

应用统计

全国硕士研究生入学考试应用统计教辅系列 <

应用统计考研 通关宝典

> 主编 索瑜

知识罗列

例题讲解

真题演练

决战考研

应用统计(同等学力)

应用统计(专业学位)

应用统计(学术型)

应用统计



北京理工大学出版社
BEIJING UNIVERSITY OF TECHNOLOGY PRESS

版权专有 侵权必究

图书在版编目(CIP)数据

应用统计考研通关宝典 / 索瑜主编. —北京: 北京理工大学出版社, 2021. 4

ISBN 978-7-5682-9746-2

I. ①应… II. ①索… III. ①应用统计学-研究生-入学考试-自学参考资料 IV. ①C8

中国版本图书馆 CIP 数据核字(2021)第 068041 号

出版发行 / 北京理工大学出版社有限责任公司

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010)68914775(总编室)

(010)82562903(教材售后服务热线)

(010)68948351(其他图书服务热线)

网 址 / <http://www.bitpress.com.cn>

经 销 / 全国各地新华书店

印 刷 / 天津市新科印刷有限公司

开 本 / 787 毫米×1092 毫米 1/16

印 张 / 18

字 数 / 449 千字

版 次 / 2021 年 4 月第 1 版 2021 年 4 月第 1 次印刷

定 价 / 72.80 元

责任编辑 / 孟祥雪

文案编辑 / 孟祥雪

责任校对 / 周瑞红

责任印制 / 李志强

图书出现印装质量问题,请拨打售后服务热线,本社负责调换

前言

我们的一生中有太多考试,考研只是其中的一个。然而,无论大家最终能否顺利跨过研究生这座大门,备考过程中都会感到迷茫与煎熬。作为一名奋斗在考研一线的教学教师,因为理解考生,所以能够从大家的角度出发,因为亲身经历过,所以更希望大家能够顺利通关,而本系列丛书的真正意义也在于此。

选择考研,意味着踏上了一段新的征程,这个过程中大家可能会感到迷茫、煎熬、困惑与失落,但我们一定要坚持、勤奋、勇敢,不畏艰辛地战胜困难。我们会陪大家一同前行,不负韶华,静待花开!

本书以贾俊平老师所著《统计学》、茆诗松老师所著《概率论与数理统计教程》为基础,参照《全国硕士研究生入学统一考试应用统计硕士专业学位统计学考试大纲》(简称《432 考纲》)进行编写,考点突出,更加有助于考生备考。

本人通过多年来的教学和对学员的了解,在编写过程中充分运用以往经验,并致力于帮助自我复习较慢的学生、跨专业的学生、起步较晚的学生、对知识掌握不牢固以及很难对知识进行归类的学生迅速提高分数,帮助其形成统计学思维,掌握答题技巧并能迅速找到考查重点。

本书有以下三大特点。

(1)全面。本书涵盖《432 考纲》中规定的全部考点,做到了复习的全面性。

(2)讲练结合。本书引入了经典真题及常考题型,并对真题进行了详细的讲解,可以达到学与用相结合,复习过程中非常有助于考生对知识点的理解。

(3)模块化。本书中所有章节按照考点进行编写,并详细地注明了每个考点在考试过程中的考频、考试难度以及考试题型,可以更进一步帮助考生明确考点的重要程度,厘清复习思路。

本书既可以作为考研“432 统计学”备考用书,也可以作为统计学专业的辅导用书。对于备考“432 统计学”的同学,可以在阅读完参考教材之后阅读本书,也可以将本书与参考教材同步结合使用。另外,在阅读本书的同时,结合《应用统计考研名校真题汇编》进行巩固练习,更有助于对知识点的理解与掌握。

