

五年制高等职业教育公共基础课程学习用书

数 学

学习指导用书

第三册

主 编 涂荣豹 李广全

3

江苏凤凰教育出版社 凤凰职教

图书在版编目(CIP)数据

数学学习指导用书. 第三册 / 涂荣豹, 李广全主编
. —南京 : 江苏凤凰教育出版社, 2024. 7(2025. 6 重印)
ISBN 978-7-5743-0878-7
I. ①数… II. ①涂… ②李… III. ①数学—高等职业教育—教学参考资料 IV. ①O1
中国国家版本馆 CIP 数据核字(2024)第 029845 号

五年制高等职业教育公共基础课程学习用书	
书 名	数学学习指导用书(第三册)
主 编	涂荣豹 李广全
责任编辑	宋 强
装帧设计	黄雨薇
出版发行	江苏凤凰教育出版社
地 址	南京市湖南路 1 号 A 楼, 邮编: 210009
出 品	江苏凤凰职业教育图书有限公司
网 址	http://www.fhmooc.com
照 排	南京普胜印刷技术有限公司
印 刷	南通韬奋印刷有限公司
厂 址	南通市通州区圩亭河西路 5 号, 邮编: 226300
电 话	0513-85675269
开 本	880 毫米×1 240 毫米 1/16
印 张	11.5
版次印次	2024 年 7 月第 1 版 2025 年 6 月第 2 次印刷
标准书号	ISBN 978-7-5743-0878-7
定 价	36.80 元
批发电话	025-83677909
盗版举报	025-83658893

如发现质量问题, 请联系我们。
【内容质量】电话: 025-83658873 邮箱: sunyi@ppm.cn
【印装质量】电话: 025-83677905

前言

本书是五年制高等职业教育公共基础课程教材《数学(第三册)》的配套学习用书,根据教材章节顺序编写,设有“课标要求”“知识导读”“学法建议”“例题导练”“达标练习”等栏目.

1. 课标要求. 将课标中的教学内容要求以及相关学业质量要求细化到每一课时.
2. 知识导读. 帮助学生整理课本中的重要概念、公式、定理、法则和重要结论.
3. 学法建议. 指出本课学习内容的重点、难点、关键点以及学习中需要注意的方面.
4. 例题导练. 在教材的基础上,选取相近的若干典型例题,选题时注意体现“三性”:
(1) 基础性. 重在帮助学生把握五年制高职数学学习的基础,使学生扎扎实实地掌握五年制高职数学基础知识和基本技能.
(2) 实用性. 最大程度地满足“教、学、练、考”的要求.
(3) 方向性. 主要表现在以课标为准绳,体现培育学生数学学科核心素养的特点,注重启发兴趣、诱导思维、探究规律,提高学生解决实际问题的能力;内容呈现上做到“思路分析”精准,解答简洁、规范,“思维诊断”画龙点睛.

5. 达标练习. 分为 A、B 两组. A 组以模仿教材中的例题、练习为主,为基本题;B 组在题型、方法、技巧等方面略有提高,难度、深度以课标学业水平二的学业质量描述为依据.

本书由涂荣豹、李广全主编,第 11 章由郑步春编写,第 12 章由丁阳、颜伟、张明琦编写,第 13 章由陈金国、王锋峰编写,第 14 章由周新华编写,第 15 章由吴萍编写.

由于编者水平有限,书中难免有不足和欠妥之处,恳请广大读者批评指正.

编者

2024 年 7 月

目 录

第 11 章 概率与统计初步

11.1 随机事件	1
11.2 古典概型	10
11.3 概率的简单性质	14
11.4 抽样方法	18
11.5 统计图表	29
11.6 样本均值与标准差	43

第 12 章 排列组合与二项式定理

12.1 分类、分步计数原理	50
12.2 排列与排列数公式	54
12.3 组合与组合数公式	62
12.4 排列组合的应用	72
12.5 二项式定理	75

第 13 章 数据表格信息处理

13.1 数据表格、数组	82
13.2 数组的运算	88
13.3 数据的图示	100
13.4 散点图及其数据拟合	108
13.5 用软件处理数据表格	112

第 14 章 三角计算

14.1	和角公式	116
14.2	倍角公式	127
14.3	正弦型函数	134
14.4	解三角形	146
14.5	三角计算的应用	156

第 15 章 坐标轴平移与参数方程

15.1	坐标轴平移	162
15.2	参数方程	168

第 11 章 概率与统计初步

11.1 随机事件

11.1.1 随机事件的概念

课标要求 KEBIAO YAOQIU

1. 理解确定性现象与随机现象的含义,理解必然事件、不可能事件及随机事件的意义.
2. 会判断随机现象、随机事件等问题,会判断必然事件和不可能事件.

知识导读 ZHISHI DAO DU

1. 随机现象:在一定的条件下,事先不能断定出现哪种结果的现象.
2. 确定性现象:在一定的条件下,事先就能断定出现或者不出现某种结果的现象.
3. 随机试验:对相同条件下可以重复进行的、结果有多种可能,并且具体结果不可预知的随机现象的观察或试验的统称.
4. 随机事件:在一定条件下可能发生也可能不发生的事件.
5. 必然事件:在一定条件下必然会发生的事件.
6. 不可能事件:在一定条件下一定不会发生的事件.
7. 基本事件:在每次试验中必有一个发生,也仅有一个发生的事件.
8. 复合事件:由几个基本事件组成,如果其中有一个发生,则该事件也一定发生的事件.



学法建议

XUEFA JIANYI

本课时学习重点是了解确定性现象与随机现象的意义；难点是理解随机试验的特点；关键是理解随机试验与随机事件的关系，能判定一个事件是随机事件、必然事件或不可能事件。

学习中要注意以下几点：

1. 对于现实世界中各种各样的现象，从结果能否预知的角度出发，可以分为两大类。一类现象的结果是无法预知的，即随机现象；一类现象的结果总是确定的，即确定性现象。
2. 研究随机现象，通常要进行观察。观察一定条件下随机现象发生的过程称为试验。如果一个试验可以在相同条件下重复进行，每次试验的结果不止一个，但事先能明确试验所有可能的结果，而且进行试验之前不能确定哪一个结果会出现，但必然会出现结果中的一个，那么称它为随机试验，简称试验。
3. 随机试验的特点：(1) 可重复性，即可以在同一条件下重复进行试验；(2) 可知性，即每次试验的结果不止一个，并且能事先明确试验的所有可能结果；(3) 不确定性，即试验之前无法确定哪一个结果会出现，但必然会出现结果中的一个。
4. 随机试验的结果称为随机事件。每次试验必然会发生的事件称为必然事件，每次试验肯定不会发生的事件称为不可能事件。
5. 基本事件的特征是，在每次试验中必有一个发生，也仅有一个发生。
6. 复合事件的特征是，由几个基本事件组成，如果其中有一个发生，则该事件也一定发生。



例题导练

LITI DAOLIAN

例 1 下列事件中，哪些是随机事件，哪些是必然事件，哪些是不可能事件？

- (1) 在标准大气压下，温度高于 0°C 时冰融化；
- (2) 如果 $a > 6$ ，那么 $a - 6 > 0$ ；
- (3) 从 3 件次品、1 件正品共 4 件产品中抽取 2 件产品，抽到的都是正品；
- (4) 从分别标有 1, 2, 3, 4, 5 的 5 张标签中任取一张，取到 6 号签；
- (5) 买一张体育彩票，中奖；
- (6) 将一颗骰子掷两次，所得点数之和大于 7。

【思路分析】 判断一个事件是随机事件、必然事件，还是不可能事件，可从它们的概念出发。

解 (1)(2) 中的事件必然会发生，因此是必然事件；



(3)(4)中的事件一定不会发生,因此是不可能事件;

(5)(6)中的事件可能发生也可能不发生,因此是随机事件.

【思维诊断】必然事件、不可能事件的特点是,要么发生,要么不发生,具有确定性.随机事件的特点是出现什么结果,具有多种可能,具体出现哪种结果有可能发生,也有可能不发生,即具有不确定性(随机性).

例 2 袋中有标号分别为 $1, 2, \dots, 10$ 的球各一只,从中取出一只球.记事件 $A_1 = \{\text{取出的是 1 号球}\}$, $A_2 = \{\text{取出的是 2 号球}\}$, $\dots$, $A_{10} = \{\text{取出的是 10 号球}\}$, $B = \{\text{取出的球标号不超过 3}\}$, $C = \{\text{取出的球标号不超过 10}\}$, $D = \{\text{取出的是 11 号球}\}$. 这些事件分别是什么事件?

