计、假设



检验、分类数据分析、方差分析、相关分析与回归分析、时间序列分析。第四部分包含第12章，主要学习和了解指数的相关知识。

要学好统计学并能灵活运用，一方面要掌握基本的统计理论与方法；另一方面要学会将实际问题抽象为统计问题，并能掌握基本的使用统计软件的方法。我们以“围绕经管类专业建设培养统计技能”为目标，以“建设统计学省级一流课程”为依托，组织相关教师编写了本教材。本教材具有以下特点：一是理论体系框架结构合理，内容丰富且完整。按照统计学的定义，教材涵盖了收集、处理、分析和解释数据等内容，并详细介绍了描述统计分析方法和推断统计分析方法；二是采用了与时俱进和生动的案例与例题，这些来源于现实的例题既能提高学生的学习兴趣，又能帮助学生培养将实际问题抽象为统计学问题的能力，提升经管类专业学生的数据素养能力，同时也实现了教学内容和思政教育的有机结合，实现思政教育与专业课程教育同向、同行、协同育人；三是借助于 Excel、SPSS 等软件，将统计数据的分类、汇总和结果用图形展示，并应用于假设检验、方差分析等推断统计方法，提升学生将统计学理论方法应用于实际问题的能力；四是每章结束时讲述了统计学家的故事，通过统计学家的生平和学术贡献，实现知识传授和价值引领的统一，切实提高立德树人成效；五是为贯彻落实国家文化数字化战略，本书还配有大量微课视频，扫码即可观看。

本书可作为高等学校经济管理类以及人文、社会科学等学科的本、专科教材或参考书，也可以作为从事统计工作或统计分析工作人员的参考书。

本书由曾霞、王磊和朱冬辉担任主编。各章执笔人分别为：第1章、第2章朱冬辉，第3章耿智琳，第4章肖宪标，第5章、第6章、第7章曾霞，第8章、第9章王磊，第10章李文君，第11章赵琼，第12章肖宪标。

编者

2021年12月

目 录

第1章 绪论	1
育人目标	1
案例 你知道统计对我们很重要吗?	1
1.1 统计与统计学	2
1.1.1 什么是统计	2
1.1.2 统计学及其分类	4
1.1.3 统计的应用	5
1.2 统计数据	5
1.2.1 分类数据和数值型数据	6
1.2.2 观测数据和实验数据	7
1.2.3 截面数据、时间序列数据和面板数据	7
1.3 几个基本概念	8
1.3.1 总体与样本	8
1.3.2 统计指标与统计指标体系	9
1.3.3 变量与变量值	10
本章小结	11
思政园地	12
思考与练习	13
第2章 统计数据的搜集	15
育人目标	15
案例 数据究竟美在哪?	15
2.1 数据的来源	17
2.1.1 数据的间接来源	17
2.1.2 数据的直接来源	18
2.2 统计调查	19
2.2.1 普查	19



2.2.2	概率抽样与非概率抽样	20
2.2.3	统计调查方案的设计	26
2.3	数据搜集的方法与道德准则	27
2.3.1	数据搜集方法	27
2.3.2	数据搜集中的道德准则	29
2.4	实验法	30
2.4.1	实验组和对照组	31
2.4.2	实验法中的统计	32
2.4.3	实验法产生的问题	32
2.5	数据的误差	34
2.5.1	抽样误差	34
2.5.2	非抽样误差	35
2.5.3	误差的控制	37
	本章小结	38
	思政园地	39
	思考与练习	40
第3章	数据的整理与展示	45
	育人目标	45
	案例 如何展示普通高等教育学校（机构）情况？	45
3.1	数据的预处理	46
3.1.1	数据的审核	46
3.1.2	数据的筛选与排序	46
3.2	品质数据的图表	47
3.2.1	分类数据的图表	47
3.2.2	有序分类数据的图表	50
3.3	数值型数据的图表	52
3.3.1	数据的分组	52
3.3.2	数值型数据的图表	53
3.4	Excel 实践与提高	58
	本章小结	68
	思政园地	69
	思考与练习	70



第 4 章 数据分布特征的度量	72
育人目标	72
案例 2021 年上半年我国居民收入和消费支出情况分析	72
4.1 集中趋势的度量	74
4.1.1 分类数据：众数	75
4.1.2 有序分类数据：中位数和分位数	76
4.1.3 数值型数据：平均数	77
4.2 离散程度的度量	82
4.2.1 分类数据：异众比率	83
4.2.2 有序分类数据：四分位差	83
4.2.3 数值型数据：极差、平均差、方差和标准差	84
4.2.4 相对位置的度量：标准分数、经验法则、切比雪夫不等式	87
4.3 偏态与峰态的度量	88
4.3.1 偏态	88
4.3.2 峰态	90
4.3.3 实践与提高	90
本章小结	92
思政园地	93
思考与练习	94
第 5 章 抽样分布理论	97
育人目标	97
案例 如何评价某种疫苗的有效性?	97
5.1 正态分布与二项分布	97
5.1.1 正态分布	98
5.1.2 二项分布	99
5.2 参数与统计量	100
5.2.1 参数	100
5.2.2 统计量	100
5.3 抽样分布理论	102
5.3.1 抽样分布的定义	102
5.3.2 χ^2 分布	102
5.3.3 t 分布	103



5.3.4 F 分布	105
5.4 样本均值的抽样分布与中心极限定理	107
5.4.1 样本均值的抽样分布	107
5.4.2 中心极限定理	107
5.5 样本比例的抽样分布	110
5.6 样本方差的抽样分布	110
本章小结	110
思政园地	111
思考与练习	112
第6章 参数估计	114
育人目标	114
案例 某款车型的汽车安全吗?	114
6.1 参数估计的一般问题	115
6.1.1 估计量与估计值	115
6.1.2 点估计与区间估计	115
6.1.3 评价估计量的准则	119
6.2 一个总体参数的区间估计	121
6.2.1 总体均值的区间估计	121
6.2.2 总体比例的区间估计	124
6.2.3 总体方差的区间估计	125
6.3 两个总体参数的区间估计	126
6.3.1 两个总体均值之差的区间估计	127
6.3.2 两个总体比例之差的区间估计	131
6.3.3 两个总体方差之比的区间估计	132
6.4 样本容量的确定	133
6.4.1 估计一个总体均值时的样本容量	133
6.4.2 估计一个总体比例时的样本容量	134
本章小结	135
思政园地	137
思考与练习	139
第7章 假设检验	142
育人目标	142



