

应用统计（第五版） 习题与参考解答



丁岚编著

对外经济贸易大学出版社

应用统计 (第五版)

习题与参考解答

丁 岚 编著

对外经济贸易大学出版社
中国·北京

图书在版编目 (CIP) 数据

应用统计 (第五版) 习题与参考解答 / 丁岚编著. —
北京: 对外经济贸易大学出版社, 2012

ISBN 978-7-5663-0352-3

I. ①应… II. ①丁… III. ①应用统计学 - 高等学校
- 教学参考资料 IV. ①C8

中国版本图书馆 CIP 数据核字 (2012) 第 143676 号

© 2012 年 对外经济贸易大学出版社出版发行

版权所有 翻印必究

应用统计 (第五版) 习题与参考解答

丁 岚 编著

责任编辑: 郭华良 汪 洋

对外经济贸易大学出版社
北京市朝阳区惠新东街 10 号 邮政编码: 100029
邮购电话: 010-64492338 发行部电话: 010-64492342
网址: <http://www.uibep.com> E-mail: uibep@126.com

唐山市润丰印务有限公司印装 新华书店北京发行所发行
成品尺寸: 185mm × 230mm 12.75 印张 248 千字
2012 年 7 月北京第 1 版 2012 年 7 月第 1 次印刷

ISBN 978-7-5663-0352-3

印数: 0 001 - 3 000 册 定价: 23.00 元

前 言

应用统计是经济与管理类以及其他社会学科专业必修的一门专业基础课程，是通用的方法论学科，在经济和社会生活中应用范围非常广泛。近年来，随着经济学、管理学、社会学等学科定量分析方法的普及和深入，统计方法已经成为正确解读、认识、分析经济现象和社会现象必不可少的工具。应用统计具有很强的操作性和实践性。要学好这门课程，必须通过大量的练习，才能理解统计分析中常用的概念、基本原理，才能熟练掌握和应用其基本分析方法。

为了配合本科层次应用统计教学的需要，帮助学生更好地掌握这门课程，提高学生经济定量分析的能力，我们编写了《应用统计（第五版）习题与参考解答》。本书是对外经济贸易大学出版社2010年出版的《应用统计（第五版）》配套教辅用书。本书共十一章，每章包括五部分内容：

- 一、学习目的和要求
- 二、主要概念和内容
- 三、难点和重点
- 四、练习题
- 五、参考答案

书后还附有六份模拟试题及参考答案，便于学生进行自测。

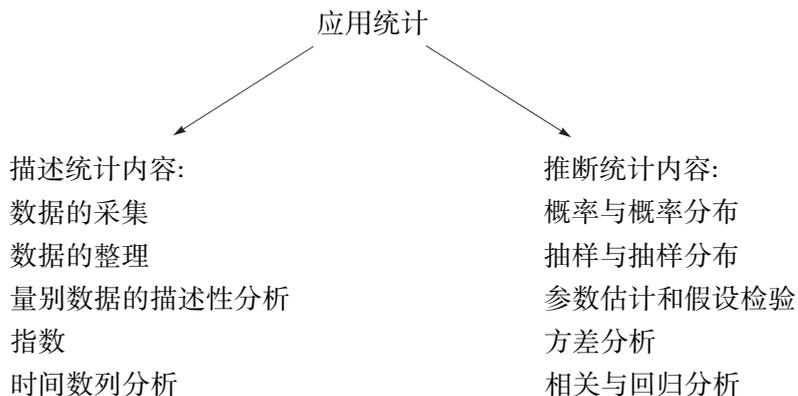
陈昊、师源、李宽、刘元、刘舒予、吴静涓等六位硕士生参加了本指导书的编写和上机计算等工作。本书在编写过程中，参考了统计界许多同仁的成果，选用了一些分析案例，引用了一些观点，在此一并表示衷心感谢。