由于作者水平有限,书中难免存在错误和疏漏之处,敬请读者谅解并提出宝贵意见,以便进一步修订和改进。

索 瑜

目 录

第一篇 描述性统计	1
第一章 导论	1
考点一 统计学的基本概念	2
考点二 统计数据类型	3
考点三 统计中的几个基本概念	5
答案与解析	8
第二章 数据的搜集	10
考点一 数据的来源	11
考点二 调查数据	12
考点三 数据搜集方法	17
考点四 实验数据	19
考点五 误差分析与控制	20
答案与解析	22
第三章 数据的图表展示	25
考点一 数据的预处理	26
考点二 品质数据的整理与展示	27
考点三 数值型数据的整理与展示	32
考点四 合理使用图表	40
答案与解析	41
第四章 数据的概括性度量	44
考点一 集中趋势的度量	45
考点二 离散程度的度量	54
考点三 分布形态的度量	59
答案与解析	61
第二篇 概率基础	63
第五章 概率与概率分布	63
考点一 随机事件及其概率	64
考点二 概率的运算法则	69
考点三 随机变量及其分布	74
考点四 随机变量的数学期望与方差	78

考点五 常用离散分布与连续分布	83
考点六 随机变量函数的分布	90
答案与解析	92
第三篇 推断性统计	97
第六章 统计量及其抽样分布	97
考点一 统计量	98
考点二 关于分布的几个概念	101
考点三 由正态分布导出的几个重要分布	103
考点四 样本均值的分布与中心极限定理	108
考点五 其他抽样分布	110
答案与解析	113
第七章 参数估计	116
考点一 参数估计的基本原理	117
考点二 点估计	119
考点三 一个总体参数的区间估计	125
考点四 两个总体参数的区间估计	131
考点五 样本量的确定	137
答案与解析	140
第八章 假设检验	144
考点一 假设检验的基本问题	145
考点二 一个总体参数的检验	152
考点三 两个总体参数的检验	158
考点四 P 值决策问题	165
答案与解析	168
第九章 分类数据分析	170
考点一 χ^2 统计量与拟合优度检验	171
考点二 列联分析:独立性检验	173
考点三 列联表中的相关测量	176
答案与解析	179
第十章 方差分析	181
考点一 方差分析引论	182
考点二 单因素方差分析	186
考点三 双因素方差分析	194
答案与解析	202

第四篇 计量分析	205
第十一章 一元线性回归	205
考点一 变量间关系的度量	206
考点二 一元线性回归	211
考点三 利用回归方程进行预测	220
考点四 残差分析	222
答案与解析	223
第十二章 多元线性回归	226
考点一 多元线性回归模型	227
考点二 回归方程的检验	229
考点三 多重共线性	237
考点四 变量选择与逐步回归	239
答案与解析	240
第五篇 时间序列基础	243
第十三章 时间序列分析和预测	243
考点一 时间序列及其分解	244
考点二 时间序列的预测	251
答案与解析	258
第六篇 指数基础	261
第十四章 指数	261
考点一 指数基本问题	262
考点二 总指数编制方法	263
考点三 指数体系	268
考点四 几种典型的指数	272
答案与解析	275
参考文献	279

