

图书在版编目(CIP)数据

高中同步测控优化设计. 数学. 必修 :Ⅲ. 新课标北师大版 /任志鸿主编.
- 2 版. - 海口 :南方出版社 2004. 12(2006. 11 重印)
(志鸿优化系列丛书)
ISBN 7-80660-498-7

I. 高... II. 任... III. 数学课 - 高中 - 教学参考资料 IV. G634

中国版本图书馆 CIP 数据核字(2004)第 132124 号

责任编辑 余云华

装帧设计 邢 丽

志鸿优化系列丛书

高中同步测控优化设计. 数学. 必修 :Ⅲ

任志鸿 主编

南方出版社 出版

(海南省海口市海府一横路 19 号华宇大厦 12 楼)

邮编 570203 电话 0898 - 65371546

山东鸿杰印务有限公司印刷

山东世纪天鸿书业有限公司总发行

2006 年 11 月第 4 版 2006 年 11 月第 1 次印刷

开本 890 × 1240 1/16

印张 6.75 字数 181 千字

定价 10.50 元

(如有印装质量问题请与承印厂调换)



目录

conten

第一章 统 计	1
§ 1 从普查到抽样	2
§ 2 抽样方法	4
§ 3 统计图表	12
§ 4 数据的数字特征	17
§ 5 用样本估计总体	22
§ 6 统计活动 结婚年龄的变化	28
§ 7 相关性	32
§ 8 最小二乘估计	37
本章测评	42
第二章 算法初步	46
§ 1 算法的基本思想	46
§ 2 算法的基本结构及设计	52
§ 3 几种基本语句	58
本章测评	62
第三章 概 率	65
§ 1 随机事件的概率	66
§ 2 古典概型	71
§ 3 模拟方法——概率的应用	77
本章测评	81
模块综合测评	84

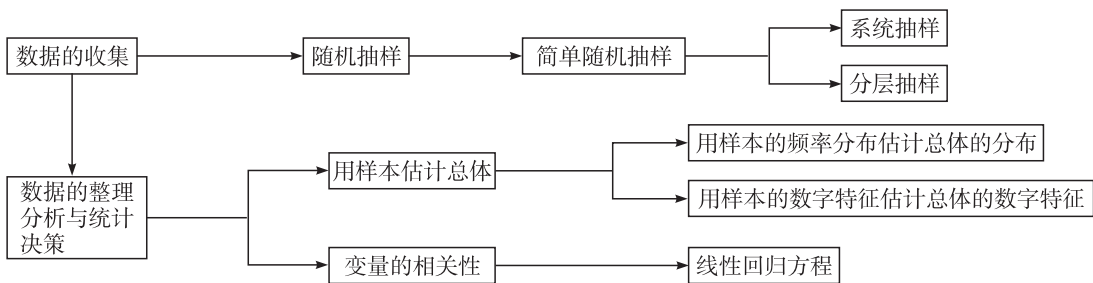
第一章 统 计



三维目标

- 1. 从现实生活或其他学科中提出具有一定价值的统计问题 ,理解随机抽样的必要性和重要性 ,学会用简单随机抽样方法从总体中抽取样本 ,通过对实例的分析 ,了解分层抽样和系统抽样方法.
- 2. 通过实例体会分布的意义和作用 ,在表示样本数据的过程中 ,学会列出频率分布表、画频率分布直方图 ,频率折线图、茎叶图 ,体会它们的特点 ,并能通过实例理解样本数据标准差的意义和作用 ,学会计算数据标准差. 从而培养数形结合的思想.
- 3. 能根据实际需求合理地选取样本 ,从样本数据中提取基本的数字特征 ,进一步体会用样本估计总体的思想 ,体会样本频率分布和数字特征的随机性 ,解决一些简单的实际问题.
- 4. 在对数据处理过程中学习一些数据处理的方法 ,并运用所学知识、方法解决实际问题 ,体会统计思维与确定性思维的差异 ,注意到统计结果的随机性 ,统计推断是有可能犯错误的 ,体会统计的作用和基本思想 ,从而形成初步评价的意识. 通过实习作业培养独立思考、合作与交流的意识.
- 5. 经历用不同估算方法描述两个变量线性相关的过程 ,体会最小二乘法的思想 ,能根据给出的线性回归方程系数公式建立线性回归方程 ,从而体会数学的科学价值、应用价值等.
- 6. 现代社会是信息化的社会 ,人们常常需要收集数据 ,根据所获得的数据提取有价值的信息 ,作出合理的决策. 统计是研究如何合理收集、整理、分析数据的学科 ,它可以为人们制定决策提供依据. 因此 ,统计基础知识已经成为一个未来公民必备的常识. 通过本章对数据的收集、整理和分析 ,可增强社会实践能力、提高分析问题和解决问题的能力.

知识网络



§ 1 从普查到抽样

知识梳理

1. 在我们现实生活中有很多问题需要普查,但普查存在两个缺陷,即工作量大、对检验对象具有破坏性,所以很多时候,普查是行不通的.

2. 在实际调查中,通常从调查对象中按照一定的方法抽取一部分进行调查或观测,获取数据,并以此对调查对象的某项指标做出推断,这就是抽样调查.其中调查对象的全体称为总体,被抽的一部分称为个体.

3. 抽样调查与普查相比,有较突出的两个优点:
(1)迅速、及时;(2)节约人力、物力和财力.

知识导学

学习本节时,同学们要尽可能借助于熟知的实际调查的例子,理解学习普查和抽样调查的特点和区别.人口普查是大家熟知的实例,从中一方面可以感知我国政府进行普查的必要性和重要性,另一方面了解普查工作的烦琐和艰辛.同时联系我们日常生活中的某项调查、产品质量抽查、性能检测等调查活动的实例,感悟抽样调查的优越性,普查的不可实施性.

在统计中,有时由于检验对象的量很大,在很多情况下,很难做到对所有考查的对象作全面的观测,有时根本无法施行.例如测试灯泡的寿命、医生检验人的血液中血脂的含量、判断山东省的成年人平均身高是否为全国之最等,这些试验有的是破坏性的,有的由于测试的总体包含的成员数量很大,如果逐一测试,要消耗大量的时间、人力、物力,得不偿失.一个行之有效的方法是从总体中选取部分个体,记录下来,并从这组数据来推断总体的情况.

抽样调查,就是通过从总体中抽取一部分个体进行调查,借以获得对整个总体特征的了解.如何抽取样本,直接关系到对总体估计的准确程度.抽样时要保证每一个个体都可能被抽到,即每一个个体被抽到的机会是均等的.

疑难突破

1. 抽样是了解数据的一种有效方法,但不是所有的问题都能通过抽样进行,有些数据必须采用普查的方法进行.那么在实际中,怎样合理地选择这两种方法呢?

剖析:普查和抽样是我们调查问题的常用方法,我们

必须掌握它们的特点并灵活应用.合理地进行选择才能更加快捷地得到更可靠的数据,下面列表表示抽样与普查的特点:

调查方法 特点	普 查	抽 样
优点	1. 所取得的资料更加全面系统 2. 调查某时段的总体的数量	1. 迅速、及时 2. 节约人力、物力、财力,对个体信息的了解更详细
缺点	耗费大量的人力、物力、财力	获取的信息不够全面系统
适用范围	总体容量不大,要获取详实、系统、全面的信息	1. 大批量检验 2. 破坏性检验 3. 不必要普查等

2. 在数据的收集,抽样是一种很重要的方法,但是抽样不能无原则地进行,那么,要使抽样调查的数据更具有合理性,还应该注意哪些问题呢?

剖析:抽样是为了通过样本来了解整体的数据特征,抽样不能是无原则的,也要遵循一定的规律,在实际中要注意以下几点:

首先,抽样调查就是通过从总体中抽取一部分个体进行调查,借以获得对整个总体特征的了解.所以,如何抽取样本,直接关系到对总体估计的准确程度,因此抽样时要保证每一个个体都可能被抽到,每一个个体被抽到的机会是均等的,要尽可能地避免人为因素的干扰,同时,还要注意到要尽可能地控制抽样调查中的误差.

其次,在抽样调查中应注意:抽取的样本具有全面性、代表性、随机性;片面地选取样本是不合理的,得出的答案也是不科学的.因此抽样调查的关键在于合理选取样本.

典题精讲

例1 为了准确调查我国某一时期的人口总量、人口分布、民族人口、城乡人口、受教育的程度、迁移流动、就业状况、人口住房等多方面情况,需要什么样的统计方法呢?

思路分析 凡是大批量的,或有破坏性的检验通常用抽样调查的方法,而在总体容量不是很大的情况下,要获得更系统的信息,通常用普查的方法.

解 要求全面准确调查我国的人口状况,因此应当用普查的方法进行调查.