【思路分析】袋中有标号分别为 $1, 2, \dots, 10$ 的球各一只,从中取出一只球,事先能明确所有可能取到的球的标号为 $1, 2, 3, \dots, 10$, 而且取出球之前不能确定会取到几号球,即事件 $A_1, A_2, \dots, A_{10}$ 可能发生,也有可能不发生,具有不确定性.

解 事件 $A_1, A_2, A_3, \dots, A_{10}, B$ 都是随机事件;

事件 C 是必然事件;

事件 D 是不可能事件.

【思维诊断】在每次试验中必有一个发生,也仅有一个发生的事件,称为基本事件.而事件 B 是由 A_1, A_2, A_3 这 3 个基本事件组成的,如果 A_1, A_2, A_3 中有一个发生,则事件 B 也一定发生,像这样包含若干个基本事件的事件称为复合事件.

例 3 将一枚质地均匀的硬币抛两次,观察两次正反面朝上的情况.

(1) 写出所有的基本事件;

(2) 如果 A 表示事件“第一次正面朝上”,写出组成 A 的基本事件.

【思路分析】将一枚质地均匀的硬币抛两次,每一次都可能是正面朝上,也可能是反面朝上,所以共有 4 个基本事件.

解 (1) 将一枚质地均匀的硬币抛两次,所有的基本事件是 $A_1 = \{\text{正, 正}\}$, $A_2 = \{\text{正, 反}\}$, $A_3 = \{\text{反, 正}\}$, $A_4 = \{\text{反, 反}\}$.

(2) A 表示“第一次正面朝上”,组成 A 的基本事件是 A_1 和 A_2 .

【思维诊断】本题可通过画树形图求解.画树形图的关键,一是确定层数,二是确定每层分叉的个数.

达标练习

A 组

1. 指出下列事件各是什么事件.

(1) 标准大气压下,加热到 100°C 时,水沸腾;

- (2) 篮球队员在罚球线上投篮一次,未投中;
- (3) 掷一次骰子,朝上一面的点数是 6;
- (4) 任意画一个三角形,其内角和是 360° ;
- (5) 经过有交通信号灯的路口,遇到红灯;
- (6) 射击运动员射击一次,命中靶心.

2. 从 $1, 2, 3, \dots, 10$ 十个数字中任取 3 个数字,那么“这 3 个数字的和大于 6”这一事件是().

- A. 必然事件
- B. 不可能事件
- C. 随机事件
- D. 以上选项均不正确

3. 有两个正四面体玩具,其四个面上分别标有数字 $1, 2, 3, 4$,下面做掷这两个正四面体玩具的试验:用 (x, y) 表示结果,其中 x 表示第 1 个正四面体玩具朝下一面的点数, y 表示第 2 个正四面体玩具朝下一面的点数,试写出:

- (1) 试验的基本事件;
- (2) 事件“ $x+y>3$ ”包含的基本事件;
- (3) 事件“ $x=y$ ”包含的基本事件.

B 组

1. 一只口袋内装有大小相同的 5 个球,其中 3 个为白球,球面上分别标有数字 $1, 2, 3$; 2 个为黑球,球面上分别标有数字 $4, 5$. 现从口袋中每次任取一球,每次取出不放回,连续取两次,并用 (x, y) 表示结果,其中 x 表示第 1 次取出的球上的数字, y 表示第 2 次取出的球



上的数字.

- (1) 列出这个试验的所有基本事件.
- (2) “取出的两个球均为白球”这一复合事件包含哪几个基本事件?
- (3) “取出的两个球均为黑球”这一复合事件包含哪几个基本事件?

2. 随机掷两枚质地均匀的正方体骰子, 骰子的六个面上分别刻有 1 到 6 的点数, 观察两颗骰子朝上一面的点数情况.

- (1) 写出所有的基本事件;
- (2) 如果 A 表示事件“两颗骰子朝上一面的点数都是奇数”, 写出事件 A 包含的基本事件.

3. 一只口袋中装有红、白、绿三个小球, 另一只口袋中装有红、白两个小球. 现从两只口袋中各取一个小球, 观察取出的两个小球的颜色情况.

- (1) 写出所有的基本事件;
- (2) 如果 A 表示事件“两个小球颜色相同”, 写出事件 A 包含的基本事件.

11.1.2 频率与概率

 课标要求 KEBIAO YAOQIU

1. 理解随机事件发生的不确定性.
2. 理解概率的统计定义,会根据概率的统计定义计算概率.
3. 能说出概率与频率的区别与联系.

 知识导读 ZHISHI DAO DU

1. 事件 A 发生的频率:在相同条件下重复进行 n 次试验,事件 A 发生的次数是 m ,就称比值 $\frac{m}{n}$ 为事件 A 发生的频率.
2. 随机事件的概率:对于给定的随机事件 A ,在相同条件下,随着试验次数的增加,事件 A 发生的频率会在某个常数附近摆动并趋于稳定.可以用这个常数来刻画随机事件 A 发生的可能性大小,这个常数称为随机事件 A 的概率,记作 $P(A)$.这种概率称为统计概率.
3. 概率与频率的关系:随机事件的概率刻画的是随机事件发生的可能性大小,是频率的稳定值,是确定的数;而频率是概率的近似值,在相同条件下进行的重复试验,得到的频率不一定相同.

 学法建议 XUEFA JIANYI

本课时学习重点是理解随机事件发生的不确定性,难点是理解概率的意义,关键是认识概率与频率的区别与联系.

学习中要注意以下几点:

1. 理解频率与概率这两个不同的概念.
 - (1) 频率是一个不确定的数,是在多次重复试验中某事件发生的次数与试验次数的比值,这个比值随着试验次数的增加而不断变化,即频率与试验次数有关.因此,频率在一定程度上反映事件发生的可能性大小,但无法从根本上刻画事件发生的可能性大小.
 - (2) 概率是一个确定的数.因为事件发生的可能性大小是客观存在的,即概率与试验次数是无关的.概率是针对大量试验而言的,随着试验次数的不断增加,某个随机事件发生的频率稳定在一个固定值.

2. 在实际应用中,通常将试验次数最多的频率值的最后一个有效数字四舍五入,作为概率的估计值.

例题导练

例 1 某足球运动员进行脚内侧射门训练,结果如表 11-1 所示.

表 11-1

射门次数 n	10	20	25	30	35	40
射进门内次数 m	8	18	22	27	31	37

(1) 计算表中射进门内的各个频率值(保留两位有效数字).

(2) 这名足球运动员射门一次,射进门内的概率估计值是多少?

【思路分析】 要计算某事件发生的频率,只要计算某事件发生的次数与试验次数的比值.由于当试验次数足够多的时候概率的估计值近似等于频率,通常将试验次数最多的频率值的最后一个有效数字四舍五入,作为概率的估计值.

解 (1) 射进门内的各个频率值如下表.

表 11-2

射门次数 n	10	20	25	30	35	40
射进门内次数 m	8	18	22	27	31	37
射进门内频率 $\frac{m}{n}$	0.80	0.90	0.88	0.90	0.89	0.93

(2) 从上表可看出,尽管射门次数 n 不同,射进门内次数 m 不同,但射进门内频率 $\frac{m}{n}$ 呈现一定的规律性和稳定性,即在数值 0.9 附近波动.因此这个足球运动员射门一次,射进门内的概率估计值为 0.9.

【思维诊断】 随机事件在一次试验中是否发生虽然不能事先确定,但是在大量重复试验的情况下,它的发生呈现出一定的规律性,可以用事件发生的频率去“测量”,因此可以通过计算事件发生的频率去估算概率.此类题目的解题方法是:先利用频率的计算公式依次计算出各个频率值,然后根据概率的定义确定频率的稳定值即为概率.

例 2 计算农作物种子发芽率规定的天数是:玉米、大豆、麦类、向日葵等为 7 天;高粱等为 8 天;水稻、绿豆、花生等为 10 天.水稻在相同条件下的发芽试验,经过 10 天水稻发芽终期正常发芽结果如表 11-3 所示,则水稻发芽终期正常发芽的概率估计值是多少?

表 11-3

每批粒数 n	100	300	400	600	1 000	2 000	3 000
发芽的粒数 m	96	282	382	570	948	1 912	2 850

【思路分析】 要计算概率的估计值,先要计算不同批次的水稻种子的发芽频率.

解 不同批次的水稻种子的发芽频率如下表.

