

普通高等教育“十二五”规划教材

公共基础课系列

统计学

S t a t i s t i c s

——数据处理与分析

王德发 刘小锋 主编



 上海财经大学出版社

普通高等教育“十二五”规划教材·公共基础课



统计学

——数据处理与分析

王德发 刘小锋 主编



上海财经大学出版社

图书在版编目(CIP)数据

统计学:数据处理与分析/王德发,刘小锋主编. —上海:上海财经大学出版社,2013. 9

(普通高等教育“十二五”规划教材·公共基础课系列)

ISBN 978-7-5642-1680-1/F·1680

I. ①统… II. ①王…②刘… 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(2013)第 160902 号

- ☐ 责任编辑 施春杰
- ☐ 封面设计 钱宇辰
- ☐ 投稿热线 apin001@163.com
- ☐ 订购电话 021-65904705
- ☐ 责任校对 赵伟 林佳依

TONGJIXUE

统 计 学

——数据处理与分析

王德发 刘小锋 主编

上海财经大学出版社出版发行
(上海市武东路 321 号乙 邮编 200434)

网 址: <http://www.sufep.com>

电子邮箱: webmaster @ sufep.com

全国新华书店经销

上海华教印务有限公司印刷装订

2013 年 9 月第 1 版 2013 年 9 月第 1 次印刷

787mm×1092mm 1/16 18 印张 460 千字
印数: 0 001—4 000 定价: 38.00 元

P 前言

REFACE

本书是 2012 年 10 月上海财经大学出版社出版的王德发主编的《统计学》的配套学习辅导书。统计学作为一门方法论科学,统计方法繁多,适用条件各异,怎样根据数据的特性找到最恰当的统计方法?这是有一定难度的。本书汇集了近 600 道习题,可以帮助学生系统地理解统计学的基本思想和方法,也可以作为统计学教师和统计工作者的参考资料。此外,由于计算机科学的快速发展,各种统计软件不断涌现,为此,本书针对 Microsoft Office 中的 Excel 统计模块的数据处理功能进行介绍,在加强统计学抽象概念直观化的基础上,提高学生的数据处理能力。

本书主要分为两大部分:第一部分由每个章节的学习目的和要求、内容提要、课后习题答案、课外习题和课外习题参考答案五块内容组成;第二部分是 Excel 统计数据处理,即统计学的实验指导。另外,本书还附加了一份综合案例和一份考试样卷及样卷答案。

本书第一部分各章节的分工如下:刘旻编写第一、二、三章;奚欢编写第五、九章;罗平编写第六、十一章;胡志明编写第七、八章;陈婧编写第四章;詹欢编写第十章。王德发教授编写考试样卷及答案;刘小锋编写第二部分和综合案例。