应用统计课程主要介绍进行数据分析所必须使用的定量统计分析方法。这些方法分为描述统计方法和推断统计方法。描述统计方法是指对调查数据进行整理和计算统计量数等加工处理，用来描述总体特征的方法；推断统计方法是根据调查的样本数据去推断总体数量特征的方法。要学好本课程，首先应该从总体上把握进行统计分析的基本流程和应用统计课程的逻辑框架；其次，掌握每一流程中的各种具体分析方法，再根据数据的特征正确地选择统计分析方法。

统计分析的基本流程:

数据的采集→数据的整理→数据的分析

应用统计内容总框架图如下:



丁 岚

2012 年 3 月

目 录

绪论	(1)
第一章 数据的采集和整理	(5)
第二章 量别数据的描述性分析	(18)
第三章 概率与概率分布	(30)
第四章 抽样分布	(47)
第五章 参数估计	(56)
第六章 假设检验	(69)
第七章 卡方检验和方差分析	(86)
第八章 简单相关与回归分析	(97)
第九章 多元相关与回归	(115)
第十章 指数	(128)
第十一章 时间数列分析	(139)
模拟试题一	(147)
模拟试题一参考答案	(150)
模拟试题二	(152)
模拟试题二参考答案	(156)
模拟试题三	(158)
模拟试题三参考答案	(163)
模拟试题四	(165)
模拟试题四参考答案	(170)
模拟试题五	(173)
模拟试题五参考答案	(177)
模拟试题六	(179)
模拟试题六参考答案	(182)

绪 论

一、学习目的和要求

本章从整体上对应用统计的内容框架和有关基本概念进行了介绍,为以后各章的学习奠定基础。通过本章学习,应该理解统计学的含义和统计学的研究对象的特点,统计学研究的基本方法和基本内容,统计学与其他学科的关系以及统计数据的分类。

二、主要概念和内容

(一) 统计学发展的三个阶段

初创阶段:17 世纪中叶到 19 世纪中叶。这一阶段,统计学主要用来描述和比较各个国家的综合国力及人口研究,从研究对象上说是实质而不是方法论,从方法上看是经验研究多于理论研究,尚未进入系统的理论研究。1671—1676 年英国的威廉·配第著写的《政治算术》一书成为统计实务上升到理论和方法的标志。

近代统计学阶段:19 世纪中叶到 20 世纪 40 年代。1869 年凯特勒发表了《社会物理学》,揭开了近代统计学发展的序幕。这一阶段已经建立起系统的理论和方法,随后的许多数学家将数学方法在统计分析中大量应用,并且实现了由实质性科学向方法论科学的转变。

现代统计学阶段:20 世纪 50 年代至今。统计理论和方法的应用有了广泛的发展,进一步开发出一系列新的统计研究方法,并且将统计研究方法与电子计算机应用密切结合,开发了一系列功能较为齐全、数据处理能力较强的统计专用软件(如 SAS, SPSS, STATISTICA 等),使统计学成为一门与现代化计算工具密切相关的方法论学科。

(二) 统计学的定义

统计学是在质与量的辩证统一中,通过大量观察和平均方法,对客观总合现象的数量特征进行研究的方法论学科。

(三) 统计学研究的基本方法

统计学研究的基本方法是大量观察法和平均法。

(四) 统计学从内容上分类

统计学从内容上分为描述统计学和推断统计学。

(五) 统计学中常用的几个术语

1. 总体: 同质客观现象个体所组成的集团。
2. 总体单位: 组成总体的个体成员。
3. 样本: 从总体中按照某种抽样方法抽中的个体单位。
4. 标志: 说明总体单位属性的名称, 表现为数量上不等的称为数量标志; 表现为状态上差异的称为品质标志。