案例 有疵点电路板的比率超过了设计规定比率的证据是什么?	142
7.1 假设检验的一般问题	143
7.1.1 假设检验问题的提出	143
7.1.2 假设的陈述	143
7.1.3 两类错误与显著性水平	145
7.1.4 检验统计量与拒绝域	146
7.1.5 根据 P 值做决策	147
7.2 一个总体参数的假设检验	149
7.2.1 总体均值的检验	149
7.2.2 总体比例的检验	153
7.2.3 总体方差的检验	154
7.3 两个总体参数的假设检验	156
7.3.1 两个总体均值之差的检验	156
7.3.2 两个总体比例之差的检验	161
7.3.3 两个总体方差之比的检验	163
本章小结	164
思政园地	165
思考与练习	167
第 8 章 分类数据分析	169
育人目标	169
案例 你在线上购买生鲜食品吗?	169
8.1 分类数据与卡方检验	170
8.1.1 分类数据	170
8.1.2 卡方检验	170
8.2 拟合优度检验	170
8.2.1 定义	170
8.2.2 Excel 实践与提高	172
8.3 列联分析: 独立性检验	173
8.3.1 列联表	173
8.3.2 列联表的分布	174
8.3.3 独立性检验	174
8.3.4 卡方分布的期望值准则	175
8.3.5 SPSS 实践与提高	176



8.4 列联表中的相关测量	179
8.4.1 φ 相关系数	179
8.4.2 列联相关系数	180
8.4.3 V 相关系数	180
8.4.4 SPSS 实践与提高	181
本章小结	182
思政园地	183
思考与练习	184
第9章 方差分析	186
育人目标	186
案例 三种不同类型的疫苗对新冠病毒的免疫度是否不同?	186
9.1 方差分析引论	186
9.1.1 问题的提出	187
9.1.2 方差分析及相关概念	187
9.1.3 方差分析的基本思路	188
9.1.4 方差分析的基本假定	189
9.2 单因素方差分析	189
9.2.1 数据结构形式	189
9.2.2 分析步骤	190
9.2.3 案例与软件实现	192
9.2.4 关系强度的测量与多重比较	194
9.3 双因素方差分析	197
9.3.1 双因素方差分析及其类型	197
9.3.2 无交互作用的双因素方差分析	198
9.3.3 有交互作用的双因素方差分析	203
本章小结	207
思政园地	207
思考与练习	208
第10章 相关分析与回归分析	212
育人目标	212
案例 回归与电影	212
10.1 相关分析	214
10.1.1 变量间的关系	214



10.1.2	相关关系的种类	215
10.1.3	相关分析的内容	216
10.1.4	相关关系的描述——相关表和相关图	216
10.1.5	相关程度的测度——相关系数	219
10.2	一元线性回归分析	224
10.2.1	一元线性回归模型	224
10.2.2	一元线性回归模型参数的估计	226
10.2.3	一元线性回归模型的检验	229
10.2.4	利用回归方程进行预测	235
10.3	多元线性回归分析	237
10.3.1	多元线性回归模型	238
10.3.2	多元线性回归模型参数的估计	238
10.3.3	多元线性回归模型的检验	241
10.3.4	关于多重共线性问题	246
10.3.5	利用多元线性回归方程进行预测	254
	本章小结	255
	思政园地	256
	思考与练习	258
第 11 章	时间序列分析	263
	育人目标	263
	案例 如何预测下一年份各季度的国内生产总值?	263
11.1	时间序列概述	264
11.1.1	时间序列的概念	264
11.1.2	时间序列的构成要素	264
11.2	时间序列的描述性分析和长期趋势分析	267
11.2.1	发展水平和增长量	268
11.2.2	发展速度和增长速度	268
11.2.3	应注意的问题	271
11.2.4	移动平均法和指数平滑法	271
11.2.5	趋势方程拟合法	273
11.3	时间序列季节变动分析和预测	275
11.3.1	同期平均法	276
11.3.2	移动平均趋势剔除法	276



本章小结	280
思政园地	280
思考与练习	281
第 12 章 指数	284
育人目标	284
案例 你知道哪些指数呢?	284
12.1 指数的一般问题	286
12.1.1 指数的定义	287
12.1.2 指数的分类	288
12.2 总指数的编制	288
12.2.1 综合指数的编制	289
12.2.2 平均指数的编制	295
12.2.3 几种常见指数介绍	298
12.3 指数体系与因素分析	300
12.3.1 指数体系	300
12.3.2 因素分析	301
本章小结	306
思政园地	307
思考与练习	308
附录 用 Excel 生成概率分布表	312
附录 1 标准正态分布表	312
附录 2 标准正态分布的分位数表	313
附录 3 t 分布临界值表	315
附录 4 χ^2 分布临界值表	316
附录 5 F 分布临界值表	318
参考文献	322

第 1 章 绪 论



育人目标



1. 通过统计学的基本理论与研究方法的学习，更好地利用统计学的基本方法去了解 and 认识我国经济建设与发展的成就。

2. 党的二十大报告指出：在基本实现现代化的基础上，我们要继续奋斗，到本世纪中叶，把我国建设成为综合国力和国际影响力领先的社会主义现代化强国。综合国力、国际影响力和现代化强国如何体现？可以由哪些指标来反映？通过统计学的学习将能很好地回答这些问题。