第一篇 描述性统计

第一章 导论

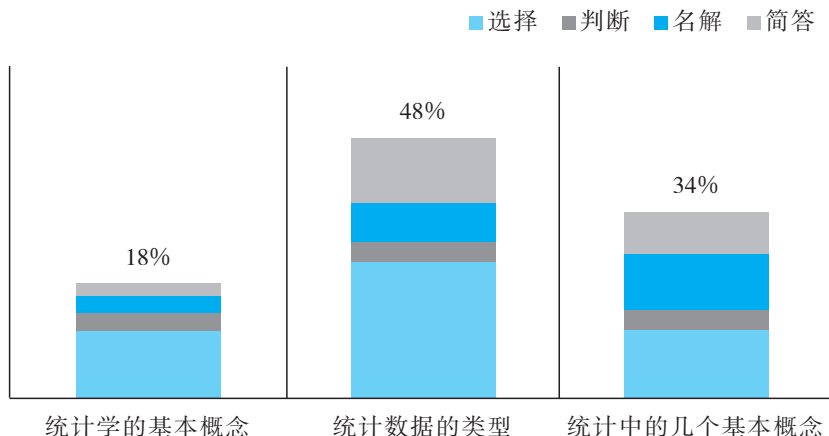


扫码回复“导论”
听讲解

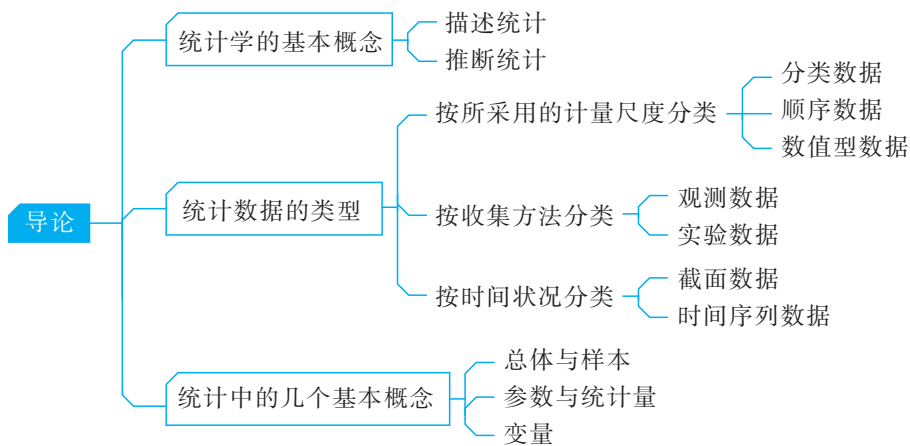
本章综述

本章主要是对统计学知识体系的介绍以及基础知识的讲解,根据历年考情提要,要求考生具备如下能力:理解统计学的含义;理解描述统计和推断统计;了解统计学的应用领域;理解并掌握统计数据类型;理解并掌握统计中的几个基本概念。

考频统计



知识框架



考点一 统计学的基本概念

考情提要

难度分析	☆
考频分析	☆☆
高频考点分析	(1)理解统计学、描述统计、推断统计的概念。 (2)描述统计与推断统计的区分。 (3)概念的简述
题型分析	选择:2012、2014 央财;2015、2016 重庆大学;2014 山东大学;2017 北工大;2019 北京林业大学;2014、2015、2017、2018 安徽大学。 判断:2014 重庆大学;2017 央财。 名解:2020 东北财经;2019 山东财经。 简答:2019 武汉理工

考点精讲

在我们的工作或管理中,总会面对各种各样的数据。我们需要分析这些数据,从中得出某些结论以帮助我们做出决策。统计就是用来处理数据的,它是关于数据的一门学问。统计学提供的就是一套有关数据收集、数据处理、数据分析的方法。

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。数据分析所用的方法大体上可以分为**描述统计**与**推断统计**两大类。

描述统计是研究数据收集、处理、汇总、图表描述、概括与分析等统计方法。其内容包括如何取得研究所需要的数据,如何用图表形式对数据进行处理和展示,如何通过对数据的汇总、概括与分析,得出所关心的数据特征。

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。其内容包括**参数估计**和**假设检验**两大类。参数估计是利用样本统计量推断所关心的总体参数;假设检验则是利用样本信息判断对总体的某个假设是否成立。

真题精练

1. 统计学是关于数据的科学,其内容包括()。
- A. 数据收集和整理
B. 数据整理和分析
C. 数据分析和解释
D. 数据收集、整理、分析和解释

【2016 重庆大学】

2. 研究如何利用样本数据来推断总体特征的统计学方法是()。

- A. 描述统计
B. 理论统计
C. 推断统计
D. 应用统计

【2012 中央财经大学、2014 山东大学】

考点二 统计数据类型

考情提要

难度分析	☆☆
考频分析	☆☆☆
高频考点分析	(1)不同数据类型的理解与简述。 (2)根据描述区分不同数据类型
题型分析	选择: 2011、2012、2014、2018、2019 央财;2016 重庆大学;2015、2017 山东大学;2012、2013 首经贸;2017 浙工商;2015 华中农业;2019 中南财;2019、2020 北工大;2019 安徽师范;2019 湖南师范;2019 上海师范大学。 判断:2017 央财;2017 重庆大学。 名解:2017、2019 武汉理工;2015、2016、2018 首经贸。 简答:2017 央财;2016 重庆大学;2013、2016 浙工商;2019 河北大学;2017、2019 东北财经;2019 云南大学;2017 安徽大学;2020 南审