表 11-4

每批粒数 n	100	300	400	600	1 000	2 000	3 000
发芽的粒数 m	96	282	382	570	948	1 912	2 850
发芽的频率 $\frac{m}{n}$	0.960	0.940	0.955	0.950	0.948	0.956	0.950

从上表可看出,尽管水稻在相同条件下的发芽试验每批粒数 n 不同,发芽终期正常发芽的粒数 m 不同,但发芽的频率 $\frac{m}{n}$ 呈现一定的规律性和稳定性,即在数值 0.95 附件波动.因此水稻发芽终期正常发芽的概率估计值为 0.95.

【思维诊断】 对于不同批次的水稻种子的发芽率往往误差会比较大,为了减少误差,我们经常采用多批次求平均数的方法计算频率来估计概率. $(96+282+382+570+948+1\,912+2\,850) \div (100+300+400+600+1\,000+2\,000+3\,000) \approx 0.951$, 当 n 足够大时,发芽的频率逐渐稳定于 0.95,故用频率估计概率,水稻发芽的概率估计值是 0.95.

例 3 下列三个命题是否正确?

- ① 设有一批产品,其次品率为 0.05,则从中任取 200 件,必有 10 件是次品.
- ② 做 100 次抛硬币的试验,结果 51 次出现正面朝上,因此,出现正面朝上的概率是 $\frac{51}{100}$.
- ③ 随机事件发生的频率就是这个随机事件发生的概率.

【思路分析】 次品率是次品数量与全部产品数量的比率.

解 命题①不正确,次品率是次品数量与全部产品数量的比值,并不是针对 200 件产品来说的,所以任取 200 件,不一定有 10 件是次品.

命题②③混淆了频率与概率.

达标练习

A 组

1. 某人将一枚质地均匀的硬币连续抛了 10 次,正面朝上的情形出现了 7 次,则下列说法中正确的是().

A. 正面朝上的概率为 0.7

B. 正面朝上的频率为 0.7

C. 正面朝上的概率为 $\frac{1}{2}$

D. 正面朝上的概率接近于 0.7

2. 根据某教育研究机构的统计资料,在校学生近视的概率为 40%,某眼镜商要到一中学给学生配眼镜,若已知该校学生总人数为 1 200,则该眼镜商应准备的眼镜数量为().

A. 460 副

B. 480 副

C. 不少于 480 副

D. 不多于 480 副

B 组

1. (1) 抛一枚硬币时,理论上“正面朝上”的概率为 $\frac{1}{2}$,那么抛 100 次硬币,你能保证恰好 50 次正面朝上吗?

(2) 如果中奖的概率为 $\frac{1}{1\,000}$,那么买 1 000 张奖券就一定能中奖吗?

2. 若某篮球运动员三分球的命中率为 60%,则下列说法中正确的是().

A. 他每投 10 次三分球,必正好有 6 次命中

B. 他每投 10 次三分球,必正好有 4 次不能命中

C. 如果他前 6 次三分球均命中,则第 7 次三分球必不能命中

D. 他连投 10 次三分球,命中次数可能少于 6 次

11.2 古典概型

 课标要求 KEBIAO YAOQIU

1. 理解等可能事件的意义,会把事件分解成等可能基本事件.
2. 理解古典概型的基本特点,掌握等可能事件的概率计算方法.
3. 会判断随机事件中的基本事件和古典概型,会求简单随机事件的古典概率.

 知识导读 ZHISHI DAO DU

1. 古典概型的定义:如果一个随机试验可能出现的结果只有有限个,即基本事件总数是有限的,并且每个基本事件发生的可能性相同,那么称这样的随机试验为古典概型试验,简称古典概型.

2. 古典概型的特征:随机试验可能出现的结果只有有限个,即基本事件总数是有限的,并且每个基本事件发生的可能性相同.

3. 古典概型的概率计算公式:

$$P(A) = \frac{\text{事件 } A \text{ 包含的基本事件数}}{\text{基本事件总数}} = \frac{m}{n}$$

 学法建议 XUEFA JIANYI

本课时学习重点是理解等可能事件的意义,会把事件分解成等可能基本事件,理解古典概型的基本特点;难点是将实际问题抽象为数学模型;关键是找出古典概型的基本事件总数与事件包含的基本事件数.

学习中要注意以下几点:

1. 通过重复试验观察一个事件发生的频率,可作为求概率近似值的一个具体方法,但通过大量重复试验用频率来表示概率有局限性,主要是由于进行大量重复试验工作量大,结果有一定的摆动性,有些试验还具有破坏性.

2. 古典概型的特点:在每次试验中,不同的试验结果只有有限个,即基本事件只有有限个,并且每个试验结果出现的可能性是相等的,即基本事件的发生是等可能的.由于这样的两个特点,求事件的概率可以不通过大量重复试验,而只要通过对一次试验中可能出现的结果进行分析和计算即可.

3. 运用古典概型的概率计算公式求事件发生的概率时,一般先要判断该问题是否是古典概型问题,然后再明确问题研究的试验和事件是什么.

例题导练

例 1 一个袋中装有 6 个球,其中 4 个是白球,2 个是红球.从袋中随机摸出 2 个球,求下列事件的概率.

- (1) 摸出的 2 个球都是白球;
- (2) 摸出的 2 个球 1 个是白球,另 1 个是红球.

【思路分析】 记“摸出的 2 个球都是白球”为事件 A ,“摸出的 2 个球 1 个是白球,另 1 个是红球”为事件 B ,为了利用公式 $P(A) = \frac{m}{n}$,先确定等可能的基本事件的总数 n ,再确定事件 A 和事件 B 包含的基本事件数 m .

解 设 4 个白球的编号为 1,2,3,4;2 个红球的编号为 5,6.从袋中的 6 个球中任意摸出 2 个,可能的结果[记“摸出 1,2 号球”为 $(1,2)$,其余情况类推]有 $(1,2), (1,3), (1,4), (1,5), (1,6), (2,3), (2,4), (2,5), (2,6), (3,4), (3,5), (3,6), (4,5), (4,6), (5,6)$,共 15 种.

(1) 从袋中的 6 个球中任意摸出 2 个,满足“2 个球都是白球”条件的结果有 $(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)$,共 6 种,故 $P(A) = \frac{6}{15} = \frac{2}{5}$.

(2) 从袋中的 6 个球中任意摸出 2 个,满足“1 个是白球,另 1 个是红球”条件的结果有 $(1,5), (1,6), (2,5), (2,6), (3,5), (3,6), (4,5), (4,6)$,共 8 种,故 $P(B) = \frac{8}{15}$.

例 2 把一个体积为 64 cm^3 的正方体木块表面涂上红漆,然后锯成 64 个体积均为 1 cm^3 的小正方体,并从中任取一块,试求:

- (1) 这一块没有涂红漆的概率;
- (2) 这一块恰有一面涂红漆的概率;
- (3) 这一块恰有两面涂红漆的概率;
- (4) 这一块恰有三面涂红漆的概率;
- (5) 这一块恰有四面涂红漆的概率.

【思路分析】 由于每一块小正方体被取到的可能性相等,并且等可能的基本事件的数量是有限的(64 个),故本题属于古典概型问题.关键是求出随机事件包含的基本事件数 m .

解 把体积为 64 cm^3 的正方体木块锯成 64 块体积为 1 cm^3 的小正方体,其中没有涂红漆的有 8 块,恰有一面涂红漆的有 24 块,恰有两面涂红漆的有 24 块,恰有三面涂红漆的有 8 块,恰有四面涂红漆的木块不存在.所以

- (1) “这一块没有涂红漆”的概率为 $P_1 = \frac{8}{64} = \frac{1}{8}$;

(2) “这一块恰有一面涂红漆”的概率为 $P_2 = \frac{24}{64} = \frac{3}{8}$;

(3) “这一块恰有两面涂红漆”的概率为 $P_3 = \frac{24}{64} = \frac{3}{8}$;

(4) “这一块恰有三面涂红漆”的概率为 $P_4 = \frac{8}{64} = \frac{1}{8}$;

(5) “这一块恰有四面涂红漆”是不可能事件,其概率为 $P_5 = 0$.

例 3 从有 2 件次品和 3 件正品的 5 件商品中抽取 2 件(不考虑抽取顺序),求抽取的 2 件中有 1 件次品的概率.

【思路分析】 对商品标注号码,列举出基本事件.

解 将 5 件商品标号,次品标号为 1,2,正品标号为 3,4,5,则抽取的可能结果有 (1,2),(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5),即基本事件总数 $n=10$. 每一基本事件是等可能发生的,故属古典概型.

抽取的 2 件中有 1 件次品的可能结果有 (1,3),(1,4),(1,5),(2,3),(2,4),(2,5),共 6 种.