绿色通道 要获得更系统的信息,在对总体没有破坏性的前提下,如果想获得第一手的统计数据及资料,普查无疑是一个非常好的方法,但这一方法任务量大,要耗费大量的人力、物力、财力等.

变式训练 某工厂要检验一批产品的质量,决定从这批产品中任意抽取 10 个进行检验,以判断产品的质量如何.在这个题目中,总体是_____,样本是_____.

思路解析 总体是指所要研究对象的全体,简称总体或母体,称组成总体的每个成员(单个元素)为个体,称从总体中随机抽出来的若干个个体的集合为样本或子样,称样本中个体的个数为样本大小或样本容量.

答案 :一批产品 10 个产品

例 2 某饮料厂,出厂一批罐装饮料,检验员为了检验产品的质量应采用什么方法检验,并说明其合理性.

思路分析 破坏性检验只能用抽样调查方法进行抽样调查.

解 这是大批量的破坏性的检验,不可能进行普查,因此应当采取抽样的方法进行检验,对随机抽取的部分产品进行检验,得到的检验结果,就可以看作是这批罐装饮料的质量情况.因为同一批产品,从中随机抽取一部分,能代表全体产品的质量,所以是合理的.

绿色通道 对于检验的总体,如果这种检验带有破坏性,那么这种普查就不合理,也不可行,抽样检验是合理而且必要的.可以设想,产品的质量合格率仅是一个近似值,它不能代表所有批次的产品均有这个比率,这就是抽样存在的问题.

问题探究

问题 我国第五次人口普查中漏登的人数大约是多少?你对人口普查的漏登率是如何认识的?

导思 :了解第五次人口普查的情况,并加以分析.

探究 根据人口普查登记质量抽查表明漏登率为 1.81%.第五次全国普查结果,我国人口达到 12.95 亿.因此我国第五次人口普查中漏登的人数大约是二千三百四十三万九千五百人.漏登率是指某地区漏登人数与应统计人数之比,是通过进一步抽查某些地区漏登人数后,计算出来的结果.由于流动人口的存在,和工作中等原因,出现漏登是不可避免的.

自主广场

我夯基 我达标

1 医生要检验人血液中血脂的含量,采取的调查方法应该是_____ ()

- A. 普查
- B. 抽样调查
- C. 既不能普查也不能抽样调查
- D. 普查与抽样调查都可以

思路解析 医生要检验人血液中血脂的含量,不可能将病人的所有血液都抽出来检查,因此不能用普查,只需抽出适量的血液作为样本来检验即可.

答案 B

2 若要调查某城市家庭的收入情况,在该问题中,总体是_____ ()

- A. 某城市
- B. 某城市的所有家庭的收入
- C. 某城市的所有人口
- D. 某城市的工薪阶层

思路解析 要调查某城市家庭的收入情况,在该问题中,总体是某城市的所有家庭的收入,应选 B.

答案 B

3 抽样调查在抽取调查对象时_____ ()

- A. 按一定的方法抽取
- B. 随便抽取
- C. 全部抽取
- D. 根据个人的爱好抽取

思路解析 抽样调查在抽取调查对象时必须能能保证所抽取出的样本具有代表性,使每个个体被抽入样的可能性相等,因此抽样时一定要注意事先设计好抽样的程序,按既定的程序进行抽样.

答案 A

4 普查是一项非常艰巨的工作,当_____时,普查无疑是一项非常好的调查方式;当_____时,普查工作量就很大,在很多情况下,_____工作难以实现.

思路解析 在总体容量不是很大的情况下,普查是全面获取信息最可靠的方法.

答案 普查的对象很少 普查的对象很多 普查

5 你班的班主任想全面了解你班学生的学习和思想状况,请你帮助班主任设计一个调查方案.

思路分析 在总体容量不是很大的情况下,普查是全面获取信息最可靠的方法,它有两个特点:

- (1) 所得资料更加全面系统;
- (2) 能够得到某个时期的信息总量.

解 因为一个班的人数不是太多,为了帮助班主任全面了解班里学生的学习和思想状况,可以采取普查的方法进行调查.你可以先设计一个问卷,包括同学们

对学习的各种看法,同学们的爱好、心理和思想状况等,然后发放给每一个学生,并全部收回,然后进行统计.这样就可以全面了解每个学生的学习和思想状况了.

- 6 在食品质量检验中,为了检验某批次袋装牛奶(10万包)的细菌超标情况,请你说出检验方法,并说明其合理性.

思路分析 这是大批量且有破坏性的检验问题,只能进行抽样调查,因为同一批次牛奶细菌超标情况没有大的差异,所以这样检验是科学合理的.

解 要检验某批次袋装牛奶的细菌超标情况几乎不可能将每一包牛奶进行检验,也就是不可能进行普查,因此,我们只要抽取少量进行检验就可以了,然后推断这批袋装牛奶的细菌是否超标,并对超标情况进行统计,认为这批牛奶的细菌超标情况基本如此.

我综合 我发展

- 7 某玻璃厂要检验一批次(10万块)玻璃的质量(包括硬度、承受压力能力)应如何检验?并说明其合理性.

思路分析 抽样调查与普查相比有很多优点,最突出的有两点:(1)迅速、及时;(2)节约人力、物力、财力.对一个问题的调查,要具体问题具体分析,根据普查与抽查的特点,选用科学合理的方法.

解 我们知道,要检验玻璃的质量,不可能将每块玻璃都进行试验,因此我们检验这批玻璃时,可以抽取少量进行试验,由此来推断玻璃的质量.

- 8 如果现在有一项调查,调查你们学校学生的家庭平均月收入情况,那么你会怎样做?将你的想法写成调查方案,并与同学交流你的调查方案与想法,看看是否需要改进的地方.

思路分析 由于学校人数较多,用普查的方法工作量太大,所以可以用抽样调查的方法.下面是三组可能的普查方案:

(1)组同学可能想先确定每个班要抽查的人数,然后用随机抽样的方法,抽取部分同学进行问卷调查,最后汇总各班情况进行统计,这是一个比较合理的方法.

(2)组同学可能想先找到全校学生的学籍号,然后隔一定人数选出一位同学,这样找出了你要调查的样本,然后进行问卷调查,最后进行统计,得出结果,这也是一个不错的方法.

(3)组同学可能想到,每位同学的家庭收入不同,先选10个家庭收入较高的调查,再选10个家庭收入中等的调查,最后选10个家庭收入较低的调查.

解 设计合理的调查方案是调查的基础,是统计活动中非常重要的环节.

(1)组同学在一个班抽取的被调查人,一定要随机抽取,可以用抓阄的方法.

(2)组同学的方法是比较科学的,以后我们还会学习这一抽样方法.

(3)组同学这种方法不是很合理,因为三种情况的家庭并不均等,应需要改进.

§2 抽样方法

知识梳理

1. 简单随机抽样

(1)设一个总体含有 N 个个体,从中逐个不放回地抽取 n 个个体为样本($n \leq N$),如果在抽取过程中,总体内的每个个体被抽到的概率相同,这样的抽样方法就叫做简单随机抽样.

(2)最基本、最常用的简单随机抽样方法有两种,即抽签法和产生随机数的方法.

(3)抽签法的实施步骤可以归结为:①编号;②制签;③搅匀;④抽签;⑤获样.

(4)产生随机数法就是利用工具(转盘或摸球、随机数表、科学计算器或计算机)产生的随机数进行

抽样.

(5)简单随机抽样具有简便易操作的优点,在总体容量不大的情况下是行之有效的.

2. 分层抽样

(1)当总体中的个体由明显差异的几部分组成时,为了使抽取的样本更好地反映总体的情况,通常采用分层抽样法.

(2)实施分层抽样时,应将总体中各个个体按某种特征分成若干个互不重叠的几部分,每一部分叫做层,在各层中按层在总体中所占的比例进行简单随机抽样.

3. 系统抽样

(1)当总体容量和样本容量都很大时,不再适宜用简

单随机抽样和分层抽样,而系统抽样就可以解决这个问题.

(2)一般地,要从容量为 N 的总体中抽取容量为 n 的样本,采用系统抽样可按下列步骤进行:

- ①编号:(有时可直接使用个体自身所带的号码,如学号、准考证号、门牌号等).
- ②分段:对编号进行分段,要保证“等距”分段.
- ③确定起始号码:在第一段用随机抽样的方法确定起始的个体号码.
- ④按事先指定的规则抽取样本,通常是起始号码分段间隔,依次获取样本.

4. 三种抽样方法的比较

(1)三种抽样方法的区别与联系如下表:

类别	共同点	各自特点	相互联系	适用范围
简单随机抽样	抽样过程中每个个体被抽取的可能性相等	从总体中逐个抽样		总体中的个体数较少
系统抽样		将总体均分成几部分,按事先确定的规则在各部分抽取	在起始部分抽样时采用简单随机抽样的方法抽样	总体中的个体数较多
分层抽样		将总体分成几层,分层进行抽取	各层抽样时采用简单随机抽样或系统抽样的方法进行抽样	总体由差异明显的几部分组成

(2)简单随机抽样是系统抽样和分层抽样的基础,三种抽样方法的共同特点是:

- ①都是等可能性抽样;
- ②被抽取样本的总体的个数有限;
- ③逐个地进行抽取;
- ④都是不放回抽样.