考点精讲

统计数据是对现象进行测量的结果。按照不同的分类方法,统计数据可以分成不同的类型。

1. 按所采用的计量尺度分类

分类数据是只能归于某一类别的非数字型数据,它是对事物进行分类的结果,数据表现为类别,是用文字来表述的。如,人口按照性别分为男、女两类。

顺序数据是只能归于某一有序类别的非数字型数据。如,考试成绩可以分为优、良、中、及格、不及格等。

数值型数据是按数字尺度测量的观察值,其结果表现为具体的数值。

分类数据和顺序数据说明的是事物的品质特征,通常是用文字来表述的,其结果均为类别,因而也被统称为**定性数据**或**品质数据**。数值型数据说明的是现象的数量特征,通常是用数值来表现的,因此也可称为**定量数据**或**数量数据**。

知识点补充

根据研究对象计量的不同精确程度,计量尺度由低到高,由粗略到精确可以分为四个层次:

(1)定类尺度:按照客观现象的某种属性分类。各类别无好坏、优劣、大小之分。如,金融业、旅游业等。

(2)定序尺度:对客观现象各类之间等级差或顺序差的一种度量。采用定序尺度可以将研究对象分为各种类别,同时不同类别之间可以进行好坏、优劣、大小比较。如,产品分为一等品、二等品、三等品、次品等。

(3)定距尺度:度量的是现象类别与次序之间的间距。不仅可以用数值表示各类别的不同和顺序的大小,而且可以反映各现象之间数量上的差异。但定距尺度没有绝对的零点,如温度。定距尺度主要进行加减运算。

(4)定比尺度:用确切的数字表示现象之间的差别。但定比尺度有绝对的零点,如GDP、价格。定比尺度不仅可以进行加减运算,还可以进行乘除运算。

2. 按收集方法分类

观测数据是通过调查或观测而收集到的数据,这类数据是在没有对事物人为控制的条件下得到的,有关社会经济现象的统计数据几乎都是观测数据。

实验数据是在实验中控制实验对象而收集到的数据。自然科学领域的大多数数据都为实验数据。

3. 按时间状况分类

截面数据是在相同或近似相同的时间点上收集的数据,这类数据通常是在不同的空间获得的,用于描述现象在某一时刻的变化情况。

时间序列数据是在不同时间收集到的数据,这类数据是按时间顺序收集到的,用于描述现象随时间变化的情况。

真题精练

1. 按照被描述的对象与时间的关系,统计数据可分为()。

- A. 顺序数据和数值型数据
- B. 观测数据和实验数据
- C. 截面数据和时间序列数据
- D. 低维数据和高维数据

【2016 重庆大学】

2. 一项民意调查的目的是确定年轻人愿意与其父母讨论的话题。调查结果表明,45%的年

年轻人愿意与其父母讨论家庭财务状况,38%的年轻人愿意与其父母讨论有关教育的话题,15%的年轻人愿意与其父母讨论爱情问题。该调查所搜集的数据是()。

- A. 分类数据 B. 顺序数据 C. 数值型数据 D. 实验数据

【2015 山东大学】

3. 下列哪一种情形涉及定性数据的收集? ()

- A. 质量控制工程师测量灯泡的寿命
B. 社会学家通过抽样调查来估计广州市市民的平均年收入
C. 运动器材厂家在区分各大俱乐部棒球选手是左撇子还是右撇子的时候做的调查
D. 婚礼策划公司通过抽样调查来估计上海市市民举办婚礼的平均开销

【2017 山东大学】

4. 在一次问卷调查中要求被调查者直接填写出个人的民族、婚姻状况、居住地的邮政编码、年龄和收入,以下说法不正确的是()。

- A. 民族是定性变量
B. 邮政编码是定量变量
C. 年龄的计量尺度是定比尺度
D. 收入数据是定量数据

【2011 中央财经大学】

考点三 统计中的几个基本概念

🕒 考情提要

难度分析	☆☆
考频分析	☆☆☆
高频考点分析	(1)理解总体、样本、参数、统计量、变量等概念。 (2)理解各个概念之间的具体关系。 (3)正确识别题中的样本与总体
题型分析	选择:2011 央财;2013、2015—2017 浙工商;2019 云南大学;2019 安徽师范;2019 浙江财经;2019 上海海事大学。 判断:2019 央财;2014 重庆大学。 名解:2016 外经贸;2015 武汉理工;2017 首经贸;2018、2019 河北大学;2011、2014、2015 东北财经。 简答:2011 重庆大学;2015、2020 浙工商;2018 东北财经;2017 北工大;2015 暨南大学