在本书的编写过程中,得到了上海财经大学浙江学院金融与统计系的汤美微、应芥舟和陈婧(注:我系有两位同名同姓,都取名陈婧的教师)的支持和帮助,在此一并致以谢意。

编者

2013 年 3 月 8 日

第一部分 学习指导

第一章 绪论/3

- 一、学习目的和要求/3
- 二、内容提要/3
- 三、课后习题答案/4
- 四、课外习题/5
- 五、课外习题参考答案/8

第二章 统计调查/12

- 一、学习目的和要求/12
- 二、内容提要/12
- 三、课后习题答案/13
- 四、课外习题/15
- 五、课外习题参考答案/17

第三章 统计数据处理/21

- 一、学习目的和要求/21
- 二、内容提要/21
- 三、课后习题答案/22
- 四、课外习题/23
- 五、课外习题参考答案/26

第四章 数据描述性分析/30

- 一、学习目的和要求/30
- 二、内容提要/30
- 三、课后习题答案/35
- 四、课外习题/40
- 五、课外习题参考答案/45

第五章 概率论基础/49

- 一、学习目的和要求/49

二、内容提要/49

三、课后习题答案/52

第六章 抽样分布/54

一、学习目的和要求/54

二、内容提要/54

三、课后习题答案/58

四、课外习题/61

五、课外习题参考答案/64

第七章 参数估计/70

一、学习目的和要求/70

二、内容提要/70

三、课后习题答案/72

四、课外习题/78

五、课外习题参考答案/81

第八章 假设检验/86

一、学习目的和要求/86

二、内容提要/86

三、课后习题答案/91

四、课外习题/102

五、课外习题参考答案/106

第九章 相关分析与回归分析/112

一、学习目的和要求/112

二、内容提要/112

三、课后习题答案/115

四、课外习题/137

五、课外习题参考答案/141

第十章 时间序列分析/144

一、学习目的和要求/144

二、内容提要/144

三、课后习题答案/149

四、课外习题/155

五、课外习题参考答案/160

第十一章 指数分析/166

- 一、学习目的和要求/166
- 二、内容提要/166
- 三、课后习题答案/168
- 四、课外习题/171
- 五、课外习题参考答案/175

第二部分 Excel 统计数据处理

- 一、Excel 统计数据处理概述/183
- 二、统计数据的收集(教材第二章)/185
- 三、统计数据处理(教材第三章)/188
 - (一) 数据编码/188
 - (二) 数据的录入/188
 - (三) 数据的筛选/190
 - (四) 数据排序/192
 - (五) 设计条件格式/193
 - (六) 数据分类汇总/195
 - (七) 创建图表和增强图表/196
 - (八) 透视表/197
 - (九) 直方图/200
- 四、数据描述性分析(教材第四章)/203
 - (一) 函数方法/203
 - (二) 描述统计工具量的使用/204
 - (三) 两组或两组以上数据描述统计结果分析/205
- 五、概率论基础(教材第五章)/206
 - (一) 二项分布概率的计算和概率曲线图的编制/206
 - (二) 正态分布概率的计算和概率曲线图的编制/210
 - (三) 动态概率分布曲线的编制/212
 - (四) 全概率公式和贝叶斯公式的应用/220
- 六、抽样分布(教材第六章)/223
- 七、参数估计(教材第七章)/226
 - (一) 用公式计算出置信区间/226
 - (二) 编制活动工作表进行参数的统计量估计/228
- 八、假设检验(教材第八章)/231
 - (一) 编制活动工作表进行一个总体参数的假设检验/231
 - (二) Z -检验工具的使用/232
 - (二) t -检验工具的使用/234

(三) F -检验工具/239

(四) 方差分析/240

九、相关分析与回归分析(教材第九章)/242

(一) Excel 的回归分析工具/242

(二) 相关分析/248

十、时间序列分析(教材第十章)/252

(一) 用 Excel 进行预测/252

(二) 用 Excel 进行时间序列分析/257

十一、指数分析(教材第十一章)/261

(一) 用 Excel 计算总指数/261

(二) 用 Excel 计算平均数指数/262

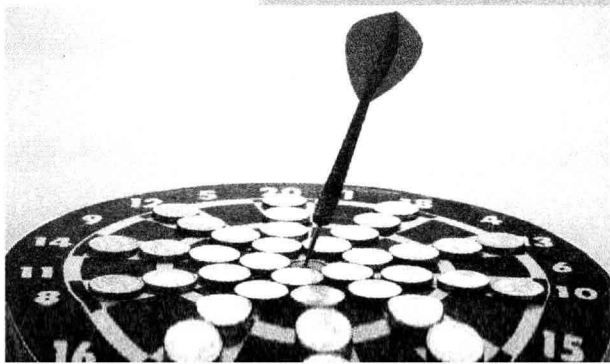
综合案例分析/263

第三部分 样卷

《统计学》期末考试样卷/275

《统计学》期末考试样卷答案/278

第一部分 学习指导



第一章

绪 论

一、学习目的和要求

- (1) 了解“统计”一词的渊源及其含义;
- (2) 了解统计学是如何产生的;
- (3) 了解统计学的性质和特点有哪些;
- (4) 掌握统计学的内容和一些基本概念。

二、内容提要

(一) “统计”一词的渊源及其含义

“统计”一词的渊源:“统计”(statistics)一词最早由德国马尔堡和格丁根大学教授阿痕瓦尔(Gottfried Achenwall, 1719~1772)提出,意思是指各种现象的状态和状况。我国古代的“统计”一词,原为“总计、合计”之意。具有现代意义的“统计”一词,则是 20 世纪初从日本传过来的。