知识导学

学习本节前应先复习初中所学的总体、个体、样本、样本容量的概念,并借助于实例了解抽取样本的必要性和重要性.

学习本节中的三种抽样方法(简单随机抽样、系统抽样、分层抽样)时,应结合生活中的实例去体会它们各自的特点、使用范围和操作步骤.

简单随机抽样(抽签法、产生随机数法)是获取样本最基本的抽样方法,尽管它有局限性(仅适宜总体数较少时),但它却是任何抽样方法的基础,也是不可缺少的方法,所以务必熟记它们的操作方法.

记忆小窍门:抽签法操作要点是编号、写签、搅匀、抽取.

随机数表法操作要点是编号、选起始数、读数、获取样本.

系统抽样和分层抽样是在总体中个体数较大和总体中个体具有明显的差异时,采用的抽样方法.

学习过程中,要弄懂系统抽样和分层抽样的适用范围、概念和操作步骤,结合实例理解其中的关键步骤,哪些步骤属于简单随机抽样.初步理解后,要及时结合具体问题,运用系统抽样法和分层抽样法设计抽样方案,完成抽样过程,加以巩固.

记忆小窍门:系统抽样的步骤可概括为:①编号,②定间隔分段,③在第一段内定起始号,④后加间隔得每段号码,获取样本.

当用分层抽样进行抽取样本时,需注意以下几个问题:

(1)必须先判断所考查的总体中的个体是否具有明显的差异,若是则用分层抽样,若不是,应用简单随机抽样或系统抽样.

(2)为获取各层入样数目,需先正确计算出抽样比 $k = \frac{\text{样本容量}}{\text{总体容量}}$.若 k 与某层个体数的积不是整数时,可四舍五入取整.

(3)体会分层抽样仍属于等可能抽样,用分层抽样获取的样本更能很好地代表总体.

疑 难 突 破

1. 系统抽样中如何对总体中的每个个体进行合理分段?

剖析:系统抽样操作的要领是先将个体数较多的总体分成均衡的若干部分,然后按照预先制定的规则,从每一部分中抽取 1 个个体,得到所需样本.由于抽样的间隔相等,因此系统抽样也被称作等距抽样(或叫机械抽样).所以系统抽样中必须对总体中的每个个体进行合理(即等距)分段.

若从容量为 N 的总体中抽取容量为 n 的样本,用系统抽样时,应先将总体中的各个个体编号,再确定分段间隔 k ,以便对总体编号进行分段.

当 $\frac{N}{n}$ 是整数时,取 $k = \frac{N}{n}$ 为分段间隔即可,如 $N = 100, n = 20$,则分段间隔 $k = \frac{100}{20} = 5$,也就是将 100 个个体平均分为 5 段.

特别地,当 $\frac{N}{n}$ 不是整数时,应先从总体中随机剔除一

些个体,使剩余个体数 N' 能被 n 整除,这时分段间隔 $k = \frac{N'}{n}$,如 $N = 101$, $n = 20$,则应先在总体中剔除 1 个个体,使剩余的总体容量(即 100)能被 20 整除,从而得出分段间隔 $k = \frac{100}{20} = 5$,也就是说,只需将 100 个个体平均分为 5 段。

一般地,用简单随机抽样的方法从总体中剔除部分个体,其个数为总体中的个体数除以样本容量所得的余数。

上述过程中,总体中的每个个体被取出(或被剔除)的可能性相等,也就是每个个体不被选取(或不被剔除)的可能性也相等,所以在整个抽样过程中每个个体被抽取的机会仍然都相等,这说明使用系统抽样法抽取样本的过程是公平的。

2. 分层抽样中各层入样的个体数应如何确定?

剖析:当已知总体由差异明显的几部分组成时,应将总体分成互不交叉的几部分,然后按照各部分所占的比例,从各部分中独立抽取一定数量的个体,再将各部分抽出的个体合在一起作为样本,这就是分层抽样。其中所分成的每一部分叫层。

由于层与层之间有明显的区别,而层内个体间差异很小,为了使样本更能充分地反映总体的情况,抽取样本时,必须照顾到各个层的个体。所以每层中所抽取的个体数应按各层个体数在总体中所占的比例抽取,也就是各层抽取的比例都等于样本容量在总体中的比例,即 $\frac{\text{样本容量}}{\text{总体容量}}$ 。这样抽取能使所得到的样本结构与总体结构基本相同,可以提高样本对总体的代表性。

在实际操作时,应先计算出抽样比 $k = \frac{\text{样本容量}}{\text{总体容量}}$,获得各层入样数的百分比,再按抽样比确定每层需要抽取的个体数。

例如:一批灯泡 400 只,其中 20 W、40 W、60 W 的数目分别为 200 只、150 只和 50 只,现用分层抽样的方法获取一个容量为 40 的样本,则三种灯泡依次抽取的个数可如下确定:

求出抽样比 $k = \frac{\text{样本容量}}{\text{总体容量}} = \frac{40}{400} = \frac{1}{10}$,则应抽取各种瓦数的灯泡只数分别为

$$200 \times \frac{1}{10} = 20, 150 \times \frac{1}{10} = 15, 50 \times \frac{1}{10} = 5.$$

典题精讲

例 1 一单位有职工 160 人,其中业务人员 96 人,管理人员 40 人,后勤服务人员 24 人,为了了解职工的收入情况,从中抽取一个容量为 20 的样本,分别按下述三种方法抽取:

①将 160 人从 1 至 160 编上号,再用白纸做成 1~160 号的签 160 个放入箱内拌匀,然后从中抽 20 个签,与签号相同的 20 个人被选出。

②将 160 人从 1 至 160 编上号,按编号顺序分成 20 组,每组 8 人,1~8 号,9~16 号,...,先从第一组中用抽签方式抽出 k 号,其余组 $(k+8n)$ 号($n=1, 2, \dots, 19$)亦被抽到,如此抽取 20 人。

③按 20: 160 = 1: 8 的比例,从业务人员中抽取 12 人,从管理人员中抽取 5 人,从后勤人员中抽取 3 人,都用随机数法从各类人员中抽取所需,他们合在一起恰好 20 人。

问:(1)上述三种方法中,总体、个体、样本分别是什么?

(2)上述三种方法中各自采取何种抽取样本的方法?

思路分析:三种常用抽样方法各具优点,它们之间既有联系又有区别,应注意对三种抽样方法进行比较。

解:(1)这三种抽取方式中,其总体都是该单位 160 名职工的收入,其个体都是单位各个职工的收入,其样本为所抽取的 20 名职工的收入。

(2)第一种为简单随机抽样,第二种为系统抽样,第三种为分层抽样。

简单随机抽样的步骤:①将总体中的个体编号;②选定开始的数字;③获取样本号码。

系统抽样的步骤:①确定分组情况和抽样距;②编号;③在第 1 段中用简单随机抽样确定起始的个体编号;④按照事先研究的规则抽取样本。

分层抽样的步骤:①分层;②按比例确定每层抽取个体的个数;③各层抽样(方法可以不同);④汇合成样本。

绿色通道 针对不同的问题,要选择科学实用的抽样方法。

简单随机抽样的方法包括抽签法和产生随机数法,它们主要适用于总体个体数较少的情况。抽签法的一般步骤:①编号;②准备抽签工具;③实施抽签。在抽签前一定要先把号签搅拌均匀,并且每次只能抽一个。抽签法最大的优点是简便易行,但此种方法不宜适用于总体数量较大的对象,一般适用于个体数量较少的对象。产生随机数法的一般步骤:①编号;②定位置,位置是任意的;③读数,要按一定顺序读数,顺序也是任意的,可以自上至下、自下至上、自左至右、自右至左。取数过程中,要把不符合要求的数(超过最大编码)和与前面重复的数去掉。

采用系统抽样时,要严格按其步骤进行。

变式训练 1 2006 天津河东一模 6 某公司在甲、乙、丙、丁四个地区分别有 150 个、120 个、180 个、150 个销售点,公司为了调查产品销售的情况,要从这 600 个销售点中抽取一个容量为 100 的样本,记这项调查为①,在丙地区中有 20 个物品型销售点,要从中抽取 7 个,调查其

销售收入和售后服务等情况,记这项调查为②.则完成①②这两项调查宜采用的抽样方法依次是……()

- A. 分层抽样法,系统抽样法
B. 分层抽样法,简单随机抽样法
C. 系统抽样法,分层抽样法
D. 简单随机抽样法,分层抽样法

思路解析:在①中,由于不同地区的产品销售情况差异较大,为了抽样的公平性,应采用分层抽样法.在②中,总体个体差异不大,总体中的个体数量也不大,故适宜用简单随机抽样.故选 B.