(六) 数据的类型

1. 质别数据和量别数据
2. 截面数据和时间序列数据
3. 初级数据和次级数据

三、难点和重点

本章是对全书的总括, 是学好应用统计课程的基础, 因此, 对本章的基本概念、常用术语, 如总体、总体单位、样本、标志等应该理解并记忆, 以达到熟练应用的程度。

四、练习题

(一) 判断题

1. 统计学是研究社会领域和自然领域的规律的实质性科学。
2. 大量观察法是统计研究的最基本的方法。
3. 总体单位的属性既可以用数值表示, 也可以用文字表示。
4. 对任一事物进行统计研究的前提是构成统计总体的各个个体单位必须具有差异性。
5. 构成总体的各个个体单位至少在某一属性上相同, 才能准确地确定总体范围。
6. 统计学与哲学相同, 属于认识客观世界的最一般的方法论科学。
7. 运用统计方法可以研究和解释事物发展的原因和规律。
8. 总体的同质性、变异性、大量性是平均法的应用前提。
9. 在用计算机数据进行汇总时, 往往将男性用“1”来表示, 女性用“0”来表示, 所以可将性别视为量别变量。
10. 在人工干预或操作情况下收集的数据称为实验数据。

(二) 单项选择题

1. 调查城镇居民对政府服务的满意度,需要抽取足够多的居民,这种方法称为()。
A. 随机抽样法 B. 统计分组法 C. 回归分析法 D. 大量观察法
2. 用部分数据推断总体数据的方法,属于()。
A. 理论统计学的内容 B. 应用统计学的内容
C. 推断统计学的内容 D. 描述统计学的内容
3. 构成一个统计总体,必须具有若干特征,下列不属于统计总体特征的是()。
A. 数量性 B. 同质性 C. 变异性 D. 大量性

(三) 多项选择题

1. 应用统计学的内容包括()。
A. 决策的方法 B. 数据的描述方法
C. 数据的推断方法 D. 数理分析方法
E. 数据的整理方法
2. 统计数据必须具有的特点是()。
A. 数量性 B. 客观性 C. 具体性 D. 通用性
E. 总和性
3. 描述统计学的内容包括()。
A. 收集数据的方法 B. 整理数据的方法
C. 用集中趋势量数描述数据 D. 用离中趋势量数描述数据
E. 用样本数据对总体数据进行推断
4. 推断统计学的内容包括()。
A. 收集数据的方法 B. 怎样从总体中抽取样本
C. 用集中趋势量数描述数据 D. 怎样控制样本对总体的代表性误差
E. 用样本数据对总体数据进行推断
5. 统计数据的特点是()。
A. 大量性 B. 客观性 C. 具体性 D. 总合性
E. 数量性
6. 统计数据按来源分类,可以分为()。
A. 质别数据 B. 次级数据 C. 量别数据 D. 初级数据
E. 截面数据

五、参考答案

(一) 判断题

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

(二) 单项选择题

1. D 2. C 3. A

(三) 多项选择题

1. BCE 2. ABCE 3. ABCD 4. BDE 5. BCDE 6. BD

第一章 数据的采集和整理

一、学习目的和要求

本章内容包括统计工作第一阶段和第二阶段:数据的采集阶段和整理阶段。

掌握数据采集的主要方法,如普查、抽样调查及统计报表等方法,了解各种不同方法的特点及适用环境;了解抽样误差的种类;掌握统计调查方案的基本内容;能够根据研究主题设计调查问卷,能够熟练将调查数据录入统计分析软件(SPSS) 中。

掌握统计数据整理的基本方法;了解统计数列的种类;掌握质别数据和量别数据的统计分组方法;掌握频数分布表的编制方法;了解统计表的结构;掌握统计图的绘制方法。

熟练运用统计分析软件(SPSS) 对统计数据进行整理和绘图。

二、数据采集主要概念和内容

(一) 统计数据的来源

统计数据的来源包括采集(统计调查、试验) 和业务记录。

(二) 统计调查的方式

统计调查的方式包括普查、抽样调查和统计报表。

1. 普查

普查是为了某一特定目的而专门组织的一次性全面调查。其特点是规定调查数据所属的标准时间,以避免数据的重复和遗漏。

2. 抽样调查

抽样调查属于非全面调查,是世界各国广泛使用的一种调查方式。即从研究总体中抽取部分单位进行调查。在统计工作中,通常需要研究的总体单位都非常多,大多数情况下受到时间和资金的限制必须通过抽样的方法了解总体。因此,抽样调查是统计研究最常用的一种获得数据的方式。常用的抽样调查方法有随机抽样和非随机抽样。随机抽样包括简单随机抽样、系统抽样、分层抽样、整群抽样、多阶段抽样;非随机抽样有便利抽