案例

你知道统计对我们很重要吗？

统计资料表明：我国人均国内生产总值（GDP）由 2000 年的 7 942 元增长到 2018 年的 64 521 元，按 2018 年汇率折算为 9 771 美元，已达到或超过中等收入国家水平；按 2000 年不变价格计算，人均 GDP 年均增长速度为 8.99%，略低于 GDP 增长速度；人均可支配收入由 2000 年的 3 721.3 元增长到 2018 年的 28 228 元，按 2000 年不变价格计算，年均增长速度为 8.57%，城镇居民可支配收入年均增长 7.44%，农村居民可支配收入年均增长 7.48%。从恩格尔系数的大小看，2000 年到 2018 年恩格尔系数呈逐年下降之势，其中，城镇居民恩格尔系数由 2000 年的 39.2% 降到 2018 年的 27.72%，农村居民恩格尔系数由 2000 年的 49.1% 降到 30.07%，表明我国居民的生活水平不断提高。根据联合国粮农组织恩格尔系数的划分标准，恩格尔系数在 40% ~ 49% 为小康，30% ~ 39% 为相对富裕，20% ~ 30% 以下为富裕，20% 以下为极其富裕。这说明，改革开放给我国居民生活水平带来了极大的提高，居民生活水平向相对富裕的阶段迈进。GDP 指标是怎样计算出来的？增长速度是一个什么样的指标？恩格尔系数为什么能反映一个国家或地区居民生活水平？

国际上通常把基尼系数 40% 作为收入分配差距的“警戒线”，基尼系数在 40%



以上表示收入差距较大。2000年以来,我国基尼系数长时期都处在“警戒线”以上^①,最高时2008年达到了49.1%,2018年只是略有下降,仍为46.8%,收入分配呈现出较大的差距,这表明如何缩小居民之间的收入分配差距仍然是我国政府管理部门和经济学界值得关注的一个重要问题。为什么国际上将基尼系数作为衡量一个国家或地区收入分配差距的一个重要指标?为什么将基尼系数40%设置为收入分配差距的“警戒线”。

2008年美国“次贷危机”发生后,我国经济建设面临复杂而严峻的经济形势。国际局势复杂多变,贸易保护主义抬头,经济下行压力加大,国内经济体制改革进入“深水区”,生产方式面临升级换代期,国内的经济发展水平、发展速度和发展质量处于不同层面,使得不同区域之间或区域内部不同省份的经济建设呈现出较大的发展差异,而且这种差异变化随着经济规模的不断增大变得越来越明显且越来越复杂。这种区域经济发展差异究竟呈现出怎样的特征,在不同区域之间或同一区域不同省份之间又表现为什么样的状态?

上述问题必须依靠统计学的专门知识来加以分析解答,统计对我们很重要。因此,诺贝尔经济学奖得主、著名经济学家萨缪尔森在其经典的教科书——《经济学(第12版)》中特别提到:“在许多与经济学有关的学科中,统计学是特别重要的。”自诺贝尔经济学奖设立以来,三分之二以上的诺贝尔经济学奖获奖者的研究成果与统计有关。

1.1 统计与统计学



19世纪英国著名的科幻作家赫伯·乔治·威尔斯(Herbert George Wells)曾经说过:“统计思维总有一天会像读与写一样成为一个有效率的公民的必备能力。”统计是什么?与什么打交道?统计学又是关于什么的科学?下面将对这些问题展开讨论。

1.1.1 什么是统计

日常生活中我们经常会遇到这样的问题,或在报纸、杂志上看到这样的报道和描述:无论你在百货商店购物还是在网上冲浪,总有人在收集与我们购物或上网有关的资料;又比如你可能是大学毕业参加工作没有多少年的白领工作者,准备购置一套房子作为结婚用房,这个时候需要你做一定的市场调查,也就是说你需要了解目前市场上流行什么样式的房子,居住面积多大?房子的价格如何?配套设施及周边环境怎样?你有足够的支付能力吗?等等。这里就涉及搜集统计数据、统计分析判断等相关问题,

^① 2015年中国基尼系数为0.462 创12年来最低 [EB/OL]. (2016-01-19). 中国经济网.



有了这些数据就能够为我们提供相应的指导意见。

作为一名潜在的汽车购买者和一名具有良好社会责任的公民，你愿意为保护地球生态环境做出应有的贡献。根据最新的研究结果，消费者的行为对自然资源到底有怎样的影响呢？你应该使用汽油发动机的汽车，还是新能源汽车，或者电动车、自行车呢？

报纸、杂志或互联网上的统计结果会影响你的消费观。

世界银行世界发展指数（WDI）数据显示，2019年美国国内生产总值（GDP）为214 277亿美元，增速为2.3%；而中国GDP为143 429亿美元，增速为6.1%，GDP规模稳居世界第二位，为位居第三位日本的GDP总额的2.82倍。但人均GDP中国仅为美国的16%、日本的26%，与世界平均水平11 436美元/人相比，少了1 174美元。这样的结果意味着我们需要用统计的方法来对所研究的事物、现象进行分析比较。

在认识纷繁复杂的当今世界时，统计学自有其一席之地。对于列入国家食品管理部门食品名单上的食品是否含有转基因；新冠疫苗的接种数量与接种普及程度如何，疫苗（Vero细胞）的预防效果是否有差异；学前教育是否影响孩子们上学后的学习表现；维生素C是否有助于预防疾病等；这些都可以运用统计方法进行分析。总之，不论何时何地，只要有数据，统计就能派上用场。

卡尔文勋爵指出：“当你能衡量你所谈论的事并且能用数字把它表示出来时，可以说你对它有一些了解；当你不能衡量它并且不能用数字表示出来时，你对它的了解是肤浅而不能令人满意的；你的了解可能是知识的开端，但很难说你已经进入了科学的阶段。”^①这里所谈论的数字也就是我们日常生活中经常用到的统计数据，同时他也告诉我们，要对客观事物进行科学的了解和认知，需要利用统计数据进行分析。当然，在利用统计工具展示数据资料的信息时，是需要掌握一定的数据分析技巧的。