答案 B

变式训练 2 在下列问题中,各采用什么抽样方法抽取样本较为合适?

(1)从 20 台彩电中抽取 4 台进行质量检验;

(2)科学会堂有 32 排座位,每排有 40 个座位(座号为 1~40),一次报告会坐满了听众,会后为听取意见留下了第一列的所有 32 名听众进行座谈;

(3)光远中学有 180 名教职工,其中教师 136 名,管理人员 20 名,后勤服务人员 24 名,今从中抽取一个容量为 15 的样本.

解:(1)所述问题中总体中的个体数和样本容量均较少,故宜用简单随机抽样法.

(2)所述问题具有总体中的个体数较多,且每个个体无明显差异的特点,所以适宜用系统抽样法进行抽取样本.

(3)所述问题的总体中的个体具有明显差异,即出现了 3 个阶层,因此适宜用分层抽样法获取样本.

例 2 对某单位 1 000 名职工进行某项专门调查,调查的项目与职工任职年限有关,人事部门提供了如下资料:

任职年限	人数
5 年以下	300
5~10 年	500
10 年以上	200

试利用上述资料,设计一个样本容量为 100 的抽样方法.

思路分析:由于调查的项目与职工任职年限有关,而任职年限又分为三个层次,符合分层抽样的要求,故采取分层抽样的方法.

解:进行分层抽样可分为三层,

因为抽取样本容量与总体个体数之比为

$$\frac{100}{1\,000} = \frac{1}{10},$$

所以每层应抽取的个体数应为

$$5 \text{ 年以下 } 300 \times \frac{1}{10} = 30(\text{人}),$$

$$5 \sim 10 \text{ 年 } 500 \times \frac{1}{10} = 50(\text{人}),$$

$$10 \text{ 年以上 } 200 \times \frac{1}{10} = 20(\text{人}).$$

然后再在每层用简单随机抽样的方法进行抽样,将各层抽取的个体合在一起,就是所要抽取的样本.

绿色通道 由于分层抽样充分利用了我们所掌握的信息,使样本具有较好的代表性,而在各层中进行抽样时,大多数情况下采用简单随机抽样,有时也会用到其他方法,这样需根据问题的需要来决定.

分层抽样适用于总体由差异明显的几部分组成的情况.分层抽样时,分层所抽取的个体数等于该层个体总数与抽样比的乘积,抽样比 $= \frac{\text{抽取样本容量}}{\text{总体个体数}}$.

变式训练 1 **2006 四川高考卷 5** 甲校有 3 600 名学生,乙校有 5 400 名学生,丙校有 1 800 名学生,为统计三校学生某方面的情况,计划采用分层抽样法,抽取一个容量为 90 人的样本,应在这三校分别抽取学生……()

- A. 30 人,30 人,30 人
B. 30 人,45 人,15 人
C. 20 人,30 人,10 人
D. 30 人,50 人,10 人

思路解析:甲校有 3 600 名学生,乙校有 5 400 名学生,丙校有 1 800 名学生,为统计三校学生某方面的情况,计划采用分层抽样法,抽取一个容量为 90 人的样本,应在这三校分别抽取学生 30 人,45 人,15 人,选 B.

答案 B

变式训练 2 **2006 山东高考卷 13** 某学校共有师生 2 400 人,现用分层抽样的方法,从所有师生中抽取一个容量为 160 的样本,已知从学生中抽取的人数为 150,那么该学校的教师人数是_____.

思路解析:抽取教师为 $160 - 150 = 10(\text{人})$,所以学校教师人数为 $2\,400 \times \frac{10}{160} = 150(\text{人})$.

答案 150 人

例 3 某校高三年级共有 430 名学生,为了对某次考试的数学成绩作质量分析,打算从中抽取 40 人作样本,请你设计一个系统抽样,抽取上面所需的样本.

思路分析:总体中的每个个体,都必须是等可能地入样,为了实现“等距”入样,且又等可能性,因此,应先剔除 30 个个体,然后再分段,定起始位置.

解:先用简单随机抽样从总体中剔除 30 个个体(可用产生随机数法,也可用抽签法),再将剩下的 400 名学生进行编号 1 2 3 … 400,然后将总体分为 40 个部分,其中每个部分包括 10 个个体,如第 1 部分的个体编号为 1 2 3 … 10,从中随机抽取一个号码,比如 6,那么就可以从第 6 号开始,每隔 10 个抽取 1 个,这样得到容量为 40

的样本 6, 16, 26, 36, ..., 396.

绿色通道 利用抽签法抽取样本时,编号应从 1 开始,而利用随机数抽取样本时,编号应从 0 开始.例如将 100 个个体编号成 00, 01, 02, ..., 99, 而不是编号成 0, 1, 2, ..., 99, 以便于运用随机数表.将起始号码选为 00, 而不是 01, 可使 100 个个体都可用两位数字号码表示,否则将会出现三位数字号码 100. 利用随机数表产生随机数是最常用的产生随机数的方法,要掌握此种方法的步骤.

当总体容量和样本容量都很大时,采用简单随机抽样或分层抽样,都是非常麻烦的,系统抽样正好能解决这个问题.

用系统抽样抽取一定容量的样本时,首先要分清总体中的个数是否能被样本容量整除,否则就会出现抽样距不等的情况,就不合乎系统抽样的原则,还要注意总体的排列有没有明显的周期性,这时抽样距的选取要恰当,要打破周期性,如果总体事先排好序,要先打乱顺序,再抽样,以达到抽取的样本具有广泛的代表性.

变式训练 下列抽样试验中,最适宜用系统抽样法的是 ()

- A. 某市的 4 个区共有 2 000 名学生,且 4 个区的学生人数之比为 3: 2: 8: 2, 从中抽取 200 人入样
- B. 某厂生产的 2 000 个电子元件中随机抽取 5 个入样
- C. 从某厂生产的 2 000 个电子元件中随机抽取 200 个入样
- D. 从某厂生产的 20 个电子元件中随机抽取 5 个入样

思路解析 A 中总体有明显层次,不适用系统抽样法; B 中样本容量很小,适宜用简单随机抽样法中的随机数表法; D 中总体数很小,故适宜用抽签法. 只有 C 比较符合适用系统抽样法. 故选 C.

答案 C

例 4 2006 湖北高考卷 17 某单位最近组织了一次健身活动,活动分为登山组和游泳组,且每个职工至多参加了其中一组. 在参加活动的职工中,青年人占 42.5%, 中年人占 47.5%, 老年人占 10%. 登山组的职工占参加活动总人数的 $\frac{1}{4}$, 且该组中,青年人占 50%, 中年人占 40%, 老年人占 10%. 为了了解各组不同的年龄层次的职工对本次活动的满意程度,现用分层抽样的方法从参加活动的全体职工中抽取一个容量为 200 的样本. 试确定:

- (1) 游泳组中,青年人、中年人、老年人分别所占的比例;
- (2) 游泳组中,青年人、中年人、老年人分别应抽取的

人数.

思路分析 考虑到参加活动的职工差异比较明显,所以应将全体参加活动的职工合理分层,再从各层中按比例抽取入样人数. 为此需先计算出青年人、中年人、老年人分别所占的比例.

解 (1) 设登山组人数为 x , 游泳组中,青年人、中年人、老年人各占比例分别为 a 、 b 、 c ,

$$\text{则有 } \frac{x \cdot 40\% + 3xb}{4x} = 47.5\%, \quad \frac{x \cdot 10\% + 3xc}{4x} = 10\%,$$

$$\text{解得 } b = 50\%, \quad c = 10\%.$$

$$\text{故 } a = 100\% - 50\% - 10\% = 40\%,$$

即游泳组中,青年人、中年人、老年人各占比例分别为 40%、50%、10%.

$$(2) \text{ 游泳组中,抽取的青年人数为 } 200 \times \frac{3}{4} \times 40\% =$$

$$60(\text{人}), \text{ 抽取的中年人数为 } 200 \times \frac{3}{4} \times 50\% = 75(\text{人}), \text{ 抽}$$

$$\text{取的老年人数为 } 200 \times \frac{3}{4} \times 10\% = 15(\text{人}).$$

黑色陷阱 解答该题的关键是第(1)问,最容易出现的错误是不能正确运用变量将登山组人数和青年人、中年人、老年人各占比例数表示出来,必会导致表达混乱,直至出错. 其根本原因在于没有正确地阅读和理解题意,找不出解决问题的入手点. 合理的解题方案是弄懂题意后,选择登山组人数(设为 x)为出发点,进而可表示出参加活动的全体职工数($4x$),那么“在参加活动的职工中,中年人占 47.5%, 老年人占 10%”才能转化为相应的数学表达式,并由此求出游泳组中,青年人、中年人、老年人分别所占的比例.