记 $A = \{\text{抽取的 2 件中有 1 件次品}\}$,则事件 A 包含 10 个基本事件中的 6 个,即 $m=6$,故事件 A 发生的概率为 $P(A) = \frac{m}{n} = \frac{6}{10} = \frac{3}{5}$.

【思维诊断】 本例为不返回抽样问题,所谓不返回抽样就是指从 n 个不同元素中任意取出 1 个元素,不放回,再从剩下的 $n-1$ 个元素中任意取出 1 个元素,如此取下去,直到满足要求为止.



达标练习



DABIAO LIANXI

A 组

1. 下列试验,每一试验可能出现的结果(基本事件)的个数是有限个还是无限个? 每一试验的每一基本事件发生的可能性是否相同?

(1) 掷一颗质地均匀的骰子一次,观察骰子朝上一面的点数;

(2) 有方块 1,2,3 和红桃 4,5 共五张扑克牌,将它们牌点向下置于桌面上,现从中任意抽取一张,观察抽到的一张是什么牌;

(3) 袋中有红、白、黄三种不同颜色的球各一个,这三个球除颜色外,外形、质量等完全相同,从中任取一个球,观察球的颜色.



2. 在 10 件产品中,有 8 件是合格品,2 件是次品,从中任意抽取 2 件进行检验,求:

- (1) 2 件都是次品的概率;
- (2) 2 件中恰好有 1 件是合格品的概率;
- (3) 2 件中至多有 1 件是合格品的概率.

3. 将一颗骰子先后掷两次.

- (1) 一共有多少种不同的结果?
- (2) “朝上一面的点数之和是 7”的结果有多少种?
- (3) “朝上一面的点数之和是 7”的概率是多少?

B 组

1. 在某次测验中,用 x_n 表示编号为 n 的同学的成绩,且前 5 名同学的成绩如下表:

编号 n	1	2	3	4	5
成绩 x_n	70	76	72	70	72

从前 5 名同学中,随机地选 2 名同学,求恰有 1 名同学成绩在区间 $(68, 75)$ 中的概率.

2. 一只口袋里装有 2 个白球和 2 个黑球,这 4 个球除颜色外完全相同. 四个人按顺序依次从中摸出一球,求第二个人摸到白球的概率.

11.3

概率的简单性质



学业要求

XUEYE YAOQIU

1. 理解随机事件概率的取值范围;理解互斥事件及对立事件的概念,能判断两个事件是否是互斥事件,进而判断它们是否是对立事件.
2. 能用加法公式计算两个互斥事件的概率.
3. 知道“对立事件概率之和为 1”的结论,会用相关公式进行简单的概率计算.



知识导读

ZHISHI DAODU

1. 互斥事件:在一次试验中不可能同时发生的两个事件.
2. 对立事件:在一次试验中,其中必有一个发生的两个互斥事件.
3. 概率的简单性质
 - (1) 必然事件 U 的概率为 1,即 $P(U)=1$;
 - (2) 不可能事件 $\emptyset$ 的概率为 0,即 $P(\emptyset)=0$;
 - (3) 对于任一事件 A ,有 $0 \leq P(A) \leq 1$;
 - (4) 对于互斥事件 A 和 B ,有 $P(A \cup B) = P(A) + P(B)$;
 - (5) 对于对立事件 A 和 $\bar{A}$,有 $P(A) = 1 - P(\bar{A})$.



学法建议

XUEFA JIANYI

本课时学习重点是了解互斥事件的概念以及互斥事件的概率加法公式;难点是理解互斥事件与对立事件的联系与区别;关键之处在于当直接求某一事件的概率较为复杂时,可转而求其对立事件的概率.

学习中要注意以下几点:

1. 根据概率的定义可推出随机事件概率的取值范围.
2. 互斥事件即两个事件至多一个发生,它们可能都不发生,但不可能都发生. 如果 A , B 是互斥事件,那么 A 与 B 和的概率等于事件 A 与事件 B 的概率之和.
3. 在一次试验中,两个互斥事件有可能都不发生,如果两个互斥事件在一次试验中必然有一个发生,那么这样的互斥事件就是对立事件. 相应地,对立事件的概率之和等于 1.



例题导练

例1 一名射手进行一次射击,试判断下列各组中的两个事件哪些是互斥事件,哪些是对立事件.

- (1) 事件 A “命中7环”与事件 B “命中10环”;
- (2) 事件 A “命中7环”与事件 C “命中环数小于6环”;
- (3) 事件 B “命中10环”与事件 C “命中环数小于6环”;
- (4) 事件 D “命中6、7、8、9、10”与事件 C “命中环数小于6环”.

【思路分析】 判断两事件是否是互斥事件,关键看在同一试验中是否同时发生.

解 (1) 事件 A 与事件 B 不可能同时发生,是互斥事件;
 (2) 事件 A 与事件 C 不可能同时发生,是互斥事件;
 (3) 事件 B 与事件 C 不可能同时发生,是互斥事件;
 (4) 事件 D 与事件 C 不可能同时发生,并且必有一个发生,既是互斥事件,又是对立事件.

【思维诊断】 两个互斥事件中至多有一个发生,它们可能都不发生,但不可能都发生.

例2 某射手在一次射击中命中9环的概率是0.28,命中8环的概率是0.19,命中环数小于8环的概率是0.29,计算该射手在一次射击中命中9环或10环的概率.

【思路分析】 该射手命中9环、8环和命中环数小于8环彼此互斥,因此可用概率加法公式求解.

解 记该射手在一次射击中命中10环或9环为事件 A ,命中10环、命中9环、命中8环和命中环数小于8环分别记为 A_1, A_2, A_3, A_4 .

因为 A_2, A_3, A_4 彼此互斥,所以

$$P(A_2 \cup A_3 \cup A_4) = P(A_2) + P(A_3) + P(A_4) = 0.28 + 0.19 + 0.29 = 0.76.$$

又因为 A_1 与 $A_2 \cup A_3 \cup A_4$ 为对立事件,所以

$$P(A_1) = 1 - P(A_2 \cup A_3 \cup A_4) = 1 - 0.76 = 0.24.$$

因为 A_1 与 A_2 互斥,且 $A = A_1 \cup A_2$,所以

$$P(A) = P(A_1 \cup A_2) = P(A_1) + P(A_2) = 0.24 + 0.28 = 0.52.$$

【思维诊断】 要注意厘清各个事件之间的关系,分清哪些事件是互斥的,哪些不是互斥的;在将一个事件拆分为几个互斥事件时,要做到不重不漏.

例3 向三个相邻的军火库扔一颗炸弹,命中第一个军火库的概率为0.025,命中其余两个军火库的概率均为0.1,并且若其中一个军火库发生爆炸,另外两个军火库也会发

生爆炸,求三个军火库发生爆炸的概率.

【思路分析】三个军火库要发生爆炸,只要炸弹命中一个军火库即可,因为只扔了一颗炸弹,故命中第一、第二、第三军火库的事件是彼此互斥的.

【解】以 A, B, C 分别表示炸弹命中第一、第二、第三军火库这三个事件,则 $P(A) = 0.025, P(B) = P(C) = 0.1$.

以 D 表示“军火库爆炸”这一事件,则有 $D = A \cup B \cup C$, 其中 A, B, C 是互斥事件.

因为只扔了一个炸弹,不会同时命中两个及两个以上军火库,所以

$$P(D) = P(A) + P(B) + P(C) = 0.025 + 0.1 + 0.1 = 0.225.$$

【思维诊断】对于一个较复杂的事件,一般将其分解成几个简单的事件,当这些事件彼此互斥时,原事件的概率就是这些事件的概率之和.

达标练习 DABIAO LIANXI

A 组

1. 若 100 件产品中有 5 件次品,从中任意取 2 件,判断下列每对事件是不是互斥事件.

(1) $A = \{\text{恰有 1 件是次品}\}, B = \{\text{恰有 2 件是次品}\};$

(2) $C = \{\text{至少有 1 件次品}\}, D = \{\text{全是正品}\};$

(3) $E = \{\text{至少有 1 件正品}\}, F = \{\text{至少有 1 件次品}\}.$

2. 掷一颗骰子,写出下列事件的对立事件.

(1) $A = \{\text{出现 5 点}\};$

(2) $B = \{\text{出现偶数点}\};$

(3) $C = \{\text{出现的点数大于 4}\}.$



A. 对立事件

B. 不可能事件

C. 互斥事件

D. 必然事件

1. 给出下列命题:

① 对立事件一定是互斥事件;

② 若 A, B 为两个随机事件, 则 $P(A \cup B) = P(A) + P(B)$;

③ 若事件 A, B, C 彼此互斥, 则 $P(A) + P(B) + P(C) = 1$.