考点精讲

1. 总体与样本

总体是包含所研究的全部个体(数据)的集合。总体通常由所研究的一些个体组成,组成总体的每个元素称为**个体**。比如,要检验一批灯泡的使用寿命,这批灯泡构成的集合就是总体,每个灯泡就是一个个体。

根据总体包含的单位数目是否可数分为**有限总体**与**无限总体**。有限总体是指总体的范围能够确定,而且元素的数目是有限可数的;无限总体是指总体所包括的元素是无限的、不可数的,其区分目的是判别在抽样中每次抽取是否独立。

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**。抽样的目的是根据样本提供的信息推断总体的特征,根据样本统计量去估计总体参数。比如,从一批灯泡中随机抽取 100 个,这 100 个灯泡就构成了一个样本,然后根据这 100 个灯泡的平均使用寿命去推断这批灯泡的平均使用寿命。

2. 参数与统计量

参数是用来描述总体特征的概括性数字度量,是研究者想要了解的总体的某种特征值。如,总体平均数 μ 、总体标准差 σ 、总体比例 π 等。

参数是确定的,但是未知的。

统计量是用来描述样本特征的概括性数字度量,是根据样本数据计算出来的一个量,是样本的函数。如,样本平均数 \bar{x} 、样本标准差 s 、样本比例 p 等。

统计量是随机的,但是已知的。

总体、样本、参数及统计量之间的关系如图 1-1 所示。

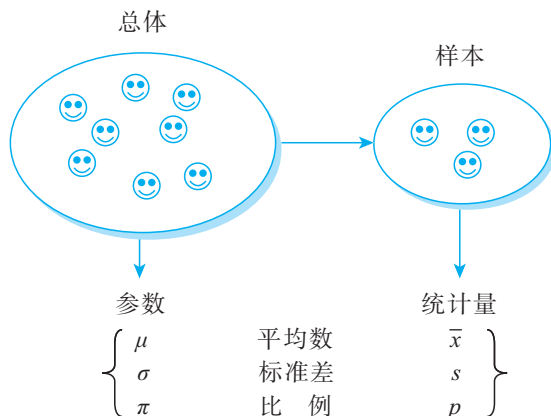


图 1-1 总体、样本、参数及统计量之间的关系

知识点补充

(1) 标志与指标。

标志是反映总体单位的属性和特征,有能够用数量表示的数量标志,还有不能用数量表示的品质标志。

指标是反映总体的数量特征,所有指标都能用数量来表示。

联系:标志与指标的关系如同总体单位和总体的关系一样,是个别和整体的关系。标志是总体指标的来源和基础,指标是标志的综合。没有标志就没有指标,在一定条件下数量标志和指标存在着变换关系。

(2) 时点数据与时期数据。

时点数据是反映社会经济现象在某一时刻或某一时点上的状况的总量数据。不具有可加性,数值的大小与登记时间的间隔长短无关,数值是间断计数的。如,某地区人口数,某商业企业商品库存额等。

时期数据是反映社会经济现象在一段时期内发展变化状况的数据。具有可加性,数值的大小与所属时间长短有直接相关关系(一般为正相关),数值是连续计数的。如,投资总额,商品销售额等。

3. 变量

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。变量的具体取值称为**变量值**。变量分类如表 1-1 所示。