“统计”一词的三种含义:统计活动、统计资料和统计科学。

(二) 统计学的产生与发展

统计学是伴随着统计活动的产生和发展而形成的,它的发展过程始终是沿着两条主线展开的:一是以“政治算术学派”为开端形成和发展起来的、以社会经济问题为主要研究对象的社会经济统计;二是以概率论的研究为开端并以概率论为基础形成和发展起来的、以方法和应用研究为主的数理统计。

在统计学的产生和发展过程中,产生了许多不同的统计学派,如形成于德国的“国势学派”,形成于英国的“政治算术学派”,产生于 19 世纪中叶的“数理统计学派”、“社会统计学派”等等。

（三）统计学的性质和特点

近年来,我国统计学界提倡讨论的观点,主张统计学只有一门,包括数理统计学和社会经济统计学。它既不同于数学,也不同于经济学等其他实质性科学,而是一门研究客观事物数量方面的独立的方法论科学。

统计学研究的特点:(1)从研究对象来看,它研究客观事物的数量方面。(2)从研究方法来看,它强调对客观事物总体进行大量观察,通过归纳和推理以获得总体数量方面的综合性认识。(3)从学科体系来看,统计学是一门多科性的科学,是一个学科“家族”。

（四）统计学的内容和一些基本概念

统计学的内容:描述统计学和推断统计学。

统计分析中的几个重要概念:总体与总体单位(个体)、指标、变量。

三、课后习题答案

（一）思考题

1. 现代的“统计”一词通常包括统计活动、统计资料和统计科学三种含义。

(1) 统计活动是指人们根据一定的目的,采用相应的统计方法收集资料、整理资料和分析资料的工作过程。

现在各国的政府统计,包括经济、社会 and 科技等方面的统计;联合国统计机构的活动,如世界经济发展和环境统计、各国经济比较等;以及工矿企业或经营服务部门对生产经营过程中的投入和产出所进行的各种计量行为都是统计活动。

(2) 统计资料是统计活动的最终成果,例如某报报道:“2011 年某市实现国内生产总值 1 543.63 亿元,按可比价格计算,比去年同期增长 13.7%,增幅比去年同期高 1.6 个百分点,比全国平均增长速度约高 4 个百分点……”

统计资料以较为专业的形式出现的有:联合国统计年鉴,统计月刊,各国出版发行的经济年鉴等。

(3) 统计学则是一门指导统计工作的方法论科学,是统计活动的经验总结。

三者关系为:统计科学与统计活动之间的关系是理论与实践的关系,统计资料则是统计活动的成果。

2. 统计学作为一门系统的科学,距今只有 300 多年的历史,产生于欧洲资本主义社会初期,经历了古典统计学、近代统计学和现代统计学三个时期,先后产生过政治算术学、国势学、社会统计学和数理统计学等重要学派。

数理统计学从其发展的时间进程看,可分为三个时代:卡尔·彼尔逊统计时代、R. A. 费雪统计时代和 J. 奈曼—E. S. 彼尔逊统计时代。1946 年,瑞典统计学家克拉默尔发表了《统计学的数学方法》一书,用测度论系统总结了数理统计的发展,标志着统计学开始由近代转向现代。因此,上述三个时代的后两个时代可归入现代统计学范畴。

3. 描述统计(descriptive statistics)是用来描绘(describe)或总结(summarize)观察量的基本情况的统计方法。

推断统计是指依据样本资料推断总体特征的技术和方法,包括参数估计和假设检验的方法。

描述统计与推断统计紧密联系,描述统计是推断统计的前提,推断统计是描述统计的发展。

4. 统计总体(population)简称总体,是统计研究所确定的客观对象,它是根据一定的研究目的,在同质基础上,由客观存在的许多个别单位所组成的一个整体。

总体单位(unit),又称个体,是那些具有某种共同性质并组成总体的个别单位,是各项统计数字的原始承担者。总体单位既有同质性,又有差异性。

同质性是各个总体单位所具有的某一种共同的属性或特征,是总体赖以形成的客观基础;因为每一个总体单位除了具有一个与其他总体单位性质相同的属性或特征以外,还有许多其他总体单位所不具备的属性或特征,所有总体单位所具有的各种不同的数量特征、属性或数量关系构成了总体内部的差异性。

5. 确定性变量受确定性的因素影响,也就是说,影响变量变化的因素是明确的、可解释的。在这种因素的影响下,变量的变化幅度、变化方向是可以确定的。例如,职工的工资总额一般受职工人数和职工平均工资两个因素的影响,它们各自的变化及两者之间的交互作用对工资总额的变动影响是确定的。