问题探究

问题 1 从总体中抽取样本,再通过对样本的研究估计总体的情况是统计的基本思想. 那么,该怎样理解从总体中的这种抽样的思想? 请结合实际例子进行说明.

导思 可从发生在我们大家周围的许多实际问题出发,来理解统计的基本思想.

探究 对于统计的基本思想应作较广义的理解. 可以结合实际例子进行理解. 例如,当在同一条件下进行 100 次抛掷同一硬币的试验,其结果可以看成是,在相同条件下进行很多次这种试验的结果组成的总体中抽取的一个容量为 100 的样本,从而我们可以利用这个样本的情况去估计总体的情况. 能反映统计的基本思想的实际问题还有很多,如彩票、产品检验、电视节目收视率、学生的体重、商品的销售、景点的游客接待等都是通过样本数据来估计总体的情况的. 在统计中涉及的抽样方法很多. 如果按抽取样本时总体中的每个个体被抽取的可能性是否相等来分,可分为等可能性抽样和不等可能性抽样. 而实

实际的抽样总要考虑公平性问题,所以,我们最常用的是等可能性抽样,也就是在抽样中,每个个体被抽到的可能性是相等的.

问题2 放回和不放回抽样是两种不同的抽样方法,在实际中要根据问题进行适当选择.请结合抽样的过程,谈谈你对抽样方法中的放回抽样和不放回抽样的理解.

导思 在我们的实际生活中放回和不放回抽样都大量存在着,不妨借助熟悉的例子加以理解.

探究 在实际应用中,多采用不放回抽样,但像抛掷硬币、抛掷骰子等都是放回抽样,放回抽样在理论研究中较为重要.事实上,只有在放回抽样的情况下,才能保证每次抽样的结果作为随机变量所取的一个值都是相互独立且分布相同的.不过在实际应用中这两类抽样的区分有时并非那么严格,特别是当总体中的个体数很多甚至无限时,被抽取的个体是否放回总体对总体分布所造成的影响很小,因此这时的不放回抽样也可看成放回抽样,从而可以运用在放回抽样情况下可以运用的结论.

问题3 三种随机抽样方法有一个共同特点,即都是等可能抽样,也就是说不仅每次从总体中抽取一个个体时,各个个体被抽到的可能性相等,而且在整个抽样过程中,各个个体被抽到的可能性也相等,从而保证了这种抽样方法的公平性.那么为什么每个个体被抽到的机会是均等的呢?

导思 大家可以借助于下面的问题理解.

在一个盒子里装有大小、形状完全相同的5个白球和1个黑球,现有6人每人摸球一次,每次摸球一个,摸后不放回,试问6人中每人摸到黑球的机会是否相同?

探究 不妨按甲、乙、丙、丁、戊、己的顺序依次摸球.下面计算他们在摸球顺序不同的情况下摸到黑球的机会是否一样.

甲第一个摸球,摸到黑球的机会是 $\frac{1}{6}$.

乙第二个摸球,摸到黑球的机会是 $(1 - \frac{1}{6}) \times \frac{1}{5} = \frac{1}{6}$.

丙第三个摸球,摸到黑球的机会是 $(1 - \frac{1}{6} - \frac{1}{6}) \times \frac{1}{4} = \frac{1}{6}$.

丁第四个摸球,摸到黑球的机会是 $(1 - \frac{1}{6} - \frac{1}{6} - \frac{1}{6}) \times \frac{1}{3} = \frac{1}{6}$.

戊第五个摸球,摸到黑球的机会是 $(1 - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6}) \times \frac{1}{2} = \frac{1}{6}$.

己第六个摸球,摸到黑球的机会是 $(1 - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6}) \times \frac{1}{1} = \frac{1}{6}$.

$$\frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} = 0$$

这就是说,6人摸球次序有先有后,但每个人摸到黑球的机会是相等的.

类似的问题还有:

“摸彩票”问题,中奖号码是随机产生的,买彩票有先后,但每位买彩票的人获奖机会是均等的.



我夯基 我达标

1 系统抽样适用的总体应是……………()

- A. 容量较少的总体
- B. 总体容量较多
- C. 个体数较多但均衡的总体
- D. 任何总体

思路解析 与简单随机抽样相比,系统抽样的适用范围应是总体中的个体数目较多且含有的个体均衡.

答案 C

2 2006年7月2日~4日光明中学进行了05~06学年度期末统一考试,该校为了了解高一年级1 000名学生的考试成绩,从中随机抽取了100名学生的成绩单,就这个问题来说,下面说法正确的是……()

- A. 1 000名学生是总体
- B. 每个学生是个体
- C. 1 000名学生的成绩是一个个体
- D. 样本的容量是100

思路解析 根据总体、个体、样本、样本容量的概念,可知1 000名学生的成绩是统计中的总体,每个学生的成绩是个体,被抽取的100名学生的成绩是一个样本,其样本容量为100,所以D是正确的.

答案 D

3 某城区有农民、工人、知识分子家庭共计2 004家,其中农民家庭1 600户,工人家庭303户.现要从中抽取容量为40的样本,则在整个抽样过程中,可以用到下列抽样方法中的……………()

- ①简单随机抽样 ②系统抽样 ③分层抽样
- A. ②③ B. ①③ C. ③ D. ①②③

思路解析 由于各家庭有明显差异,所以首先应用分层抽样的方法分别从三类家庭中抽出若干户,即32户、6户、2户.又由于农民家庭户数较多,宜采用系统抽样法,而工人、知识分子家庭户数较少,宜采用简单随机抽样法.故整个抽样过程要用到①②③三种抽样法,所以选D.

答案 D

- 4 已知某单位有职工 120 人,其中男职工 90 人,现采用分层抽样(按男、女分层)抽取一个样本,若样本中有 27 名男职工,则样本容量为……………()
- A. 30 B. 36 C. 40 D. 无法确定

思路解析:各层中的抽样比为 $\frac{27}{90} = \frac{3}{10}$,所以样本容量为 $120 \times \frac{3}{10} = 36$.

答案 B

- 5 福利彩票的中奖号码是由 1~36 个号码中选出 7 个号码来按规则确定中奖情况,这种从 36 个选 7 个号的抽样方法是_____.

思路解析:由于从 1 到 36 共有 36 个号码数,即说明总体数较小,故可采用简单随机抽样.

答案 抽签法

- 6 某学校有老师 200 人,男学生 1 200 人,女学生 1 000 人.现用分层抽样的方法从全体师生中抽取一个容量为 n 的样本,已知从女学生中抽取的人数为 80 人,则 $n =$ _____.

思路解析:由已知得 $n = \frac{200 + 1\ 200 + 1000}{1\ 000} \times 80 = 192$.

答案 192

- 7 为了检验某种作业本的印刷质量,决定从一捆(40 本)中抽取 10 本进行检查,利用随机数表抽取这个样本时,应按怎样的步骤进行?

思路分析:可先对这 40 本作业本进行统一编号,然后在随机数表中任选一数作为起始号码,按任意方向读下去,便会得到 10 个号码.

解:可按以下步骤进行:

第一步,先将 40 本作业本编号,可编为 00,01,02,…,39.

第二步,在附录 1 随机数表中任选一个数作为开始.如从第 8 行第 5 列的数 59 开始.

第三步,从选定的数 59 开始向右读下去,得到一个两位数字号码 59,由于 $59 > 39$,将它去掉;继续向右读,得到 16,由于 $16 < 39$,将它取出;继续下去,可得到 19,10,12,07,39,38,33,21,后面一个是 12,由于在前面 12 已经取出,将它去掉;再继续读,得到 34.至此,10 个样本号码已经取满,于是,所要抽取的样本号码是 16,19,10,12,07,39,38,33,21,34.

- 8 某校 500 名学生中,O 型血有 200 人,A 型血有 125 人,B 型血有 125 人,AB 型血有 50 人,为了研究血型与色弱的关系,需从中抽取一个容量为 20 的样本.若按分层抽样方法进行抽取,则各种血型的人分别抽多少?写出抽样过程.

思路分析:用分层抽样方法抽样时,注意两个步骤:

①先求出样本容量与总个数的比值,即抽样比;②按

比例分配各层所要抽取的个体数.

解: $\frac{20}{500} = \frac{2}{50}$,即抽样比为 $\frac{2}{50}$.

, $200 \times \frac{2}{50} = 8$, $125 \times \frac{2}{50} = 5$, $50 \times \frac{2}{50} = 2$.

故 O 型血抽 8 人,A 型血抽 5 人,B 型血抽 5 人,AB 型血抽 2 人.