表 1-1 变量分类

变量类型	代表	取值
分类变量	事物类别的一个名称	分类数据
顺序变量	事物有序类别的一个名称	顺序数据
数值型变量	事物数字特征的一个名称	数值型数据

数值型变量根据其取值不同可分为**离散型变量**和**连续型变量**。

其他情况分类:**随机变量**与**非随机变量**、**经验变量**与**理论变量**。

真题精练

- 对某地区工业企业职工情况进行研究,统计总体是()。
 - 每个工业企业
 - 该地区全部工业企业
 - 每个工业企业的全部职工

D. 该地区全部工业企业的全部职工

【2013、2015 浙江工商大学】

2. 某城市进行私营企业消防设备安全检查,个体是()。

- A. 每一个私营企业
- B. 每个私营企业的消防设备
- C. 私营企业的每一个消防设备
- D. 私营企业的全部消防设备

【2017 浙江工商大学】

答案与解析

考点一 统计学的基本概念

1. 【答案】D

【解析】统计学是一门研究数据的科学,目的是如何有效地收集、整理和分析这些数据,探索数据内在的数量规律性,对所观察的现象做出推断或预测,直到为采取决策提供依据。

2. 【答案】C

【解析】统计学主要分为描述统计和推断统计。描述统计是研究数据收集、处理、汇总、图表描述、概括与分析等统计方法;推断统计则是研究如何利用样本数据来推断总体特征的统计方法。

考点二 统计数据类型

1. 【答案】C

【解析】按照被描述的对象与时间的关系,可以将统计数据分为截面数据和时间序列数据。在相同或近似相同的时间点上收集到的数据称为截面数据。在不同时间收集到的数据,称为时间序列数据。

2. 【答案】A

【解析】分类数据是指只能归于某一类别的非数字型数据,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,题中按照沟通的话题将类别分开,属于分类数据。

3. 【答案】C

【解析】定性数据是描述一个事物特性的数据,不需要具体数值,通常用文字表示;定量数据通常从数值上来描述现象的数量特征。A, B, D 三项都需要用具体数值来描述,属于定量数据的收集。

4. 【答案】B

【解析】定量变量是指可以用数值表示其观察结果,而且这些数值具有明确的数值含义,不仅能分类而且能测量出来具体大小和差异的变量。而邮政编码只是用来进行分类的,无数值大小差异,因此为定类数据。

考点三 统计中的几个基本概念

1.【答案】D

【解析】统计总体是包含所研究的全部个体的集合,对某地区工业企业职工情况进行研究,其统计总体是该地区全部工业企业的全部职工。

2.【答案】C

【解析】总体是统计所需要研究的客观事物的全体,个体是组成总体的各个单位,是某种事物。某城市进行私营企业消防设备安全检查的对象是消防设备,因此个体为私营企业的每个消防设备。