随机变量受随机因素的影响,随机因素的变化具有不确定性和偶然性,如气候的变化、海水潮汐的涨落等。但是,变量的随机性并不意味着人类对之无能为力,统计要研究的正是现象的不确定性,就是要依据不确定因素的相互抵消的特性通过大量观察来研究随机现象的某种规律性。

6. 离散型变量只能取整数,如人数、企业数、学校数等,计量单位一般为自然单位,如个、匹、头等。

连续型变量则可以取任意小数,如产值、销售额、固定资产投资额等,计量单位为价值单位;或可以取任意两个变量中的一个无限小数,大多为某种测量器具的测量结果,一般用度量衡单位表示,如某种零件的大小尺寸、直径、重量、体积、容积等。

四、课外习题

(一) 单项选择题

1. 人们在使用“统计”一词时,通常有三种不同的含义,其中不正确的一项是()。

- A. 统计工作 B. 统计方法 C. 统计数据 D. 统计学

2. 统计是从社会经济现象的()。

- A. 质的方面去研究其量的方面 B. 量的方面去研究其质的方面
C. 质与量的辩证统一中研究其量的方面 D. 质与量的辩证统一中研究其质的方面

3. 根据统计方法的构成,可将统计学分为()。

- A. 描述统计学和推断统计学 B. 描述统计学和应用统计学
B. 理论统计学和推断统计学 D. 理论统计学和应用统计学

4. 统计学研究中使用的方法是()。

- A. 纯粹的演绎 B. 纯粹的归纳
C. 演绎与归纳相结合,以归纳为主导 D. 演绎与归纳相结合,以演绎为主导

5. 政治算术学派的创始人之一是()。

- A. 高尔登 B. 凯特勒 C. 约翰·格朗特 D. 彼尔逊

6. 为运用数学理论研究概率论开辟道路的论著是()。

- A. 《政治算术》 B. 《概率论书简》 C. 《论赌博》 D. 《概率论分析》

7. 在统计史上被认为有统计学之实而无统计学之名的统计学派是()。
- A. 数理统计学派 B. 政治算术学派 C. 社会统计学派 D. 国势学派
8. 统计对社会经济现象总体数量的认识是()。
- A. 从定性到定量 B. 从定量到定性 C. 从个体到总体 D. 从总体到个体
9. 最早把统计学引入我国的是()。
- A. 横山雅男 B. 高野岩三郎 C. 田井要助 D. 犹尔
10. 统计认识活动的过程逻辑上就是()的过程。
- A. 理性认识 B. 感性认识 C. 归纳推理 D. 大量观察
11. ()反映的是现象总体的相对水平或工作质量,一般表现为相对数或平均数。
- A. 数量指标 B. 质量指标 C. 实物指标 D. 价值指标
12. 被认为是推断统计学基础理论奠基者的是()。
- A. 皮埃尔 B. 汤姆斯·辛普森 C. 彼尔逊 D. 费雪
13. 第一位在数量统计和概率论领域享有国际声誉的中国数学家是()。
- A. 许宝騄 B. 金国宝 C. 朱君毅 D. 魏宗舒
14. 比利时统计学家、数学家和天文学家(),把德国的国势学、英国的政治算术和意大利、法国的古典概率论结合起来,开创了近代统计学的新纪元。
- A. 高尔登 B. 凯特勒 C. 约翰·格朗特 D. 彼尔逊
15. ()于 1809 年提出最小二乘法,发现了正态分布,从而奠定了统计估计的理论基础。
- A. 高斯 B. 凯特勒 C. 约翰·格朗特 D. 彼尔逊
16. 设某地区有 800 家独立核算的工业企业,要研究这些企业的产品生产情况,总体单位是()。
- A. 全部工业企业 B. 800 家工业企业
- C. 每一件产品 D. 800 家工业企业的全部工业产品
17. 有 200 家公司每位职工的工资资料,如果要调查这 200 家公司的工资水平情况,则统计总体为()。
- A. 200 家公司的全部职工 B. 200 家公司
- C. 200 家公司职工的全部工资 D. 200 家公司每个职工的工资
18. 以产品等级来反映某种产品的质量,则该产品等级是()。
- A. 数量标志 B. 数量指标 C. 品质标志 D. 质量指标
19. 某工人月工资为 550 元,工资是()。
- A. 品质标志 B. 数量标志 C. 变量值 D. 指标
20. 某班四名学生的金融考试成绩分别为 70 分、80 分、86 分和 95 分,这四个数字是()。
- A. 标志 B. 指标值 C. 指标 D. 变量值
21. 年龄是()。
- A. 变量值
- B. 离散型变量
- C. 连续型变量,但在应用中常按离散型变量处理
- D. 连续型变量
22. 工业企业的职工人数、职工工资是()。
- A. 连续型变量 B. 离散型变量
- C. 前者是连续型变量,后者是离散型变量 D. 前者是离散型变量,后者是连续型变量