上面提到的问题和现象都与统计有关，那什么是统计呢？我们认为“统计”一词的基本含义应该是：**统计**是人们认识客观世界总体数量特征和数量变化规律活动的统称。统计与数字打交道，是对现象的数量方面进行分析、判断，通过对现象的统计，达到用数量来揭示现象的状况、变化以及规律、未来趋势等目的。



思政元素

《中华人民共和国统计法》明确规定：统计的基本任务是对经济社会发展情况进行统计调查、统计分析，提供统计资料和统计咨询意见，实行统计监督。国家建立集中统一的统计系统，实行统一领导、分级负责的统计管理体制。国务院和地方各级人民政府、各有关部门应当加强对统计工作的组织领导，为统计工作提供必要的保障。

① 萨缪尔森. 经济学 [M]. 下册. 高鸿业, 译. 北京: 商务印书馆, 1979, 119.



1.1.2 统计学及其分类

统计学是在统计实践的基础上产生而发展起来的方法论科学。英文里的统计学 (statistics) 是由拉丁文的国家、政府 (status) 演化而来。有很长一段时间, 它被视为是有关于国家经济、人口统计及政治等各种现象所显现出的原始数据及图表。即使在今日, 许多的报道及政府公报里也常含有大量令人感兴趣的数据所制成的图表。到目前为止, 统计学界对于统计学的定义也是莫衷一是, 众说纷纭。不列颠百科全书认为统计学是收集、分析、表述和解释数据的科学; 《韦伯斯特国际辞典 (第3版)》定义统计是一门收集、分析、解释和提供数据的科学; 马里奥·特里奥拉 (Mario F. Triola) 在《初级统计学》一书中指出, 统计指的是一组方法, 用来设计实验、获得数据, 然后在这些数据的基础上组织、概括、演示、分析、解释和得出结论。

从上面的定义中, 我们不难发现, 统计学的定义基本类似, 都是与数据有关的科学。因此, 可以对统计学进行如下的定义。**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。它应该包括如下几个方面的内容:

(1) 收集数据, 即如何取得数据。这里涉及采用什么方法得到数据, 从哪些渠道取得等内容。

(2) 处理数据, 即如何整理与展示数据。这里涉及采用什么方法来整理数据, 采用什么形式 (图、表) 来展示数据等内容。

(3) 分析数据, 就是利用统计方法分析数据。这里将告诉我们根据数据特点采用什么方法分析数据。

(4) 数据解释, 它是对结果的说明。这里涉及以怎样的语言 (统计语言或经济语言) 对分析结果进行解释。

(5) 得到结论, 即从数据分析中得出客观结论。这里将告诉我们如何从统计数据和分析解释中得到客观而准确的结论等, 它需要一定的专业知识来支撑且要有敏感的洞察力 (知识、经验、判断)。

统计学按照其采用的数据分析方法可分为描述统计和推断统计。描述统计研究数据搜集、处理、汇总、图表描述、概括与分析等统计方法, 内容包括搜集数据、整理数据、展示数据和描述性分析等, 通过描述统计反映数据的基本特征, 找出数据变化的规律。推断统计是研究如何利用样本数据来推断总体特征的统计方法, 内容主要有参数估计、假设检验和模型构造与估计等, 通过推断统计可以对现象的总体特征作出推断。不论是描述统计还是推断统计都是为了揭示现象总体内在的数量规律, 只不过所依赖的分析方法不同而已。描述统计依赖现象的数据资料, 通过数据搜集、整理、显示和分析来揭示其总体数量规律性; 推断统计是基于分布理论、大数定律和中心极限定理, 利用样本信息和概率论对总体数量特征进行估计和检验等来揭示其总体数量规律。因而可以认为描述统计是推断统计的基础, 而推断统计是现代统计学的核心和标志。



1.1.3 统计的应用

统计应用非常广泛，可以毫不夸张地说只要有数据的地方就会用到统计。比如，政府部门的管理与指导、公司或企业的生产管理以及居民的日常生活，等等。下面主要谈一谈统计方法在社会经济活动中以及其他学科当中的应用。

统计在社会经济活动中的应用有以下方面：

(1) 经济发展战略。政府制定长远的经济发展战略，一方面需要对宏观经济形势进行科学的判断与分析；另一方面需要对政府自身进行合理的发展定位，因地制宜科学发展等。所有这些都需利用统计方法来进行数据分析。

(2) 产品质量管理。产品质量管理部门使用统计方法，通过实验设计、质量控制和质量检验的路径来控制产品质量，避免生产更多的不合格品，从而取得更好的经济效果。

(3) 市场研究领域。企业营销部门使用统计方法对市场进行研究，估计偏爱某种产品的顾客比例，以及他们为什么喜欢这种产品，并且进行决策，采取何种战略最有效地扩大销售。

(4) 财务分析领域。财务部门一方面使用统计方法在不同的投资组合中进行决策，以相对较低的投资风险获得较高的投资利润；另一方面使用统计方法选择样本进行审计，并在成本会计中找出成本变动的原因。

(5) 经济预测领域。企业要对未来的市场状况进行市场调查与分析，经济学家要对宏观经济的未来走势进行分析，都需要利用统计信息、统计方法进行预测与判断。

(6) 人力资源管理。企业人力资源管理部门利用统计方法对员工的年龄、性别、受教育程度、工资等进行分析，为制订科学的薪酬计划、合理的奖惩制度提供依据。

……

统计学是收集、处理、分析、解释数据并从数据中得出结论的方法论学科，已被应用于诸多的学科之中，统计方法及其分析结果对我们的影响非常大。比如说，一篇学术论文能否在重要的学术期刊上发表，很大程度上取决于是否正确地使用了统计分析方法。马寅初先生说得好：学者不能离开统计而研究，政治家不能离开统计而施政，事业家不能离开统计而执业。统计学方法应用于其他学科的研究之中，与它们结合形成相互渗透的交叉学科，因而产生一系列专门领域的统计学，诸如，经济统计学、经营管理统计学、社会统计学、医药卫生统计学、生物工程统计学、天体物理统计学、工业统计学、金融统计学，等等。本书内容属于经济统计学。