各种血型的抽取可用简单随机抽样(如 AB 型)或系统抽样(如 A 型),直至取出容量为 20 的样本.

我综合我发展

- 9 2006 重庆高考卷 6 某地区有 300 家商店,其中大型商店有 30 家,中型商店有 75 家,小型商店有 195 家.为了掌握各商店的营业情况,要从中抽取一个容量为 20 的样本.若采用分层抽样的方法,抽取的中型商店数是……………()
- A. 2 B. 3 C. 5 D. 13

思路解析:先算出各层的抽样比 $\frac{20}{300} = \frac{1}{15}$,所以所抽

取的中型商店数是 $\frac{1}{15} \times 75 = 5$.

答案 C

- 10 2005 湖北高考卷 12 某初级中学有学生 270 人,其中一年级 108 人,二、三年级各 81 人,现要利用抽样方法抽取 10 人参加某项调查,考虑选用简单随机抽样、分层抽样和系统抽样三种方案,使用简单随机抽样和分层抽样时,将学生按一、二、三年级依次统一编号为 1,2,…,270,使用系统抽样时,将学生统一随机编号 1,2,…,270,并将整个编号依次分为 10 段.如果抽得号码有下列四种情况:

①7,34,61,88,115,142,169,196,223,250;

②5,9,100,107,111,121,180,195,200,265;

③11,38,65,92,119,146,173,200,227,254;

④30,57,84,111,138,165,192,219,246,270.

关于上述样本的下列结论中,正确的是……()

A. ②③都不能为系统抽样

B. ②④都不能为分层抽样

C. ①④都可能为系统抽样

D. ①③都可能为分层抽样

思路解析:①②③都符合分层抽样的比例,①③是等距抽样,可能为系统抽样.

答案 D

- 11 2005 湖南高考卷 11 一工厂生产了某种产品 16 800 件,它们来自甲、乙、丙 3 条生产线.为检查这批产品的质量,决定采用分层抽样的方法进行抽样.已知从甲、乙、丙 3 条生产线抽取的个体数组成一个等差数列,则乙生产线生产了_____件产品.

思路解析: 设甲、乙分别生产了 $a-d$, a , $a+d$ 件, 则 $a-d+a+a+d=3a=16\ 800$, $a=5\ 600$.

答案 5 600

- 12 2005 山东高考卷 13 某学校共有教师 490 人, 其中不到 40 岁的有 350 人, 40 岁及以上的有 140 人, 为了了解普通话在该校教师中的推广普及情况, 用分层抽样的方法, 从全体教师中抽取一个容量为 70 人的样本进行普通话水平测试, 其中在不到 40 岁的教师中应抽取的人数是_____.

思路解析 算出各层的抽样比 $\frac{70}{490} = \frac{1}{7}$, 所以在不到 40 岁的教师中应抽取的人数为 $\frac{1}{7} \times 350 = 50$.

答案 50

- 13 某单位在岗职工共有 624 人, 为了调查工人用于上班途中的时间, 该单位工会决定抽取 10% 的工人进行调查. 请问如何采用系统抽样法完成这一抽样?

思路分析 因为 624 的 10% 为 62 人, 且 624 不能被 62 整除, 为了保证“等距”分段, 应先随机剔除 4 人. 这样就能使剩余的 620 人, 按每段 10 人“等距”地分为 62 段, 然后按照系统抽样的操作步骤, 确定样本.

解 采用系统抽样获取样本的操作过程如下:

第一步 将 624 名职工用随机方式编号.

第二步 从总体中剔除 4 人(剔除方法可用随机数表法), 将剩下的 620 名职工重新编号(分别是 000, 001, 002, ..., 619), 并分成 62 段.

第三步 在第一段 000, 001, ..., 009 这十个编号中, 用简单随机抽样抽取一个号码(如 002)作为起始号码.

第四步 将编号为 002, 012, 022, ..., 612 的个体抽出, 即可组成样本.

- 14 2006 山东济南一模 18 实验中学举行了为期 3 天的新世纪体育运动会, 同时进行全校精神文明擂台赛. 为了了解这次活动在全校师生员工中产生的影响, 分别在全校 500 名教职员工、3 000 名初中生、4 000 名高中生中作问卷调查, 如果要在所有答卷中抽出 120 份用于评估.

(1) 应如何抽取才能得到比较客观的评价结论?

(2) 要从 3 000 份初中生的答卷中抽取一个容量为 48 的样本, 如果采用简单的随机抽样, 应如何操作?

(3) 为了从 4 000 份高中生的答卷中抽取一个容量为 64 的样本, 如何使用系统抽样抽取到所需的样本?

思路分析 当总体是由差异明显的几个部分组成时, 一般用分层抽样抽取样本. 这样能使得到的结论比较客观, 分层时要注意用比例问题解题. 用系统抽样, 当总体容量 $\div$ 样本容量不能整除时, 要先从总体中剔除一部分个体.

解 (1) 应当采取分层抽样的方法进行抽样.

因为样本容量 = 120, 总体个数 = 7 500,

由 $120: 7\ 500 = 2: 125$, 可知应取

$$500 \times \frac{2}{125} = 8, 3\ 000 \times \frac{2}{125} = 48,$$

$$4\ 000 \times \frac{2}{125} = 64.$$

所以在教职员工、初中生、高中生中抽取的个体数分别是 8、48、64. 然后采用简单随机抽样的方式在教职员工的 500 份答卷中抽取 8 份, 在初中生 3 000 份答卷中抽取 48 份, 在高中生 4 000 份答卷中抽取 64 份, 得到容量为 120 的样本, 这样便完成了整个抽样过程, 就能得到比较客观的评价结论.

(2) 由于简单的随机抽样有两种方法, 为了保证操作过程中个体的抽取是等可能性的, 可选择使用抽签法或随机数表法.

方法一 将 3 000 份答卷都编上号码 0001, 0002, 0003, ..., 3 000, 然后做 3 000 个形状、大小相同的号签, 如用小球、卡片、纸条都可制作, 然后将这些号签放在同一个袋子里均匀搅拌, 每次从中抽出一个号签, 连续抽取 48 次, 就可得到一个容量为 48 的样本.

方法二 将 3 000 份答卷都编上号码 0001, 0002, 0003, ..., 3 000, 然后在随机数表上随机选取一个起始位置, 向右连续取数字, 以 4 个数为一组, 碰到右边线时接下一行左边线继续向右连续取数, 如从第 3 行第 8 列开始取, 所得数字如下: 7 665, 6 502, 6 710, 7 329, 0 797, 8 531, 3 553, 8 585, 9 889, 7 541, 4 101, 2 568, 5 992, 6 969, 6 682, 7 310, 5 037, 2 931, 5 571, 2 101, 4 218, 8 264, 9 817,

所取录的 4 位数如果小于或等于 3 000, 则对应此号的答卷就是被抽取的个体. 如果所取的 4 位数大于 3 000 而小于或等于 6 000, 则将其减去 3 000 所得差数就是被抽取的号码; 如果所取之数大于 6 000 而小于或等于 9 000, 则减去 6 000 后便为所取的号码, 依次类推. 如果遇到相同号码则只取一个而其余舍去. 经过这样的过程得到所取号码分别为 1 665, 0 502, 6 710, 1 329, 0 797, 2 531, 0 553, 2 585, 0 889, 1 541, 1 101, 2 568, 2 992, 0 969, 0 682, 1 310, 2 037, 2 931, 2 571, 2 101, 1 218, 2 264, 0 817, 0 555, 0 563, 2 643, 2 548, 2 462, 2 316, ..., 一直到取满 48 个号码为止.

(3) 由于 $4\ 000 \div 64 = 62.5$ 不是整数, 则应先使用简单随机抽样法从 4 000 个个体组成的整体中剔除 32 个个体, 再将剩余的 3 968 个个体进行编号: 1, 2, ..., 3 968, 然后将整体分为 64 个部分, 其中每个部分中含有 62 个个体, 如第 1 部分个体的编号为 1, 2, ..., 62. 从中随机抽取一个号码, 如若抽取的是 23, 则从第 23 号开始, 每隔 62 个抽取一个, 这样得到容量为 64 的样本: 23, 85, 147, 209, 271, 333, 395, 457, ..., 3 929.

§3 统计图表

知识梳理

1. 统计图表是表达和分析数据的重要工具,它不仅可以帮我们从数据中获取有用的信息,还可以直观、准确地理解相应的结果.

2. 常用的统计图表有条形统计图、扇形统计图、折线统计图和茎叶图.

3. 用茎叶图表示数据有两个突出的优点:一是所有数据信息不丢失,二是记录方便.

知识导学

学习本节,首先要明白在收集了样本数据后,下一步要做的工作是什么?应该是分析和处理数据,为此较理想的方法是将所得数据进行适当的整理、分析,并转化为直观的形式表现出来,以便从中获取相应的信息,帮助我们制定恰当的决策.所以学习本节时,可先回忆初中阶段所学的直观表达一组数据的方法,即用条形统计图、扇形统计图和折线统计图来表达一组数据的特征.