样(又称任意抽样、偶遇抽样)、判断抽样、配额抽样、滚雪球抽样。

3. 统计报表

统计报表是按照国家有关法规的规定,自上而下布置,自下而上逐级提供基本统计数据的一种调查方式。它以一定的原始记录为基础,按照统一的表式、统一的指标、统一的报送时间和报送程序进行填报,带有一次性布置、周期性填报的特点。

(三) 数据的误差

数据的误差包括抽样误差和非抽样误差两类。

抽样误差:指由于抽样的随机性所带来的误差,不能避免但可以计算控制。

非抽样误差:指除抽样误差之外的由于其他原因造成的样本观察结果与总体真值之间的差异,如问卷设计的偏差,访员工作责任心、业务水平,受访者提供的数据不真实等等。这类误差是无法测量的,但可以通过采取积极的措施减少它们的出现。

(四) 采集数据的主要工具

采集数据的主要工具是调查问卷或调查表。设计一份科学、合理的问卷是获得真实可靠数据的必要条件。

(五) 实验设计

实验设计是指通过对研究对象进行比较测试获得数据的方法。

(六) 统计调查方案的内容

1. 确定调查目的
2. 确定调查对象和调查单位
3. 确定调查项目和调查表(调查问卷)
4. 确定调查时间、地点和调查方式
5. 制定调查的组织实施计划

(七) 统计数据的审核

1. 审核的目的:保证数据的完整性和准确性
2. 审核的方法:逻辑审核和计算审核

三、数据整理的主要概念和内容

(一) 统计数据整理的概念

统计数据整理,即将收集来的数据编制成频数分布表的过程。

(二) 统计数据整理步骤

统计数据整理步骤分为统计数据分组、统计数据汇总、统计数据表述。

(三) 质别数据的整理方法

质别数据的整理分为三个步骤: 分组(类) → 归类 → 汇总。

(四) 量别数据整理方法

根据变量的特点分为单项式分组和组距式分组两种。单项式分组适用于离散型变量且变量值个数较少的情况; 组距式分组适用于连续型变量且变量值个数较多的情况。

(五) 组距式分组的步骤

确定组数 → 计算全距、求出组距 → 界定组限 → 分组归类 → 汇总制成频数分布表。

(六) 统计数列

统计数据整理后制成的频数分布表中的数据就是统计数列。统计数列分为两类: 截面数列和时间数列。截面数列按分组标志不同又分为品质数列和变量数列。

(七) 统计数据的表述

统计数据的表述即用统计表和统计图对已经整理好的数据进行概括性描述。

(八) 统计表

统计表是把统计调查所得的数据, 经过整理后概括地表述统计数字的一种表格。

从统计表的形式看, 统计表由总标题、横行标题、纵栏标题和指标数值四部分组成。

从统计表的内容看, 统计表可分为主词栏和宾词栏两部分。

(九) 统计图

统计图是将整理好的统计数据用图型来概括描述的方法。定性的截面数据适合用条形图、饼图; 定量的截面数据适合用直方图和频数多边形。如果希望显示各组的频数分布态势, 又希望能保留各组内的原始数据, 可以使用茎叶图。时间数列数据用线形图描述。