1.2 统计数据

统计数据是用于描述和解释研究对象所收集、分析和汇总的事实和数字。它既是对现象或事物测量的结果，又是对现象或事物进行分析的基础。比如，对一个国家经



济活动总量的测量可得到国内生产总值（GDP）；对商品价格水平变动的测量得到商品价格指数；对产品质量的测量可得到甲等品、乙等品等。下面从不同的角度对统计数据进行分类。

1.2.1 分类数据和数值型数据

按照数据计量尺度的不同，统计数据可分为分类数据和数值型数据。

分类数据是只能归于某一类别的非数字型数据，它是对事物分类的结果，具体表现为类别，一般用文字来表现。分类数据根据其具体表现是否有序可分为无序分类数据和有序分类数据。**无序分类数据**的各类别之间是不可以排序的。比如，人口按性别可分为男性和女性；企业按所有制形式可分为国有企业、集体企业和民营企业等。这些数据的具体取值不存在顺序关系。**有序分类数据**也称为顺序数据，各类别之间可排序。比如，产品按质量等级可分为一等品、二等品、三等品和不合格品；你对某项实施的政策是赞成、反对，还是保持中立等。这些变量取值是存在顺序关系的。实际工作中，为了方便统计处理，有时可用数字代码来表示各分类数据。这可根据个人习惯或约定俗成的规则进行编码。比如，用“0”表示女性，“1”表示男性；“1”代表国有企业，“2”代表集体企业，“3”代表民营企业；等等。

数值型数据是按数字尺度对事物或现象测量的观察值，结果表现为具体的数值，它是对事物或现象进行的精确测度。比如，人的体重为60千克、70千克、75千克等；某教室的面积为78.63平方米；某商场某个季度商品销售额为158.32万元等。

无序分类数据和有序分类数据都是反映事物的品质特征，说明事物的类别，都是用文字来描述，因此也可统称为**定性数据**或**品质数据**。

数值型数据反映的是事物的数量特征，通常用数值来表现，因此也可称为**定量数据**或**数量数据**。



思政元素

2021年中国GDP规模达114.4万亿元（17.7万亿美元），名义增长12.8%，稳居全球第二大经济体；中国、美国占全球GDP比重分别超18%和23%。2021年中国实际GDP增长8.1%，两年平均增长5.1%；同期，美国GDP同比增长5.7%，德国2.8%、英国7.5%、日本1.7%。国际货币基金组织（IMF）预测2021年全球经济实际增速5.9%，中国经济发展处于全球领先地位。

中美经济规模差距正在快速缩小。美国2021年GDP为23.0万亿美元，中国GDP相当于美国的约77.0%。2000—2021年，中美经济规模差距从9.1万亿美元快速缩减至5万亿美元。



中国有望在十年左右跃升为第一大经济体。如果未来几年中国经济保持年均5.5%左右的增长，则将在2030年前后超越美国成为全球第一大经济体，再回世界之巅，全球政经格局和治理版图进入新的篇章。^①

1.2.2 观测数据和实验数据

按照取得数据方法的不同，统计数据可分为观测数据和实验数据。

观测数据是通过调查或观测所搜集到的数据，也可称为调查数据。这类数据不涉及研究人员对研究对象或事物行为模式的干预，也就是说是在没有人为控制或操作的条件下得到的，随着观测对象的不同，其形式和内容是多种多样的。一般来讲，反映社会经济现象的统计数据都属于观测数据。

实验数据是指在实验中控制实验对象而得到的数据。这类数据需要对在实验中控制一个或多个变量并测量操纵的结果。比如，我们给一组植物施肥，另一组植物不施肥，那么我们就是在控制植物土壤的成分，可以测量增长率、成活率等变量；再比如，新冠肺炎疫情发生后，各国开展了新冠疫苗的实验而得到疫苗有效率的实验数据。自然科学领域的统计数据多数来自于实验数据。

1.2.3 截面数据、时间序列数据和面板数据

按照数据与时间的关系，统计数据可分为截面数据、时间序列数据和面板数据。

截面数据是在相同或近似相同的同一时间点搜集的数据，这类数据描述的是同一时间上的同一现象在不同空间上的变化情况，因此又称为静态数据。例如，2020年全国各省、直辖市、自治区的地区生产总值就属于截面数据，它反映了2020年国内生产总值（GDP）在我国不同地区的变化。

时间序列数据是不同时间上搜集的数据，这类数据描述的是同一现象在不同时间上的动态变化情况，因此又称为动态数据。例如，2000—2020年我国国内生产总值（GDP）就属于时间序列数据，它反映了国内生产总值从2000年到2020年间的动态变化。

面板数据是同时在时间和截面（空间）上取得的二维数据，它是截面数据和时间序列数据的结合，这类数据反映的是同一现象在不同时间和不同空间的变化情况。例如，2000—2020年我国各省、直辖市、自治区（不含港、澳、台地区）的地区生产总值就属于面板数据，它反映了国内生产总值从2000年到2020年之间在我国各地区的变化情况。从某一年份看，它是由31个地区生产总值数字组成的截面数据；从某一省份看，它是由21年地区生产总值数据组成的时间序列数据。该面板数据由31个省份

^① 任泽平. 中美经济规模快速缩小，中国有望十年之内成为第一大经济体 [EB/OL]. (2022-03-01). 泽平宏观, 2022年3月1日, <https://i.ifeng.com/c/8E185Rz20RD>.