其中正确命题的个数是().

A. 1

B. 2

C. 3

D. 0

2. 射箭比赛中,一名选手命中 10 环、9 环的概率分别为 0.3, 0.5, 他命中环数超过 8 环的概率有多大?

3. 如果从取出大小王的 52 张扑克牌中随机抽取一张,那么事件 A “取到红桃”的概率是 0.25,事件 B “取到方块”的概率是 0.25,求:

(1) 事件 C“取到红色牌”的概率;

(2) 事件 D “取到黑色牌”的概率.

4. 一架飞机向目标投弹,击毁目标的概率为 0.2,目标未受损的概率为 0.4,求目标受损但未被完全击毁的概率.

11.4

抽样方法

11.4.1 总体与样本

 课标要求 KEBIAO YAOQIU

1. 联系实际问题,了解随机抽样的必要性和重要性.
2. 理解总体、个体及样本的意义,能举例说明普查和抽查的优缺点.

 知识导读 ZHISHI DAO DU

1. 总体:统计学中,所考察对象某一项指标值(所要统计的目标)的全体.
2. 个体:构成总体的每一个指标值.
3. 样本:从总体中抽出若干个体所组成的集合.
4. 样本容量:样本中包含的个体数量.
5. 抽样:从总体中抽取一部分个体作为样本的过程.
6. 统计的基本思想:用样本估计总体.

 学法建议 XUEFA JIANYI

本课时学习重点是了解随机抽样的必要性和重要性,理解总体、个体及样本的意义;难点是认识样本与样本容量的区别;关键是分清获取数据采取的方法是普查还是抽查.

学习中要注意以下几点:

1. 认识抽样的必要性. 同学们在初中已经学习过一些统计知识,了解到收集数据是统计的首要任务. 当研究一个对象的某项数量指标时,通过普查可得到该项指标的全部数据,但是研究对象的容量很大或检测过程具有一定的破坏性时,通常采用抽样调查的方法.
2. 认识抽样的目的性,即从调查对象中按照一定的方法抽取一部分进行调查或观察以获取数据,并以此对调查对象的某项数量指标作出推断.
3. 正确理解总体、个体、样本、样本容量等的含义.
4. 总体有三个性质:(1) 大量性,指总体中个体的数目巨大;(2) 同质性,指总体中的每一个个体彼此之间有相同的性质;(3) 变异性,总体中的各个个体之间也是有差异的.

5. 样本与抽样的关系,可以理解为抽样是过程,样本是抽样的结果.样本需要尽量能全面地反映出总体的特性.

例题导练

例1 为了估计一批零件的平均长度,从这批零件中抽取10个,测得它们的长度(单位:cm)如下.

22.36 22.35 22.33 22.35 22.37
22.34 22.38 22.36 22.32 22.35

在此统计活动中,总体、个体、样本分别是什么?样本容量是多少?

【思路分析】首先要明确该统计活动的目的——估计这批零件的平均长度;其次要把握考察的指标值——这批零件的长度.

解 在此统计活动中,总体是这批零件的长度,个体是每个零件的长度,样本是抽取的10个零件的长度.样本容量是10.

【思维诊断】注意该统计活动的研究对象并不是零件,而是零件的长度.

例2 为了解一批电视机的寿命,从中抽取100台电视机进行试验,这个问题中的样本是().

- A. 这批电视机的寿命 B. 抽取的100台电视机
C. 100 D. 抽取的100台电视机的寿命

【思路分析】首先要分清考察的对象是电视机还是电视机的寿命,其次要把握样本是总体中所抽取的一部分个体.

解 样本是抽取的100台电视机的寿命,故选D.

【思维诊断】总体、个体与样本的考察对象是相同的,所不同的是范围的大小.

例3 某火车站为了解某月每天上午乘车人数,抽查了其中10天每天上午的乘车人数,所抽查的这10天每天上午的乘车人数是这个问题的().

- A. 总体 B. 个体 C. 一个样本 D. 样本容量

【思路分析】在区分总体、个体、样本、样本容量这四个概念时,首先找出考察的对象,需要注意考察对象是表示事物某一特征的数据,而非考察的事物,从而找出总体、个体,再根据被收集数据的这一部分对象找出样本,最后根据样本确定样本容量.

解 本题考察的对象是此火车站该月每天上午的乘车人数,故总体是此火车站该月每天上午的乘车人数,个体是此火车站该月各天上午的乘车人数,样本是所抽查的10天此火车站每天上午的乘车人数,样本容量是10. 故选C.

【思维诊断】解题时要分清具体问题中的总体、个体与样本,关键是明确考察的对象.样本容量是样本中包含的个体的数目,不能带单位.

例 4 某校为了解学生的身体状况,在学校操场对参加体育锻炼的学生进行调查.你认为这样的调查结果会怎样?

【思路分析】 要想从样本出发,对总体作出基本合理的判断(由于样本是随机的,误差是不可避免的),就要求样本能够很好地代表总体.

【解】 这项调查的总体应为该校全体学生的身体状况.该调查小组选择的受访者为在操场参加体育锻炼的学生,而这部分学生的身体状况是较好的,样本并不能很好地代表总体,所以这样的调查结果是片面的.

【思维诊断】 调查结果是根据样本数据进行推断的,因此要获取正确的调查结果,必须使样本具有代表性.



达标练习



DABIAO LIANXI

A 组

1. 下列调查是普查还是抽查?如果是抽样调查,请指出总体、个体、样本和样本容量.

(1) 调查你所在班级每名同学每年去图书馆的次数;

(2) 从一批电视机中抽取 10 台,调查这批电视机的使用寿命.

2. 为了解一批灯泡的使用寿命,从中随机抽取 20 只灯泡进行试验,在这个问题中,总体是_____,个体是_____样本是_____,样本容量是_____.

B 组

1. 检查一箱(共 1 250 件)包装食品的质量,按 2% 的比例抽查一部分,在这个问题中,总体是_____,个体是_____样本是_____,样本容量是_____.

2. 为了解某市老年人的健康状况,某天早晨对在公园晨练的 50 位老人进行了调查.你认为这样的抽样调查合适吗?请简要说明理由.

3. 要了解某五年制高职校的学生一周内用于上网的时间,对某网吧内正在上网的 10 名该校学生进行了调查.你认为这样的调查合适吗?请简要说明理由.



11.4.2 简单随机抽样



课标要求



KEBIAO YAOQIU

1. 掌握用抽签法、随机数表法进行抽样的步骤.
2. 会用简单随机抽样方法从总体中抽取样本.



知识导读



ZHISHI DAO DU

1. 一般地,从元素个数为 N 的总体中不放回地抽取容量为 n 的样本,如果每一次抽取时各个体有相同的可能性被抽到,就称这种抽样方法为简单随机抽样.

2. 常用的简单随机抽样方法有抽签法、随机数表法和计算器(或计算机)产生随机数法.

3. 抽签法的步骤:编号→制签→搅匀→抽签→取样.

第一步,编号.对总体中的每一个个体编号 $1, 2, \dots, N$.

第二步,制签.制作 N 个形状、大小相同的,写有 1 至 N 编号的号签.

第三步,搅匀.将号签放在同一箱中,并搅拌均匀.

第四步,抽签.从箱中每次抽出一个号签,连续抽取 n 次.

第五步,取样.将总体中与号签的编号一致的 n 个样本抽出.

4. 随机数表法的步骤:编号→选数→定向→取数→抽样.

第一步,编号.对总体中的 N 个个体一一编号(每个号码位数应一致).如 N 是三位数,即按 $001, 002, \dots, N$ 进行编号.

第二步,选数.在随机数表中任选一个数字作为开始.

第三步,定向.一般由左向右.

第四步,取数.从选定的数字开始按选定的方向读下去,若得到的数在编号中则取出,若得到的数不在编号中或前面已经取出,则跳过.如此继续下去,直到取够样本容量为止.

第五步,抽样.根据选定的上述 n 个号码,抽取对应的 n 个个体.



学法建议



XUEFA JIANYI

本课时学习重点是了解随机抽样的意义,难点是理解简单随机抽样的本质特征,关键是理解样本含有与总体基本相同的信息.

学习中要注意以下几点:

1. 从总体中抽取样本的目的是认识总体.从样本数据得出总体的有关信息,这就是统计推断.它的一个组成部分是估计.