(十) 量别数据分布的曲线种类

量别数据分布的曲线种类包括中央集中分布、U 型分布和 J 型分布。

四、难点和重点

数据的不同来源途经; 不同抽样调查方法的特点及适用环境。

SPSS 的使用: 定性变量和定量变量在 SPSS 的输入方式, 定性变量和定量变量在 SPSS 的命名方式、赋值和编码。

质别数据、量别数据的整理方法; 统计表的内容和结构; 根据统计数据类型正确选择统计图形。

五、练习题

(一) 判断题

1. 统计的初级数据和次级数据均来源于统计调查。
2. 进行普查时涉及的调查项目越多越好。
3. 普查和抽样调查是世界各国通用的收集统计数据的主要方式。
4. 调查项目是设计统计调查方案时应考虑的首要因素。
5. 普查和抽样调查的目的不同,前者为了了解总体,后者为了了解局部。
6. 统计数据的采集和整理是进行统计分析的前提。
7. 调查问卷是由问题和备选答案组成的文件。
8. 进行普查时需要规定标准时间,其目的是为了按时完成调查工作。
9. 所有的抽样方式都有一定的概率作为推断总体的保证。
10. 对无限总体的观察,必须采用抽样调查的方式获得数据。
11. 统计采集数据常用的调查工具是调查问卷和调查表。
12. 截面数列按分组标志不同可以分为品质数列和变量数列。
13. 同龄人的身高,用频数曲线描述一般表现为钟型分布。
14. 欲掌握我国人口是否具有老龄化特征,对人口按年龄分组,适合采用等距式分组。
15. 对在校大学生的年龄进行分组适合采用组距式分组。
16. 向上累计频数指小于上限的的频数共有多少。
17. 向下累计频数指大于下限的的频数共有多少。
18. 在组距式分组中,采用组中值作为每组一般水平的代表值。
19. 直方图仅适合用于显示连续型分组数据。
20. 茎叶图不仅能够反映数据的分布情况,还能显示数据的原始信息。
21. 频数分布表中的组距在茎叶图中称为步长。

(二) 单项选择题

1. 下列不属于原始数据来源的是()。
A. 行政记录 B. 统计调查 C. 统计年鉴 D. 实验
2. 在进行普查时,必须要规定标准时间,其目的是为了()。
A. 避免登记的重复和遗漏 B. 确定调查对象的范围
C. 确定调查单位 D. 保证调查的及时性
3. 某出口企业对其产品质量进行检验,采用的调查方式是()。
A. 普查 B. 随机抽样调查 C. 判断抽样调查 D. 方便抽样调查

4. 了解居民消费水平的变化状况,一般采用入户调查。确定样本户的方式是()。
- A. 系统抽样 B. 分层抽样 C. 整群抽样 D. 多步抽样
5. 抽样调查与普查的根本区别是()。
- A. 作用不同 B. 灵活程度不同
C. 选取观察值的方法不同 D. 精确度不同
6. 在抽样调查中,()。
- A. 不会出现登记性误差,只会出现代表性误差
B. 会出现登记性误差,不会出现代表性误差
C. 既会出现登记性误差,也会会出现代表性误差
D. 既不会出现登记性误差,也不会出现代表性误差
7. 统计调查表中的调查项目就是()。
- A. 统计指标 B. 统计数值 C. 统计分组 D. 统计标志
8. 街头拦截获得样本数据的方式属于()。
- A. 普查 B. 方便抽样调查 C. 判断抽样调查 D. 随机抽样调查
9. 普查中的标准时间是指()。
- A. 进行调查工作的时间 B. 调查数据所属的时间
C. 记录数据的时间 D. 数据上报的时间
10. 某些商务活动是通过实验设计获得原始数据的,在进行实验设计时应注意的原则不包括()。
- A. 对照原则 B. 齐同原则 C. 合理原则 D. 随机原则
11. 某次调查收集了 300 名学生的性别、专业和月手机支出额的数据,这项调查的变量个数有()。
- A. 1 个 B. 2 个 C. 3 个 D. 300 个
12. 构成截面数据的两个要素是()。
- A. 组距和各组频数 B. 组限和各组频数
C. 组中值和各组频数 D. 分组标志和各组频数
13. 在全距既定的情况下()。
- A. 组数越多则组距越大 B. 组数越多则组距越小
C. 组数越少则组距越小 D. 组数与组距无对应关系
14. 适合编制单值数列的数据是()。
- A. 连续型变量且数据变动范围较小 B. 离散型变量且数据变动范围较小