21 年的数据组成，共有 651 个观测值。

1.3 几个基本概念

统计学中概念很多，其中有几对概念非常重要，几乎贯穿于统计学的始终，因而有必要拿出来单独进行介绍。这些定义包括总体与样本、统计指标与指标体系、变量与变量值等。

1.3.1 总体与样本

1. 总体

在很多情况下，需要有关研究对象的个体的信息以便得出相关结论。比如，经济学家收集有关城镇居民家庭收入数据以了解全国居民家庭收入的变动情况；股票投资者收集上市公司财务状况的数据来分析股票价格指数的变动情况以决定是否投资；社会学家收集有关数据来了解人类行为；植物学家收集有关植物的情况以了解它们的生长状况。这时我们所关注的所有个体就形成了总体。因此可以认为，**总体**是指所研究对象的所有个体（元素）组成的集合。而构成总体的每个个体（元素）称为个体，亦可称为总体单位。在这里，所有居民家庭、所有的上市公司、所有的人口、所有的植物就称为总体，而每个居民家庭、每个上市公司、每个人、每种植物是总体中的每一个元素，称为个体。总体中所包含的个体（元素）数量可称为**总体单位数**，一般用大写的英文字母“ N ”表示。

总体范围的确定相对来讲比较容易，一般是根据研究目的来确定的。比如，要研究某城市工业生产情况，该城市所有工业企业的集合就是总体，该市每一个工业企业就是个体；要研究某高校在校生的生活费开支情况，该校所有的在校学生就形成总体，每一位在校学生就是个体；某企业要新上线一款新型汽车，其研究目的就是要了解这款新型汽车消费群体的状况，则购买这款新型汽车的潜在消费者就构成了我们的研究总体，每一位潜在消费者就是个体。

总体根据其总体单位数是否可数可分为**有限总体**和**无限总体**。有限总体是指总体中包含的总体单位数是可数的；无限总体是指总体中所包含的总体单位是不可数的。一般来讲，对社会经济领域开展研究，其总体多数属于有限总体，而对自然现象开展的研究就属于无限总体。比如，由工业企业、在校学生、潜在消费者构成的总体就是有限总体；而由于人类现有的认知能力有限，无法确定宇宙中行星的数量，因而对宇宙空间中行星的研究就属于无限总体。

2. 样本

总体是我们关注的所有个体组成的全体。对构成总体的所有个体都进行观察，通常是不现实的，也是不可能的。为此，我们只能根据研究对象的特点以及研究的要求，观察总体中一小部分个体的集合，以达到对研究现象总体的认识。这一小部分个体的



集合，就是人们常说的样本。**样本**就是从总体抽取的一部分个体（元素）的集合，构成样本的个体（元素）的数量称为**样本（容）量**，通常用小写的英文字母“ n ”表示。通过对样本中各元素的数据信息来推断总体特征。比如，从某高校在校学生中随机抽取 100 名学生，这 100 名学生就构成了一个样本，根据这 100 名学生的月均生活费支出来推断该校所有在校学生的月均生活支出。

在这里要提请大家注意的是，样本量与样本的可能数目是两个不同的定义。样本量是指样本中包含个体（元素）的数量，而样本的可能数目是指从总体中随机抽取一定量的样本可能形成多少个样本，这与抽样的方式方法有非常密切的关系。比如，从 $N=4$ 的总体中，随机抽取容量 $n=2$ 的样本，可能形成多少个样本呢？大家可思考一下！

1.3.2 统计指标与统计指标体系

1. 统计指标

对于总体和样本，我们所关心的只是它们的数据特征，而不是组成总体的个体（元素）本身。比如，要检验一批灯泡的质量，我们所关注的并非灯泡的数量多少，而是相关灯泡质量的数据，即通过一系列数据来假设、评价、度量灯泡的质量。这里的数据特征就是通常意义上的统计指标。因此，统计指标可定义为用来反映总体或样本特征的定义和数值，它是研究人员想要了解的某一特征值。比如，2021 年我国国内生产总值（GDP）为 1 133 239.8 亿元，人均 GDP 为 80 976 元，这些指标从某一侧面概括性地反映了我国国民经济发展的数量特征。由此可见，**统计指标**是由指标的定义（名称）和指标取值两项基本要素组成的。指标的定义是对所研究总体或样本数量特征的概括，也是对总体或样本的质的规定性；指标的取值反映所研究总体或样本在不同时间、地点、条件下的规模和水平。总之，统计指标是对所研究对象的具体化，也是统计对客观事物认识的起点。其实，总体或样本数量特征的概念也可以用参数或统计量来反映，关于参数与统计量的内容参见第 5 章抽样分布理论。

统计指标按照其表现形式与内容可分为总量指标、相对指标和平均指标。

总量指标表明所研究现象（总体或样本）总规模、总水平的指标，一般用绝对数形式表示，因此又称为绝对数指标，如人口总数、企业总数、职工总人数、工资总额、国内生产总值、商品销售额、商品进出口总额等，这些指标反映现象或过程的总规模和总水平。

相对指标表明现象（总体或样本）内部的数量对比关系或普遍程度的指标，一般用相对数形式表示，因此又称为相对数指标，如人口出生率、人口自然增长率、职工劳动生产率、人口密度、资产贡献率、万元 GDP 耗电量、普通高等院校生师比等，这类指标概括性地反映现象内部的数量对比关系或普遍程度。

平均指标表明现象（总体或样本）内部各个个体单位某一特征值一般水平的指标，一般用平均数形式表示，因此又称为平均数指标，如平均身高、平均体重、职工的平均工资、学生的平均成绩、灯泡的平均使用时间等。这类指标是根据所研究现象