第二章 数据的搜集

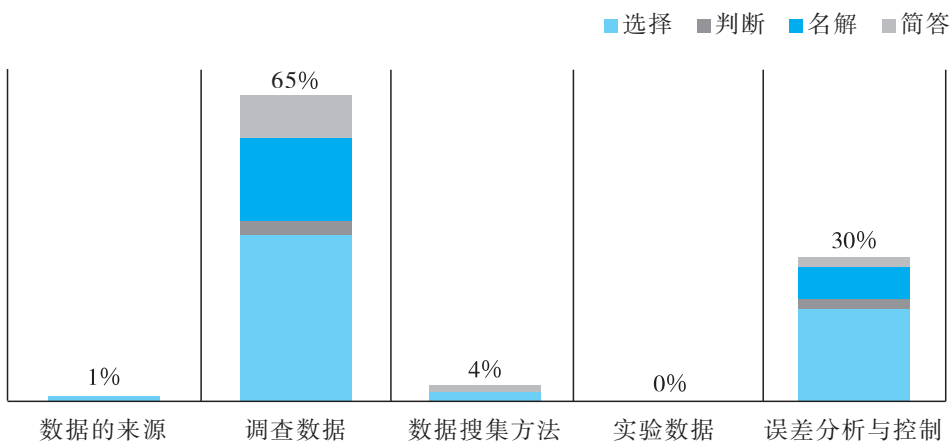


扫码回复“数据搜集”听讲解

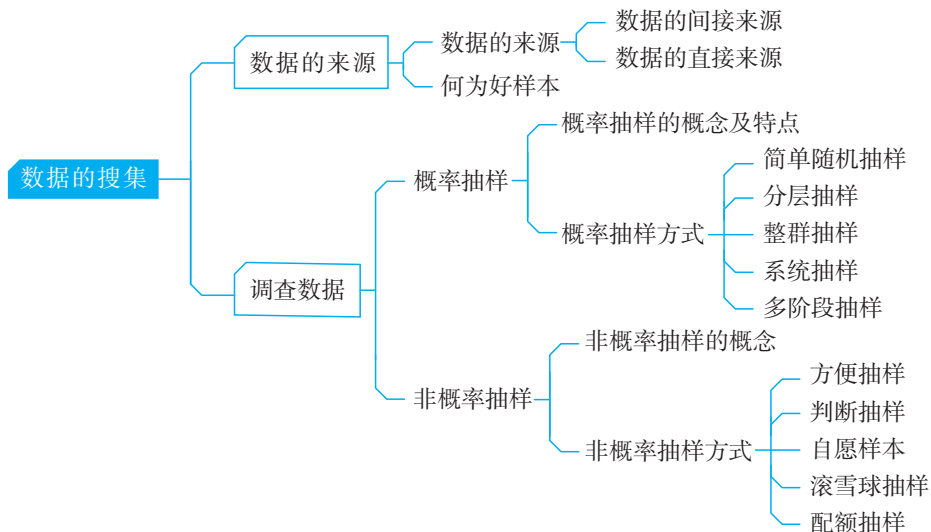
本章综述

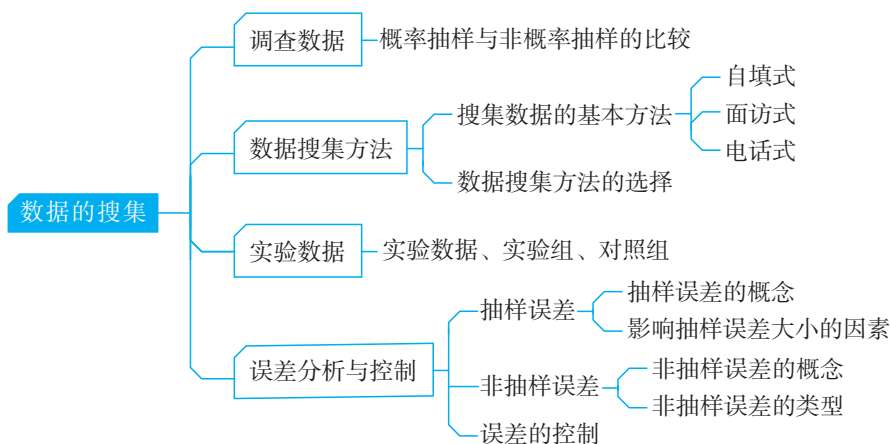
本章主要是基于统计数据的研究,从数据的搜集到研究数据本身所具有的性质。本章中历年考试重点为概率抽样与非概率抽样、获取数据的方式以及数据的误差。因此,要求考生具备如下能力:了解数据的直接来源与间接来源;掌握概率抽样与非概率抽样;了解实验数据相关内容;掌握数据的误差及误差控制方法。

考频统计



知识框架





考点一 数据的来源

考情提要

难度分析	☆
考频分析	☆
高频考点分析	(1)数据的间接来源与直接来源的特点。 (2)理解什么样的样本为好样本
题型分析	选择:2011、2012 首经贸

考点精讲

(一) 数据的来源

1. 数据的间接来源

如果与研究内容有关的原信息已经存在,我们只是对这些原信息重新加工、整理,使之成为我们进行统计分析可以使用的数据,则把它们称为间接来源的数据,也称二手数据。

从搜集的范围看,这些数据可以取自系统外部,也可以取自系统内部。数据取自系统外部的渠道有:统计部门和各级政府部门公布的有关资料;各类经济信息中心、信息咨询机构提供的市场信息和行业发展的数据情报;各类专业期刊、报纸、图书提供的文献资料;各种会议上交流的有关资料等。取自系统内部的资料,如果就经济活动而言,则主要包括业务资料,如与业务经营活动有关的各种单据、记录;经营活动中的各种统计报表等。

优点:二手资料的搜集比较方便,数据采集快、成本低。二手资料作用广泛,除了分析所要研究的问题,这些资料还可以提供研究问题的背景,帮助研究者更好地定义问题,检验和回答某些