- C. 连续型变量且数据变动范围较大 D. 离散型变量且数据变动范围较大
15. 在编制等距式分组数列时如果全距 57, 组数为 7, 为了计算方便, 组距应当取()。
- A. 8.1 B. 8 C. 7 D. 10
16. 在平面直角坐标轴中, 横轴刻度表示直方图的()。
- A. 频数 B. 组中值 C. 组距 D. 组限
17. 频数密度是指()。
- A. 组距除以频数 B. 平均每组的频数
C. 平均每组组内的频数 D. 单位组距的频数
18. 对数据进行分组时, 采用等距分组还是异组距分组, 取决于()。
- A. 变量值的多少 B. 次数的多少
C. 研究现象的特点 D. 组数的多少
19. 对工业企业按经济类型分组和按利润额分组()。
- A. 都属于质别标志分组
B. 都属于量别标志分组
C. 前者属于质别标志分组, 后者属于量别标志分组
D. 前者属于量别标志分组, 后者属于质别标志分组
20. 根据异距分组绘制直方图时在坐标纵轴刻度应当使用()。
- A. 各组的频率密度 B. 各组的频率
C. 各组的实际频数 D. 各组的组中值

(三) 多项选择题

1. 我国第五次人口普查的标准时间是 2000 年 11 月 1 日 0 时, 下列情况应登记的人口有()。
- A. 2000 年 11 月 2 日 1 时出生的婴儿 B. 2000 年 10 月 30 日 6 时出生的婴儿
C. 2000 年 10 月 30 日 14 时死亡的人 D. 2000 年 11 月 1 日 1 时死亡的人
E. 2000 年 10 月 29 日出生, 11 月 7 日 3 时死亡的婴儿
2. 在抽样调查中, 产生非抽样误差的原因有()。
- A. 错误地定义调查总体 B. 问卷设计不合理
C. 抽样方法不正确 D. 被调查人员的配合程度
E. 调查人员的工作态度和业务水平
3. 调查过程中产生登记误差的原因是()。
- A. 调查者计录数据有误 B. 调查者计算数据有误

- C. 调查者汇总数据有误
D. 被调查者提供数据有误
E. 调查样本分布不合理
4. 对采集来的初级数据在整理之前应该进行()。
- A. 相关性审核
B. 完整性审核
C. 时效性审核
D. 适用性审核
E. 准确性审核
5. 下列商务活动中,可以通过实验方法来测定其效果的有()。
- A. 员工培训方法的选择
B. 商品包装
C. 商品促销方式
D. 商品陈列布局
E. 商品分类
6. 统计调查方案包括的内容有()。
- A. 确定调查目的
B. 确定调查对象
C. 确定调查项目和调查表
D. 确定调查时间和地点
E. 确定调查的组织计划
7. 对于次级数据,应该进行以下方面的审核()。
- A. 完整性
B. 准确性
C. 时效性
D. 适用性
E. 合理性
8. 统计分组应该遵循的基本原则是()。
- A. 穷尽
B. 适用
C. 互斥
D. 合理
E. 随机
9. 关于离散型变量,下列表述正确的是()。
- A. 可以采用单值分组
B. 可以采用组距分组
C. 相邻组的上下限可以间断
D. 相邻组的上下限可以重叠
E. 可按“上限不在组内”原则处理上下限相同的变量
10. 茎叶图和箱线图不适用的数据类型是()。
- A. 已分组的分类数据和定序数据
B. 未分组的分类数据和定序数据
C. 已分组的数值型数据
D. 未分组的数值型数据
E. 频数分布表中的数据
11. 编制频数分布表的目的是()。
- A. 显示数据的差异
B. 显示数据的分布状况
C. 显示数据的中心
D. 显示数据的相互关系