(二) 多项选择题

1. 统计的含义一般有()。

- A. 统计工作 B. 统计会议 C. 统计学
D. 统计报刊 E. 统计资料

2. 政治算术学派的创始人有()。

- A. 阿罕华尔 B. 凯特勒 C. 威廉·配第
D. 恩格尔 E. 约翰·格朗特

3. 根据统计方法的构成,可将统计学分为()。

- A. 应用统计学 B. 推断统计学 C. 理论统计学
D. 描述统计学 E. 统计数学

4. 下列属于应用统计学范畴的是()。

- A. 生物统计学 B. 农业统计学 C. 卫生统计学
D. 社会统计学 E. 人口统计学

5. 对社会经济统计的产生和发展有一定影响的人物主要有()。

- A. 威廉·配第 B. 凯特勒 C. 李·普莱
D. 恩格尔 E. 约翰·格朗特

6. 下列属于描述统计中常用的图示法有()。

- A. 多边图 B. 直方图 C. 圆形图
D. 散点图 E. 全距

7. 下列属于数量指标的有()。

- A. 人口总数 B. 平均年龄 C. 国内生产总值
D. 工资总额 E. 平均身高

8. 下列属于基本统计指标的有()。

- A. 社会指标体系 B. 能源指标体系 C. 经济指标体系
D. 固定资产投资指标体系 E. 教育指标体系

9. 下列属于专题统计指标的有()。

- A. 社会指标体系 B. 能源指标体系 C. 经济指标体系
D. 固定资产投资指标体系 E. 教育指标体系

10. 下列属于离散型变量的有()。

- A. 产值 B. 人数 C. 学校数
D. 企业数 E. 销售额

(三) 判断题

1. 统计学是随着人类社会的发展和社会管理的需要而发展起来的。 ()

2. 统计数据的整理是对统计数据的加工处理过程,是统计学的核心内容。 ()

3. “政治算术学派”的创始人是威廉·配第和拉普拉斯。 ()

4. 比利时统计学家凯特勒对统计学的主要贡献是将自然科学的研究方法引进社会现象的研究中来。 ()

5. 以社会经济问题为主要研究对象的社会经济统计是以“国势学派”为开端形成和发展起来的。 ()

6. 数学与统计学一样,并不能独立地直接研究和探索客观现象的规律,而是给各学科提供了一种研究和探索客观规律的数量方法。 ()

7. 概率论最初的研究是为赌徒们找出掷骰子取胜的一套办法,所以,数理统计学研究领域最早的论著为《论赌博》。 ()

8. 运用统计方法可以研究各学科的规律和解决其具体问题。 ()

9. 统计研究事物的数量方面,指的是个别事物的数量方面,而不是总体的数量方面。 ()

10. 总供给和总需求的平衡关系属于数量界限。 ()

(四) 名词解释

- | | | |
|--------------|----------|--------------|
| 1. 描述统计 | 2. 推断统计 | 3. 参数统计学 |
| 4. 非参数统计学 | 5. 统计总体 | 6. 总体单位 |
| 7. 有限总体 | 8. 无限总体 | 9. 变量 |
| 10. 确定性变量 | 11. 随机变量 | 12. 离散型变量 |
| 13. 连续型变量 | 14. 统计指标 | 15. 国内生产总值 |
| 16. 数量指标 | 17. 质量指标 | 18. 基本统计指标体系 |
| 19. 专题统计指标体系 | 20. 正指标 | 21. 逆指标 |

(五) 简答题

1. 简述“统计”一词的含义。
2. 数理统计学从其发展的时间进程看,可以分为哪三个时代?
3. 社会经济统计研究的对象有哪些?
4. 统计学研究有什么特点?
5. 统计指标有什么特点?
6. 什么是指标体系?分类有哪些?
7. 为什么说统计学是一门多科性的科学?
8. 简述有限总体与无限总体的区别。
9. 简述统计指标的含义。
10. 总体单位有哪些特性?