条形统计图、扇形统计图和折线统计图各自有哪些特点和用途?

1. 条形统计图是用一个单位长度表示一定的数量,根据数量的多少画成长短不同的直条,然后把这些直条按照一定的顺序排列起来.其特点是便于看出和比较各种数量的多少.

2. 扇形统计图中,用圆面代表总体,圆面中的各个扇形分别代表总体中的不同部分,扇形的大小反映部分占总体的百分比的大小.扇形统计图可以很清楚地表示各部分数量同总数之间的关系.

3. 折线统计图是用一个单位长度表示一定的数量,根据数量的多少描出各点,然后把各点用线段顺次连接起来.折线统计图不但可以表示出数量的多少,而且能够清楚地表示数量增减变化的情况.

另外有时还可以用茎叶图表示样本数据,它有两个突出优点:

(1)从统计图上没有原始信息的损失,所有的数据信息都可以从茎叶图中得到;

(2)茎叶图可以随时记录,方便表示与比较.

当数据量很大或有多组数据时,茎叶图就不那么直观、清晰了.

在对一些数据进行统计时,要根据数据的特点和统

计结果的精确度进行选择合适的统计图表.如果只需大致判断一些数据的分布规律,了解数据中各元素所占比例的大小情况可以使用扇形统计图.例如统计一个农村种植的各种作物的比例.如果需要根据图表了解各数据的频率可以使用条形统计图.例如统计一批产品中优等品的频率.如果要了解数据的增减情况可以采用折线图.例如统计一个人成绩变化情况.要了解数据的全部信息可以使用茎叶图.例如篮球比赛的计分.只有合理选择图表才能使统计的数据更加直观.

疑难突破

1. 统计图表是表达和分析数据的重要工具,它不仅可以帮我们从数据中获取有用的信息,还可以帮助我们直观、准确地理解相应的结果.如何根据不同的需要选择适当的统计图进行表示?

剖析:(1)条形统计图能清楚地表示出每个项目的具体数目,当数据量很大时,条形统计图能更直观地反映数据分布的大致情况,并且能够清晰地表示出各个区间的具体数目.

(2)折线统计图是把条形统计图各个长方形上边的中点用线段连接起来得到的,它能够清晰地反映数据的变化情况.

(3)扇形统计图能清楚地表示出各部分在总体中所占的百分比.

(4)茎叶统计图没有信息的损失,所有的原始数据都可以从这个茎叶图中得到.但茎叶图只适合于表示两位有效数字的数据,当数据量很大或有多组数据时,茎叶图就不那么直观、清晰了.

(5)当在问题中收集到的数据量较多时,在用统计图表示之前,一般需要先对数据进行整理.在此基础上,再根据不同的需要选择适当的统计图表示.

2. 普查或抽样的方法收集到的数据一般乱而多,如何进行数据的处理?

剖析:(1)我们无法将其包含的信息系统理解并加以表达,这就需要对数据进行整理和分析,将其转化成可以直接利用的形式,并从中获取相应的信息.统计图表正是表达和分析数据的重要工具,并且还可以直观地、准确地理解相应结果.

(2)注意理解掌握常用统计图的特点,根据题目和需要,去选择合适的统计图,并从不同的统计图中获得各种所需要的信息.

典题精讲

例 1 某赛季甲、乙两名篮球运动员每场得分情况如下：

甲的得分 12 15 24 25 31 31 36 36 37 39 44 49 50.

乙的得分 8 13 14 16 23 26 28 33 38 39 51.

(1)你能用适当的统计图表示上面的数据吗？

(2)根据你所画的统计图，分析甲、乙运动员的得分情况.

思路分析：由于所收集数据量不是太大，而且所给的数据均为一位数或两位数，不仅符合茎叶统计图的特征，而且第(2)问需要对这两组数据进行相互的比较，所以选用茎叶统计图来表示.

解：(1)如图1-3-1所示的茎叶图中，中间的数字表示两位运动员得分的十位数，两边的数字分别表示两个人各场比赛得分的个位数.

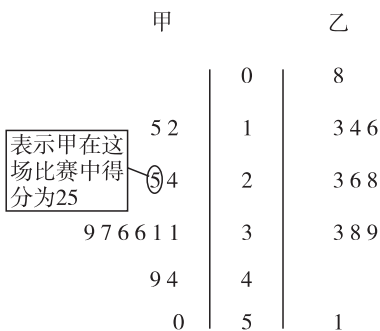


图 1-3-1

(2)根据上图可以对两名运动员的成绩进行比较.从茎叶图上可以看出，甲运动员的得分情况大致对称，中位数是 36；乙运动员的得分情况除一个特殊得分外，也大致对称，中位数是 26. 因此甲运动员的发挥比较稳定，总体得分情况比乙运动员好.

黑色陷阱：大多数同学会选用条形统计图或折线统计图，主要原因是该数据确实符合实际问题情境，但对于第(2)问甲、乙运动员的得分情况分析则图示得不到那么明显的结果.

例 2 下表给出了 2001 年 A、B 两地的降水量. (单位：mm)

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
A	9.2	4.9	5.4	18.6	38.0	106.3	54.4	128.9	62.9	73.6	26.2	10.6
B	41.4	53.3	178.8	273.5	384.9	432.4	67.5	228.5	201.4	147.3	28.0	19.1

请用适当的统计图表示上面的数据.

思路分析：题意要求能将所给的两组数据的值进行相互比较，且能观察各月份的差异及趋势，可利用条形统计图和折线统计图来表示.

解：用条形图 1-3-2 和折线图 1-3-3 表示上面的

数据.

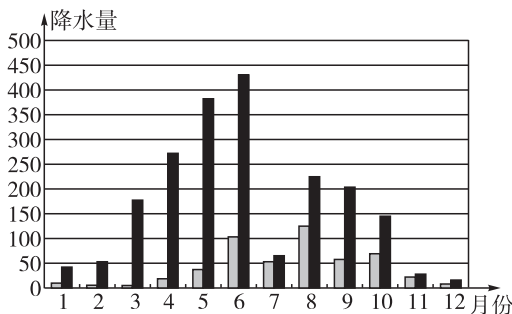


图 1-3-2

浅色为 A 地降水量，深色为 B 地降水量.

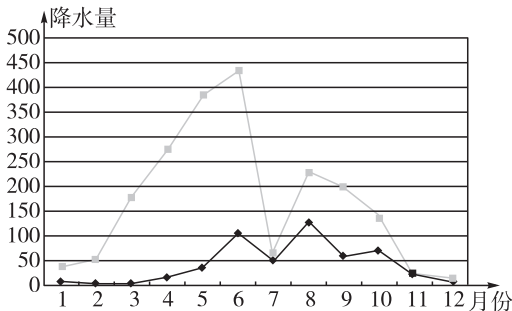


图 1-3-3

其中浅色为 B 地降水量，深色为 A 地降水量.

绿色通道：当两组数据具有可比性时，要想得出两组数据的差异比较，且能够直观观察出各组数据的分布趋势，而条形统计图与折线统计图同时具备这两种要求.

变式训练 某地农村某户农民年收入如下(单位：元)：

土地收入	打工收入	养殖收入	其他收入
4 320	3 600	2 350	850

请用不同的统计图来表示上面数据.

思路分析：题意的要求是将此四个数据用统计图展示出来，在所有的统计图中，可利用条形统计图、折线统计图、扇形统计图来表示.

解：我们分别用条形图 1-3-4、折线图 1-3-5 和扇形图 1-3-6 来表示.