2. 如何抽取样本直接关系到对总体估计的准确程度. 例如, 品尝一锅汤的味道, 需要把锅里的汤充分搅拌均匀, 即样本含有与总体基本相同的信息.

3. 了解随机抽样的公平性. 所谓随机抽样的公平性, 是指在随机抽样的过程中, 每个个体被抽到的概率都相等, 即若从含有 N 个个体的总体中抽取一个容量为 n 的样本, 在整个抽样过程中每个个体被抽到的概率都是 $\frac{n}{N}$.

4. 把握简单随机抽样的特点:

(1) 它要求被抽取样本的总体的个体数目有限, 这样便于通过随机抽取的样本对总体进行分析.

(2) 它是从总体中逐个地进行抽取, 这样便于在抽样实践中进行操作.

(3) 它是一种不放回抽样. 由于抽样实践中多采用不放回抽样, 使其具有广泛的实用性, 而且由于所抽取的样本没有被重复抽取的个体, 便于进行有关的分析和计算.

(4) 它是一种等概率抽样, 不仅从总体中抽取一个个体时, 各个个体被抽取到的概率相等, 而且在整个抽样过程中, 各个个体被抽取到的概率也相等, 从而保证抽样方法的公平性.

例题导练

例 1 某校有 1 200 名学生, 为了调查学生双休日上网情况, 该校打算抽取 50 名学生进行访谈, 以获取一个容量为 50 的样本, 此样本若采用简单随机抽样将如何获得?

【思路分析】 尽管此问题的总体中, 个体数不算“较少”, 但按简单随机抽样, 其操作过程却是保证等可能的.

解 方法 1(抽签法): 首先, 把该校学生都编上号码: 0001, 0002, 0003, $\dots$, 1200. 制作 1 200 个形状、大小相同的号签(号签可以用小球、卡片、纸条等制作), 然后将这些号签放在同一个箱子里并搅拌均匀. 抽签时, 每次从中抽出 1 个号签, 连续抽取 50 次就得到 50 名学生的编号, 通过访谈对应编号的学生, 即可获得一个容量为 50 的样本.

方法 2(随机数表法): 首先, 把该校学生都编上号码: 0001, 0002, 0003, $\dots$, 1200. 在随机数表上随机选定一个起始位置, 假如起始位置是表中的第 5 行第 9 列的数字 6, 从 6 开始向右连续取数字, 以 4 个数为一组, 所得数字如下:

6438, 5482, 4622, 3162, 4309, 9006, 1844, 3253, 2383,
0130, 3046, 1943, 6248, 3469, 0253, 7887, 3239, 7371,
3445, 9493, 4977, 2261, 8442, $\dots$

所取出的 4 位数字如果小于或等于 1200, 则与之对应的学生即为被抽取的学生; 超过 1200 的数字可以除以 1200 取其余数, 如果遇到相同的号码, 则只保留第一次取出的数字, 其余的舍去, 经过这样的处理, 被抽取的学生所对应的号码分别是:

0438,0682,1022,0762,0709,0606,0644,0853,1183,
0130,0646,0743,0248,1069,0253,0687,0839,0171,
1045,1093,0177,1061,0042,...

直到取够 50 个号码为止.

通过访谈编号对应的学生,即可获得一个容量为 50 的样本.

【思维诊断】(1) 规范且不带主观倾向的随机抽样,才能保证公平性、客观性、准确性和可信性,因此抽样的过程也反映科学的工作态度和求实的工作作风.(2) 抽签法的优点是简单易行,缺点是当总体的容量非常大时,费时、费力,如果号签搅拌得不均匀,也会导致抽样不准确.(3) 总体个数较多时随机数表法比抽签法更适用.产生随机数的方法很多,利用计算器或数学软件可以方便地产生随机数.

例 2 总体由编号为 01,02,...,19,20 的 20 个个体组成.利用下面的随机数表选取 5 个个体,选取方法为从随机数表第 1 行的第 5 列和第 6 列数字开始由左到右依次选取两个数字,则选出来的第 5 个个体的编号为().

表 11-5

7816	6572	0802	6314	0702	4369	9728	0198
3 204	9 234	4 935	8 200	3 623	4 869	6 938	7 481

A. 08

B. 07

C. 02

D. 01

【思路分析】从随机数表第 1 行的第 5 列和第 6 列数字开始由左到右依次选取两个数字,得到 65,72,08,02,63,14,07,02,43,69,97,28,01,98,...,其中 08,02,14,07,01 符合条件,故可得结论.

解 从随机数表第 1 行的第 5 列和第 6 列数字开始由左到右依次选取两个数字,第一个数为 65,不符合条件,第二个数为 72,不符合条件,第三个数为 08,符合条件,以此类推,得到符合条件的数依次为 08,02,14,07,01,故第 5 个数为 01,选 D.

【思维诊断】在随机数表中每个数出现在每个位置的概率是一样的,所以每个数被抽到的概率是一样的.

达标练习

A 组

- 关于简单随机抽样的特点,有以下几种说法,其中不正确的是().
A. 简单随机抽样要求总体的个数有限
B. 简单随机抽样是从总体中逐个抽取
C. 简单随机抽样一般情况下是不放回的抽取

- D. 简单随机抽样中,每个个体被抽到的可能性与抽取的顺序有关
2. 我国古代数学名著《九章算术》有如下“米谷粒分”问题:粮仓开仓收粮,有人送来米 1 534 石,验得米内夹谷,抽样取米一把,数得 254 粒内夹谷 28 粒,则这批米内夹谷约为().
- A. 134 石 B. 169 石 C. 338 石 D. 1 365 石
3. 某职业学校一年级新生有 240 名,为了解学生的体育运动爱好项目,试写出用简单随机抽样从中抽取一个容量为 40 的样本的步骤.

B 组

1. 假设一个总体有 5 个个体,分别记为 a, b, c, d, e ,采用逐个不放回抽取的方法,从中抽取一个容量为 2 的样本,这样的样本共有多少种可能? 写出全部可能的样本.
2. 实验室的笼子里共有 100 只小白鼠,现要从中抽取 10 只做实验用,下列两种情况是否属于简单随机抽样? 请说明理由.
- (1) 每次不经过任何挑选地抓一只,抓满 10 只为止;
- (2) 将笼中的 100 只小白鼠按 1~100 编号,任意选出编号范围内的 10 个不重复的数,把相应编号的小白鼠作为实验用的小白鼠.



11.4.3 系统抽样、分层抽样



课标要求



KEBIAO YAOQIU

1. 了解分层抽样和系统抽样方法.
2. 了解各种抽样方法的适用范围,能根据具体情况选择适当的抽样方法.



知识导读



ZHISHI DAOYU

1. 一般地,当总体所含的个体较多时,可将总体分成均衡的几个部分,然后按照预先制定的规则,从每个部分中抽取一个个体,得到所需的样本,这样的抽样方法称为系统抽样或机械抽样.

2. 从容量为 N 的总体中,用系统抽样抽取容量为 n 的样本,其步骤为:

第一步,编号. 将总体的 N 个个体编号.

第二步,分段. 将整个编号按顺序平均分成 n 段;当 $\frac{N}{n}$ 的余数为 r ($0 < r < n$) 时,从总体中剔除 r 个个体,并将剩下的 $N-r$ 个个体重新编号.

第三步,定规. 第一段中随机确定起始的个体编号 m ,然后按照逐次加 k 的原则确定后续抽取的编号. 当 $\frac{N}{n}$ 为整数时,取 $k = \frac{N}{n}$;当 $\frac{N}{n}$ 的余数为 r 时,取 $k = \frac{N-r}{n}$.

第四步,抽样. 将编号为 $m, m+k, m+2k, \dots, m+(n-1)k$ 的个体抽出.

3. 一般地,当总体由差异明显的几个部分组成时,为了使抽样更客观地反映总体情况时,可将总体中各个个体按差异特征分为层次分明且互不重叠的几部分,然后按各部分在总体中所占比例进行简单随机抽样,这种抽样方法称为分层抽样,其中所分成的各个部分称为“层”.

4. 从容量为 N 的总体中用分层抽样抽取容量为 n 的样本,其主要步骤为:

- (1) 按某种特征将总体分成若干层(部分);
- (2) 计算各层的个体数与总体个体数的比值;
- (3) 计算各层应抽取的样本容量;
- (4) 各层分别按简单随机抽样的方法抽取样本,综合各层抽样组成样本.