五、课外习题参考答案

(一) 单项选择题

- | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 1. B | 2. D | 3. A | 4. C | 5. C | 6. C | 7. B | 8. C |
| 9. A | 10. C | 11. B | 12. D | 13. A | 14. B | 15. A | 16. D |
| 17. A | 18. C | 19. B | 20. D | 21. C | 22. D | | |

(二) 多项选择题

- | | | | | |
|---------|--------|-------|----------|---------|
| 1. ACE | 2. CE | 3. BD | 4. ABCDE | 5. ACDE |
| 6. ABCD | 7. ACD | 8. AC | 9. BDE | 10. BCD |

(三) 判断题

1. $\checkmark$ 2. $\times$ 3. $\times$ 4. $\checkmark$ 5. $\times$ 6. $\checkmark$ 7. $\checkmark$ 8. $\times$ 9. $\times$ 10. $\times$

(四) 名词解释

1. 描述统计是用来描绘或总结观察量的基本情况的统计方法。描述统计的内容包括统计数据的收集、数据的加工处理、数据的显示、数据分布的特征概括等。

2. 推断统计是指依据样本资料推断总体特征的技术和方法,包括参数估计和假设检验的方法。

3. 参数统计学,又称推论统计学,依照总体条件的差异性又可分为参数统计学和非参数统计学。其中,参数统计学是指总体呈正态分布的统计推论方法。

4. 非参数统计学:推论统计学依照总体条件的差异性又可分为参数统计学和非参数统计学。其他所有应用于非正态分布总体的统计推论方法,都称为非参数统计学。

5. 统计总体(population)简称总体,是统计研究所确定的客观对象,它是根据一定的研究目的,在同质基础上,由客观存在的许多个别单位所组成的一个整体。

6. 总体单位(unit),又称个体,是那些具有某种共同性质并组成总体的个别单位,是各项统计数字的原始承担者。总体单位既有同质性,又有差异性。

7. 有限总体是指总体所包括的总体单位是可数可列的,即这些总体单位是数得清而且可以按一定的序号排列的。社会经济总体大多是有限总体,如人口数、企业数、学校数等。

8. 无限总体是指总体所包括的总体单位是不可数不可列或可数不可列的。不可数不可列是指总体所包括的总体单位数不清,自然无法编号。

9. 变量是标志的具体表现或指标的具体数值,如职工人数、年龄、工资、国内生产总值、销售总值、固定资产投资总额等。在数学中,定量变量就是变量,任意一个变量可以有一系列的取值,一般可用 X 、 Y 、 Z 来表示。

10. 确定性变量受确定性的因素影响,也就是说,影响变量变化的因素是明确的、可解释的。在这种因素的影响下,变量的变化幅度、变化方向是可以确定的。

11. 随机变量受随机因素的影响,随机因素的变化具有不确定性和偶然性,如气候的变化、海水潮汐的涨落等。

12. 离散型变量只能取整数,如人数、企业数、学校数等,计量单位一般为自然单位,如个、匹、头等。

13. 连续型变量可以取任意小数,如产值、销售额、固定资产投资额等,计量单位为价值单位;或可以取任意两个变量中的一个无限小数,大多为某种测量器具的测量结果,一般用度量衡单位表示,如某种零件的大小尺寸、直径、重量、体积、容积等。

14. 统计指标是指综合反映统计总体的某一方面数量特征的概念和数值。

15. 国内生产总值是指一个国家或一个地区在一定时期内所生产和提供的最终使用的产品和服务的总价值。

16. 数量指标反映了现象总体的总规模、总水平或工作总量,一般表现为绝对数,如人口总数、国内生产总值、工资总额等。

17. 质量指标反映的是现象总体的相对水平或工作质量,一般表现为相对数或平均数。

18. 基本统计指标体系是反映国民经济和社会发展基本情况的指标体系,包括社会指标体系、经济指标体系和科技指标体系等。