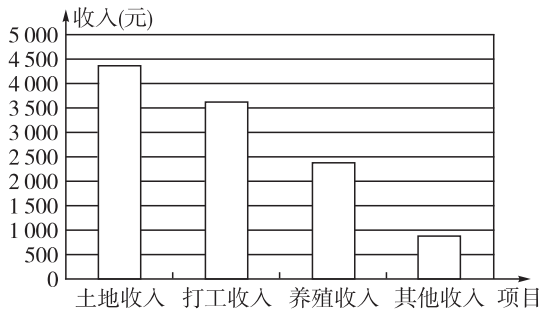


图 1-3-4

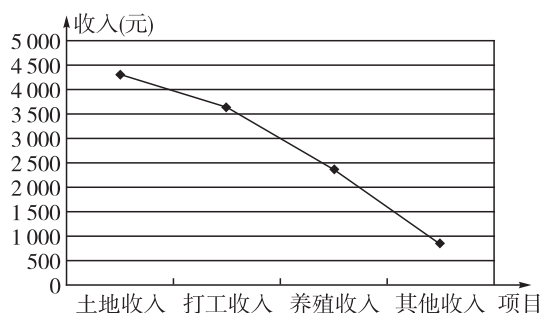


图 1-3-5

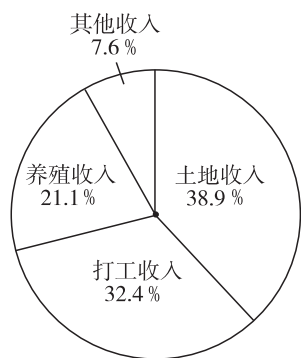


图 1-3-6

问题探究

问题 新浪科技讯 2006 年 4 月 13 日,中国互联网络信息中心(简称为:CNNIC)在京发布了北京市互联网络发展状况统计报告,数据显示,北京上网用户数为 398 万,占全国上网用户总人数的比例为 5%,占北京市人口的 28%,上网普及率居于全国首位。

北京市上网用户的职业分布如下表:

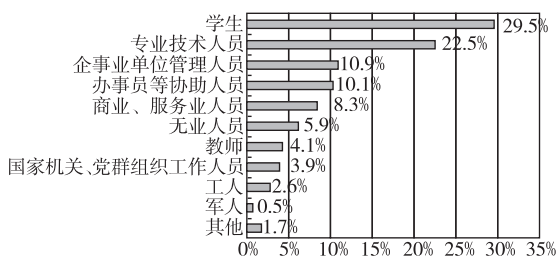


图 1-3-7

利用该图表,你能得到什么信息?你还有其他表示这些数据的方法吗?结合上述图表中的数据信息,你对各类上网人员情况有怎样的看法?

导思 由条形图可以直观、形象地看出不同职业人员上网的情况,职业分得越细,反映的总体信息越多.表达样本数据可使用不同的图表形式.从北京上网用户数和图表中的数据不难得到以下信息:各类人员上网人数及多少,上网用户的职业分配情况,各类人员上网的作用的

估计等.

探究 问题中给出了北京市各职业人员上网情况的条形图,从图中可直观、形象地看出各类上网人员所占的比例大小,学生上网的最多,占到约 30%,其次是专业技术人员,约占 23%,这两类人员约占总上网人数的 53%;工人和军人上网的人数较少,仅占 3%.

上述统计的数据,还可以用扇形统计图或折线图或茎叶图表示.下图是该问题中数据的扇形图(百分比)和折线图.

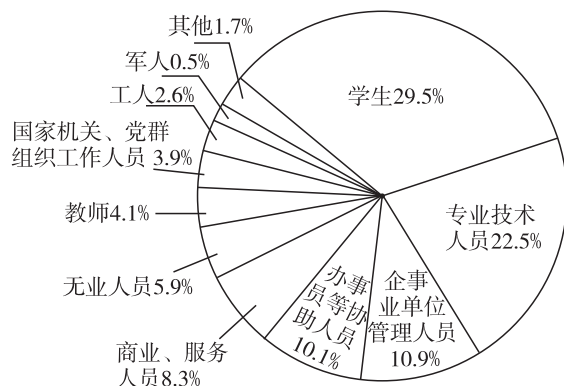


图 1-3-8

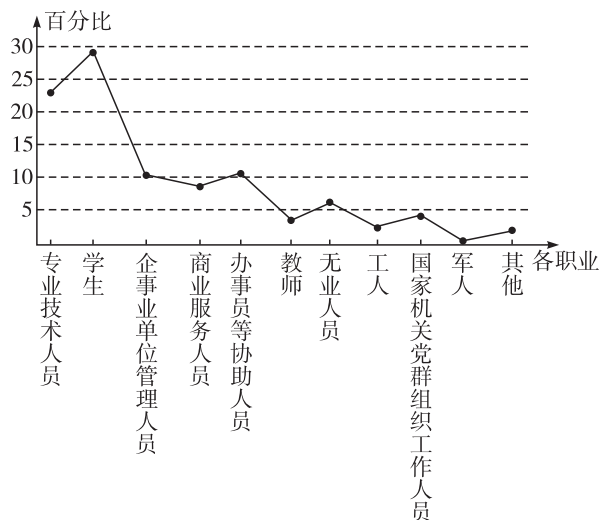
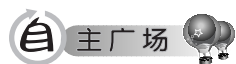


图 1-3-9

通过上述图表的信息,发现学生上网的最多,其次是专业技术人员,这两类人员约占总上网人数的一半多,工人和军人上网的人数较少,这说明近几年随着我国宽带网的普及,各类学生开始喜欢从互联网上获取信息,如学习知识、了解国家大事、相互交流等,而专业技术人员更加注意利用互联网进行资料查阅、文件下载、资源共享、不断学习;同时由于职业的特点或文化程度的限制,大部分工人、军人等特殊职业还不具备上网的条件或没有时间上网.所以从上述统计的图表中我们可以看出,目前北京市互联网在不同职业中的普及情况.进而也可知道全国各城市普及互联网的大致

情况.



我夯基 我达标

1 下面哪种统计图没有信息的损失,所有的原始数据都可以从该图中得到.....()

- A. 条形统计图
- B. 茎叶统计图
- C. 扇形统计图
- D. 折线统计图

思路解析 所有的统计图中,仅有茎叶统计图完好无损地保存着所有的数据信息.

答案 B

2 当收集到的数据量很大或有多组数据时,用哪种统计图表示较合适.....()

- A. 茎叶统计图
- B. 条形统计图
- C. 折线统计图
- D. 扇形统计图

思路解析 当收集到的数据量很大或有多组数据时条形统计图可以将各组数据分矩形趋势图显示出来.

答案 B

3 茎叶统计图表示数据有两个突出的优点:其一,____;其二,_____.

思路解析 茎叶统计图表示数据最突出的优点就是数据的原始性和其记录方便及可比较性.

答案 统计图上没有信息的损失,所有的原始数据都可以从这个茎叶统计图中得到 茎叶统计图可以随时记录,方便表示与比较

4 下面两幅统计图(如图 1-3-10、图 1-3-11),反映了某市甲、乙两所中学学生参加课外活动的情况.请你通过图中信息完成下面的问题.

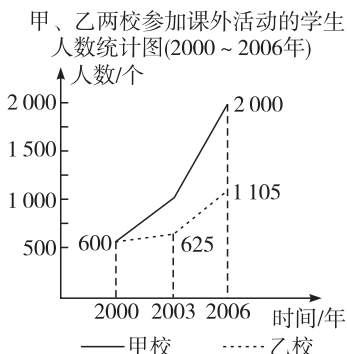


图 1-3-10

2006年甲、乙两校学生参加课外活动情况统计图

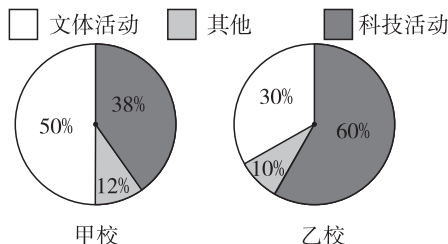


图 1-3-11

(1)通过对图 1-3-10 的分析,写出一条你认为正确的结论;

(2)通过对图 1-3-11 的分析,写出一条你认为正确的结论;

(3)2006 年甲、乙两所中学参加科技活动的学生人数共有多少?

思路分析 在统计活动中,用统计表和统计图来表示统计数据,显得更加直观、形象.本题通过多种统计图来比较甲、乙两所中学学生参加课外活动的情况.

解:(1)2000 年至 2006 年甲校学生参加课外活动的人数比乙校增长得快;

(2)甲校学生参加文体活动的人数比参加科技活动的人数多;

(3) $2\ 000 \times 38\% + 1\ 105 \times 60\% = 1\ 423$;

2006 年两所中学的学生参加科技活动的总人数是 1 423 人.

5 2005 年福州市居民的支出构成情况如下表所示.

食品	衣着	家庭设备用品及服务	医疗保健	交通和通讯	教育文化娱乐服务	居住	杂项商品和服务
40.4%	4.2%	8.9%	5.0%	8.9%	17.7%	11.5%	3.4%

用下列哪种统计图表示上面的数据较合适...()

- A. 都一样
- B. 茎叶统计图
- C. 扇形统计图
- D. 折线统计图

思路解析 扇形统计图和条形统计图均可以将统计中的所有数据所占整体百分比直观显示出来,但最佳的统计图表应当是扇形统计图,其显示得更为直观一点.

答案 C

6 2002 年 10 月 12 日《青年报》第 2 版刊载了一个扇形统计图,如图 1-3-12 所示.

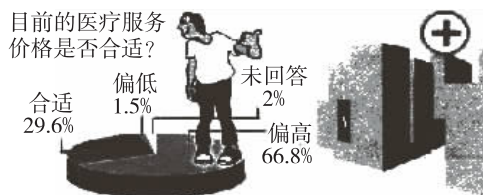


图 1-3-12

(1)从图上看,被调查者对目前的医疗服务价格是如