5. 三类抽样方法的适用范围与优点

表 11-6

类型	适用范围	优点
简单随机抽样	总体中个体个数较少	简单易行
系统抽样	总体中个体个数较多	样本分布均匀
分层抽样	总体由差异明显的几部分组成	样本有较强的代表性

学法建议

本课时学习重点是了解分层抽样和系统抽样的方法及适用范围;难点是对总体中个体的数量与组成特征进行分析,进而选择适当的抽样方法;关键是既保证抽样的随机性,又保证样本的代表性.

学习中要注意以下几点:

- 1. 当总体所含的个体较多时,将总体分成均衡的几个部分,采用系统抽样.
- 2. 当总体由差异明显的几个部分组成时,根据差异特征,采用分层抽样.
- 3. 把握系统抽样、分层抽样的方法及步骤.

例题导练

例 1 某职业学校一年级共有 20 个班级,每班有 50 名学生. 为了解一年级学生语文学习的情况,从这 1 000 人中抽取 100 人进行质量抽测,如何采用系统抽样的方法抽取?

【思路分析】 采用系统抽样方法完成从 1 000 人中抽取 100 人,需要将 1 000 人分成均衡的 100 个部分,然后按照预先制定的规则,从每个部分中抽取一人.

解 第一步,将 1 000 名学生编号为 000,001,⋯,999;

第二步,将所有编号分成 100 段;

第三步,在第一段 10 个编号中随机抽取起始号 m ,确定 k 的值为 $\frac{1\,000}{100}=10$;

第四步,将编号为 $m, m+10, m+20, \cdots, m+990$ 的学生抽出.

例 2 某单位 500 名职工中,血型为 O、A、B、AB 型的人数分别为 200 人、125 人、125 人、50 人,为了解职工健康教育情况,要抽取 40 人进行访谈,应如何抽取? 四种血型应分别抽取多少人?

【思路分析】 各部分之间有差异,是分层抽样的依据.

解 方法 1:第一步,按血型分层,将 500 名职工分成 4 层.

第二步,计算各层的人数与总人数的比.

$$\text{O 型占 } \frac{200}{500} = \frac{2}{5},$$

$$\text{A 型占 } \frac{125}{500} = \frac{1}{4},$$

$$\text{B 型占 } \frac{125}{500} = \frac{1}{4},$$

$$\text{AB 型占 } \frac{50}{500} = \frac{1}{10}.$$

第三步,计算各层应抽取的人数.

$$\text{O 型应抽 } 40 \times \frac{2}{5} = 16 \text{ 人},$$

$$\text{A 型应抽 } 40 \times \frac{1}{4} = 10 \text{ 人},$$

$$\text{B 型应抽 } 40 \times \frac{1}{4} = 10 \text{ 人},$$

$$\text{AB 型应抽 } 40 \times \frac{1}{10} = 4 \text{ 人}.$$

方法 2: 第一步, 分层. 按血型分层, 将 500 名职工分成 4 层.

第二步, 确定抽取比例. 抽取人数 $n=40$, 总人数 $N=500$, 抽取比例 $= \frac{40}{500} = \frac{2}{25}$.

第三步, 按比例确定每层应抽取人数.

$$\text{O 型应抽 } 200 \times \frac{2}{25} = 16 \text{ 人},$$

$$\text{A 型应抽 } 125 \times \frac{2}{25} = 10 \text{ 人},$$

$$\text{B 型应抽 } 125 \times \frac{2}{25} = 10 \text{ 人},$$

$$\text{AB 型应抽 } 50 \times \frac{2}{25} = 4 \text{ 人}.$$

达标练习 DABIAO LIANXI

A 组

1. 要从 1 005 名学生中按系统抽样抽取 50 人, 试叙述抽取的步骤.

2. 用系统抽样法从 160 名学生中抽取 20 人,将 160 名学生随机地按 1~160 编号,按编号顺序平均分成 20 组(1~8 号,9~16 号, $\dots$,153~160 号),若第 16 组抽出的号码为 126,则第 1 组中用抽签的方法确定的号码是().

- A. 5 B. 6 C. 7 D. 8

3. 某林场有树苗 30 000 棵,其中松树苗为 4 000 棵.为调查树苗的生长情况,采用分层抽样的方法抽取 150 棵树苗进行调查,则松树苗应抽取的数量为().

- A. 30 棵 B. 25 棵 C. 20 棵 D. 15 棵

B 组

1. 某高职院校为了解学生对参加某项社会实践活动的意向,拟采用分层抽样的方法,从该校五个年级的学生中抽取 450 人进行调查.已知该校一年级、二年级、三年级、四年级、五年级的学生人数之比为 5:4:4:3:2,则应从一年级学生中抽取_____名学生.

2. 一个总体中共有 100 个个体,随机编号为 0,1,2,99,依编号顺序平均分成 10 个小组,组号依次为 1,2,3, $\dots$,10.现用系统抽样的方法抽取一个容量为 10 的样本,规定如果在第 1 组随机抽取的号码为 m ,那么在第 k 组中抽取的号码个位数字与 $m+k$ 的个位数字相同.若 $m=6$,则在第 7 组中抽取的号码是_____.

3. 某工厂共有职工 3 000 人,其中中年、青年、老年职工人数之比为 5:3:2.若要用分层抽样的方法从所有职工中抽取 400 人进行某项调查,则中年、青年、老年职工应分别抽取多少人?



11.5 统计图表

11.5.1 频率分布表



课标要求

KEBIAO YAOQIU

1. 了解频率分布表是反映总体频率分布的表格.
2. 掌握列频率分布表的方法.



知识导读

ZHISHI DAODU

1. 当总体很大或不便获得时,可以用样本的频率分布估计总体的频率分布.我们把反映总体频率分布的表格称为频率分布表.

2. 列出样本的频率分布表的一般步骤为:

第一步,计算极差.极差是一组数据的最大值与最小值的差.

第二步,决定组距和组数.组距是分成的区间的长度,是极差除以组数所得的商(取整数).

第三步,确定各组分点.确定各组分点的原则是,既要把全部数据包括在内,又要使每个数据在一个确定的组内.一般取比数据多一位小数,且把第一组起点减小一点.

第四步,列频率分布表.一般分为三列(分组、频数、频率),最后一行是合计,频数合计值等于样本容量,频率合计值为 1.



学法建议

XUEFA JIANYI

本课时学习重点是了解频率分布表的意义,难点是理解列频率分布表的步骤,关键是通过表格形式对收集到的数据分布情况进行整理.

学习中要注意以下几点:

1. 通过随机抽样完成样本数据收集后,要通过样本数据估计总体分布情况,因此需要把收集到的数据分布情况整理出来.列频率分布表是对数据进行加工整理的一种基本方法.

2. 应把握列频率分布表每一步骤的内涵.

例题导练

例 1 某校拟调研学生的身高与运动量之间的关系,从男生中随机抽取 100 名学生的身高数据,得到如下频率分布表.

表 11-7

组号	分组	频数	频率
第 1 组	$[160,165)$	10	0.10
第 2 组	$[165,170)$	①	0.15
第 3 组	$[170,175)$	30	②
第 4 组	$[175,180)$	25	0.25
第 5 组	$[180,185)$	20	0.20
合计		100	1.0

(1) 求频率分布表中①、②位置相应的数据;

(2) 为了对比研究学生运动量与身高的关系,学校计划采用分层抽样的方法从第 2、5 组中随机抽取 7 名学生进行跟踪调研,求第 2、5 组每组抽取的学生数;

(3) 在(2)的前提下,学校决定从这 7 名学生中随机抽取 2 名学生接受调研访谈,求至少有 1 名学生来自第 5 组的概率.

【思路分析】 (1) 根据频率、频数与样本容量的关系,求出①、②的数值;

(2) 先求出第 2、5 组的人数,再根据分层抽样原理,求出第 2、5 组应抽取的人数;

(3) 用列举法求出基本事件数,计算对应的概率即可.

解 (1) 根据题意得,

位置①相应的数据即为第 2 组的频数,等于 $100 \times 0.15 = 15$;

位置②相应的数据即为第 3 组的频率,等于 $30 \div 100 = 0.30$.

(2) 第 2 组有 15 人,第 5 组有 20 人,按分层抽样方法从第 2、5 组中随机抽取 7 名学生,第 2 组中应抽取 $7 \times 15 \div (15 + 20) = 3$ 人,第 5 组中应抽取 $7 - 3 = 4$ 人.

(3) 将第 2 组的学生记为 a, b, c ,将第 5 组的学生记为 $1, 2, 3, 4$.

从这 7 名学生中随机抽取 2 名学生,基本事件有 $ab, ac, a1, a2, a3, a4, bc, b1, b2, b3, b4, c1, c2, c3, c4, 12, 13, 14, 23, 24, 34$, 共 21 个.