内部个体单位某一特征值计算得来的，其计算较为复杂，具体请见本书第4章数据分布特征的度量。

统计指标按照其反映的数量特点（性质）可分数量指标和质量指标。

数量指标表明所研究现象（总体或样本）总规模、总水平的指标，习惯上将总量指标归入数量指标的范畴。

质量指标是表明所研究现象（总体或样本）相对水平和工作质量的指标，习惯上将相对指标和平均指标归入质量指标的范畴。



思政元素

数据显示，2021年，我国国内生产总值达1 143 670亿元，比上年增长8.1%，超额完成了“国内生产总值增长6%以上”的目标。此外，2021年，我国粮食总产量为6 828.5亿千克，比上年增加133.5亿千克，增长2%，粮食总产量连续7年保持在0.65万亿千克以上，超额完成目标。

数据还显示，2021年全国居民消费价格指数（CPI）同比上涨0.9%，这一终值远低于《政府工作报告》设定的“居民消费价格涨幅3%”^①左右的目标。在当前全球通胀压力严峻的大背景下，超额完成物价调控目标实属不易。

2. 统计指标体系

一个统计指标只能反映现象的某个数量特征，说明现象某一侧面的情况。客观现象是复杂多样的，要反映其全貌，描述其发展变化的全过程，单纯依靠某单个统计指标显然是不够的，需要一系列统计指标来反映，而这一系列统计指标便形成了指标体系。**统计指标体系**是由一系列相互联系的统计指标所组成的有机整体，用以反映所研究现象各方面相互依存、相互制约的关系。例如，经济发展与经济实力的统计指标体系、居民生活与健康水平的统计指标体系、技术进步与科技创新的统计指标体系等。统计指标体系的设置不仅是对客观现象的反映，更是人们对客观现象认识的结果，随着客观形势的发展变化以及实践经验和理论认识的积累，统计指标体系也将不断改进、更新而逐步完善。

1.3.3 变量与变量值

为了更好地对我们感兴趣的对象进行研究，还需要了解的另一个重要定义是变量。**变量**是说明现象某种特征、特质或属性的概念，如人的性别、寿命、体重、受教育程度，商品销售额、产品质量等级等都是变量。变量的具体取值叫**变量值**。比如，人的性别可以取男性和女性两个值，人的体重可以是60千克、65千克、70千克等，这里

^① 2022年3月5日第十三届全国人民代表大会第五次会议上李克强总理所作《政府工作报告》。



的男性、女性、具体数字就是变量值。通常，在进行数据分析时，需要确定变量及其取值变化，因此，我们可对变量进行如下分类。

变量按照其取值的变化不同可分为分类变量和数值型变量。

分类变量是说明事物类别的名称，其取值为分类数据，可分为无序分类变量和有序分类变量。**无序分类变量**取值的各类别间是不可排序的。比如，“性别”就是一个无序分类变量，其变量取值为“男”或“女”；“产业”也是一个无序分类变量，其变量取值为“第一产业”“第二产业”“第三产业”等。这些变量取值不存在顺序关系。**有序分类变量**也称顺序变量，其取值的各类别之间可以排序。比如，“产品质量等级”就是一个有序分类变量，其取值为“甲等品”“乙等品”“丙等品”等；“你对某项实施的政策的态度”也是一个有序分类变量，其取值可以是“赞成”“中立”“反对”等，这些变量取值是可以排列顺序的。还有诸如“受教育程度”“居民生活的贫困或富裕程度”“发达国家与发展中国家的划分”等都属于有序分类变量，不胜枚举。

数值型变量是说明事物数量特征的名称，其取值为数值型数据，如“产品产量”“国内生产总值”“商品进出口额”“身高”“体重”等都属于数值型变量。数值型变量根据其取值是否连续，又可以划分为离散型变量和连续型变量。离散型变量是变量的值只能通过计数的方法取得，它是不连续的，只能取有限个值，且只能用整数来表示，能够一一列举，如“企业个数”“职工人数”“在校学生人数”等就属于离散型变量。连续型变量是在数轴上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，可以用整数或小数来表示，无法一一列举，如“年龄”“身高”“零件的尺寸”“温度”“轮胎的里程”“灯泡的使用寿命”等就属于连续型变量。

变量按照其变化是否确定可分为确定型变量和非确定型变量。

确定型变量是受确定性因素影响的变量，即影响变量值变化的因素是明确的，是可解释和控制的。**非确定型变量**是受不确定因素（随机因素）影响的变量，变量的取值是无法确定的，影响因素是不明的，也不好控制，又称为随机变量。社会经济领域内的现象既有确定性变量也有随机变量。

上面所讨论的变量都与我们所熟悉的日常生活有关，我们把这种变量称为经验变量。**经验变量**描述的是我们周围可观测到的物质世界中的事物。除了经验变量，我们还使用由统计学、数学方法构造出来的变量，这类变量可称为**理论变量**。比如，我们在以后的章节中将要介绍的 Z 变量、 t 变量、 χ^2 变量和 F 变量等。



本章小结

统计与数字打交道，统计是人们认识客观世界总体数量特征和数量变化规律活动的统称。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。其内容包括：收集数据、处理数据、分析数据、解释数据、得到结论。描述统计



是推断统计的基础，推断统计是现代统计学的核心和标志。统计应用非常广泛，只要有数据的地方就会用到统计。统计可在社会经济活动中应用，也可在其他学科中应用。本书内容属于经济统计学。

统计数据是用于描述和解释研究对象所收集、分析和汇总的事实和数字。它既是对现象或事物测量的结果，又是对现象或事物进行分析的基础。按照不同的分类规则，统计数据可以分为分类数据和数值型数据，观测数据和实验数据，截面数据、时间序列数据和面板数据。

总体是指所研究对象的所有个体（元素）组成的集合，而构成总体的每个个体（元素）称为个体，亦称为总体单位。总体中所包含的个体（元素）的数量称为总体单位数，一般用大写的英文字母“ N ”表示。总体有有限总体与无限总体之分。样本是从总体抽取的一部分个体（元素）的集合，构成样本的个体（元素）的数量称为样本（容）量，通常用小写的英文字母“ n ”表示。统计指标是用来反映总体或样本特征的定义和数值，它是研究人员想要了解的某一特征值。

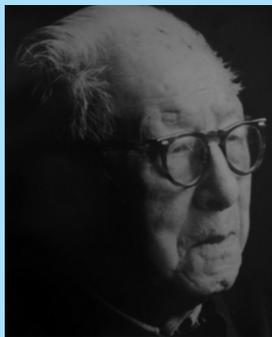
统计指标按照其表现形式与内容可分为总量指标、相对指标和平均指标。统计指标按照其反映的数量特点（性质）可分数量指标和质量指标。统计指标体系是由一系列相互联系的统计指标所组成的有机整体，用以反映所研究现象各方面相互依存、相互制约的关系。

变量是说明现象某种特征、特质或属性的概念。变量值是指变量的具体取值。变量按照其取值的变化不同可分为分类变量和数值型变量。变量按照其变化是否确定可分为确定型变量和非确定型变量。



思政园地

统计需要实事求是——人口统计学家戴世光



戴世光（1908—1999年），湖北江夏人，1931年毕业于清华大学经济系，1936年获美国密歇根大学数理统计硕士；曾任西南联合大学、清华大学、中国人民大学教授；人口学家，更是我国20世纪杰出的统计学家。

30多年前笔者机缘巧合之下，有幸聆听了戴世光教授一场学术讲座，时间大约是1987年，地点在中南财经大学供研究生上课的小教室，条件相当简陋，既没有麦克风，也没有现在的PPT，全凭口头表述。戴教授虽然是年近80的耄耋老人，但自始至终精神矍铄，声音洪亮，激情澎湃，并不时借助于粉笔将主要观点或关键词写在黑板上，其演讲的画面犹在眼前。讲座主要讲了三个问题：一是1978年冬峨眉山会议以来我国统计学界的发展变化，二是统计学与



数理统计学的相关学术观点，三是统计学需要实事求是。尤其是第三个问题，令人印象深刻，因为讲的时间比较长，对“实事求是”这一定义作了较为全面而科学的解释。“实事”就是要从实际出发，反映客观事物的基本状况；“求”就是寻找、追求、探寻；“是”就是事物的客观规律、内在的数量特征。实事求是是统计学的精髓，统计学就要实事求是。戴教授最后指出统计数据要实事求是，统计分析方法要实事求是，统计分析结论要实事求是，统计工作人员更要实事求是。这不正是现代统计学所强调的研究内容与要求吗？之后笔者又陆续拜读了戴教授的相关著作，其中两篇文章《积极发展科学的统计学为我国早日实现四个现代化服务》和《实事求是研究统计科学问题的指导思想》，其中的内容与观点更是引发了我国20世纪80年代统计学界长达十多年的学术大讨论。

戴教授，1940—1992年50余年一直从事教学科研工作，实事求是，勤于耕耘，成果丰硕，为统计学界及其从业人员实事求是地从事开展统计教学、科研、统计实务工作树立了很好的示范作用。

► 讨论：

1. 从以上示例中，你学到了什么？
2. 统计学是在统计实践的基础上产生发展起来的方法论科学，其中最为关键与重要的问题是什么？



思考与练习

一、思考题

- 1.1 什么是统计学？
- 1.2 解释描述统计和推断统计，并描述两者之间的关系。
- 1.3 举出你所知道的2个或3个统计应用的例子。
- 1.4 统计数据可分为哪几类？其分类的标准是什么？
- 1.5 解释截面数据、时间序列数据、面板数据的含义，并各举一例。
- 1.6 解释总体、个体和样本的含义，并描述它们之间的关系。
- 1.7 统计指标按计量形式与内容可分为哪几种？
- 1.8 如何理解统计指标体系，并举例说明。
- 1.9 什么是变量？如何理解理论变量与经验变量，并举例说明。

二、练习题

- 1.1 指出下面变量的类型。
(1) 性别。



(2) 年龄。

(3) 手机销售量。

(4) 为了防止新冠肺炎疫情的传播,小区物业管理部门决定对进入小区的人员进行登记,居民对这项措施的态度是(赞成、中立、反对)。

(5) 购买商品时的支付方式(现金、信用卡、支付宝或微信)。

1.2 2020年3月15日,某电视台公布了来自不同制造商的120款大屏幕电视机的检测报告,据此谈谈总体、样本和个体分别是什么。

1.3 假定120款大屏幕电视机的检测报告中,公布了每款电视机的制造商、生产成本、屏幕尺寸、类型(液晶、等离子、背投),以及综合得分(0~100分)。请问:这里涉及几个变量?哪些是分类变量,哪些是顺序变量,哪些是数值型变量?

1.4 某高校学生会从该校在校学生中随机抽取200名学生进行调查,其中,60%的学生回答他们月生活费支出在1500元以上,80%的学生回答他们的消费支付方式是校园卡。请回答以下问题:

(1) 这一研究的总体是什么,样本是什么?

(2) 月生活费支出是分类变量、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) 这一调查涉及的是截面数据、时间序列数据,还是面板数据?

1.5 《金融时报》进行了一次民意测验,共有2560名成年人进行参加了这次调查。其中一个问题是:您如何评价银行的服务态度?可能的答案是:优秀、良好、中等、较差和非常差。请回答下面的问题:

(1) 这项调查的样本容量是多大?

(2) 数据是分类型数据、顺序型数据还是数值型数据?

(3) 对这一问题,用绝对数还是相对数汇总数据,哪个更合适?

(4) 答案中有10%的人说银行服务态度优秀,有多少人给出了这种答案。

1.6 一项调查表明,消费者每月网上购物的平均支出为2500元,他们选择网上购物的主要原因是“价格便宜且方便”。回答下列问题:

(1) 这一研究的总体是什么?

(2) 消费者网上购物的主要原因是“价格便宜且方便”,它属于分类变量、顺序变量还是数值型变量?

(3) “消费者每月网上购物的平均支出为2500元”是经验变量还是理论变量?

(4) 研究者使用的是描述统计还是推断统计?



单选题



判断题