至少有 1 名学生来自第 5 组的基本事件数有 $a1, a2, a3, a4, b1, b2, b3, b4, c1, c2, c3, c4, 12, 13, 14, 23, 24, 34$, 共 18 个,对应的概率为 $P = 18 \div 21 = \frac{6}{7}$.

例 2 为了对某课题进行研究,用分层抽样方法从 A、B、C 三所高校的相关人员中抽取若干人组成研究小组,有关数据见下表,则 $x = \underline{\hspace{2cm}}$, $y = \underline{\hspace{2cm}}$.

表 11-8

高校	相关人数	抽取人数
A	18	x
B	36	2
C	54	y

若从高校 B、C 抽取的人中选 2 人作专题发言,则这 2 人都来自高校 C 的概率为

【思路分析】 依据“分层抽样从各部分抽取的个体数与该部分个体数的比相等,并等于样本容量与总体容量的比”,可求出 x 和 y . 从高校 B、C 抽取的人中选 2 人作专题发言,求这 2 人都来自高校 C 的概率是古典概型的概率问题.

解 由已知得 $\frac{x}{18} = \frac{2}{36} = \frac{y}{54}$, 解得 $x=1, y=3$.

把从高校 B 抽取的 2 人记为 a, b , 把从高校 C 抽取的 3 人记为 e, f, g .

从高校 B、C 抽取的 5 人中选 2 人作专题发言,基本事件有 $ab, ae, af, ag, be, bf, bg, ef, eg, fg$, 共 10 个.

“这 2 人都来自高校 C”包含的基本事件有 ef, eg, fg , 共 3 个, 所以这 2 人都来自高校 C 的概率 $P = \frac{3}{10}$.

达标练习

A 组

1. 从一堆苹果中任取 10 只,称得它们的质量(单位:克)如下:

125 120 122 105 130 114 116 95 120 134

样本数据落在 $[114.5, 124.5)$ 内的频率为_____.

2. 已知某样本的所有数据分别落在五个小组内,第一、二、三、五组数据的个数分别是 2, 8, 15, 5, 且第五组的频率为 0.1, 则这个样本的容量为_____, 第四组的频数和频率分别是_____和_____.

3. 已知一组数据有 40 个,把它分成六组,第一至四组的频数分别是 10, 5, 7, 6, 第五组的频率是 0.2, 故第六组的频率是_____.

4. 将一批数据分成 4 组,并列出频率分布表,其中第一组的频率是 0.23, 第二组与第四组的频率之和是 0.52, 第三组的频率是_____.

5. 某校数学复习考有 400 名学生参加,评分后校方将此 400 名学生依总分由高到低排序如下: 前 100 人为 A 组,次 100 人为 B 组,再次 100 人为 C 组,最后 100 人为 D 组. 校方进一步逐题分析学生答题情况,将各组第一题和第二题的答对率列表如下,则下列选项

中正确的是().

	A 组	B 组	C 组	D 组
第一题答对率	100%	80%	70%	20%
第二题答对率	100%	80%	30%	0%

- A. 第一题答错的同学,不可能属于 B 组
- B. 从第二题答错的同学中随机抽出一人,此人属于 B 组的概率大于 0.5
- C. 全体同学第一题的答对率比全体同学第二题的答对率低 15%
- D. 从 C 组同学中随机抽出一人,此人第一、二题都答对的概率不可能大于 0.3

B 组

1. 一个容量为 n 的样本分成若干组,已知某组的频数和频率分别是 10 和 0.4,则 $n=$ _____.
2. 将容量为 100 的样本数据分为 8 个组(如下表),则第 3 组的频率为().

组号	1	2	3	4	5	6	7	8
频数	10	13	x	14	15	13	12	9

- A. 0.03 B. 0.07 C. 0.14 D. 0.21

3. 某产品的直径长与标准值的差的绝对值不超过 1 mm 时,视为合格品,否则视为不合格品. 在近期的一次产品抽样检查中,检测人员随机抽取了某厂 5 000 件此种产品进行检测,结果发现其中有 50 件不合格品. 将这 50 件不合格品的直径长与标准值的差(单位: mm)分组,得到如下频率分布表.

分组	频数	频率
$[-3,-2)$		0.10
$[-2,-1)$	8	
$(1,2]$		0.50
$(2,3]$	10	
$(3,4]$		
合计	50	1.00

- (1) 将上面表格中缺少的数据填在相应位置;
- (2) 估计该厂生产的此种产品中,不合格品的直径长与标准值的差落在区间 $(1,3]$ 内的概率;
- (3) 现对该厂此种产品的某个批次进行检查,结果发现有 20 件不合格品,据此估算这批产品中合格品的件数.

11.5.2 频率分布直方图

课标要求

会根据频率分布表作出频率分布直方图,会用频率分布直方图对总体分布规律进行估计.

知识导读

1. 反映样本的频率分布规律的直方图称为频率分布直方图.

2. 绘制频率分布直方图的步骤

(1) 制作频率分布表,建立直角坐标系,以横轴表示数据,纵轴表示 $\frac{\text{频率}}{\text{组距}}$.

(2) 在横轴上标出各组的组分点.把横轴分成若干段,每一线段对应一个组的组距.

(3) 在坐标系中,分别将横轴上标出的点以相邻两点的线段为底,以 $\frac{\text{频率}}{\text{组距}}$ 为高作一系列矩形.

3. 频率分布直方图的性质

(1) 每个矩形的面积等于该组的频率;

(2) 所有矩形的面积之和等于 1.

学法建议

本课时学习重点是通过频率分布直方图对总体分布规律进行估计;难点是对频率分布直方图中以“ $\frac{\text{频率}}{\text{组距}}$ ”为纵轴的理解;关键是合理分组,具体问题具体分析.

学习中要注意以下几点:

1. 频率分布直方图是用矩形面积表示频率,从而直观表示样本的频率分布的趋势.
2. 样本的频率分布表与频率分布直方图都反映了总体分布的情况,两者互为补充.
3. 为了尽量减少估计错误,通常在条件许可的情况下适当增加样本容量,并且合理选用抽样方法以提高样本的代表性.

例题导练

例 1 某校高职三年级的学生进行体检,现将该年级男生的体重数据(单位:kg)进行

整理后分成五组,并绘制了频率分布直方图(如图 11-1 所示).根据一般标准,高职三年级男生的体重超过 65 kg 属于偏胖,低于 55 kg 属于偏瘦,已知图中从左到右第一、第三、第四、第五小组的频率分别为 0.25、0.20、0.10、0.05,第二小组的频数为 400,则该校高职三年级的男生总数和体重正常的频率分别为().

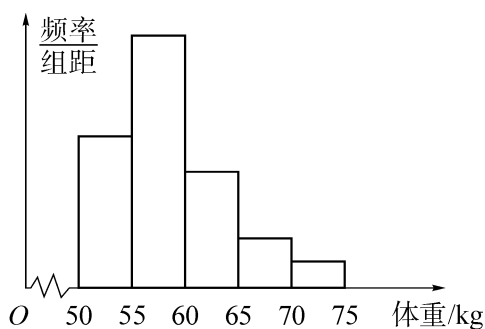


图 11-1

- A. 1 000,0.50 B. 800,0.50 C. 1 000,0.60 D. 800,0.60

【思路分析】(1) 已知第二小组的频数为 400,要求该校高职三年级的男生总数,只要先求出第二小组的频率,再用频数除以频率即可得到男生总数.

(2) 要求体重正常的频率,应先找到体重正常的小组.

解 第一、第三、第四、第五小组频率之和为 $0.25+0.20+0.10+0.05=0.60$,所以第二小组的频率为 $1-0.60=0.40$,男生总数 $=\frac{400}{0.4}=1\ 000$.

正常体重值落在第二、第三小组,两小组频率之和为 $0.40+0.20=0.60$. 综上所述,应选 C.

【思维诊断】要注意利用“频率分布直方图中,各组频率之和等于 1”这一性质.

例 2 在参加世界杯足球赛的 32 支球队中,随机抽取 20 名队员,调查得知他们的年龄分别为 25,21,23,25,27,29,25,28,30,29,26,24,25,27,26,22,24,25,26,28. 列频率分布表,并画出频率分布直方图,据此估计全体队员在哪个年龄段的人数最多,占总数的百分之几.

解 (1) 频率分布表如下.

表 11-9

分组	频数	频率
20.5~22.5	2	0.10
22.5~24.5	3	0.15
24.5~26.5	8	0.40
26.5~28.5	4	0.20
28.5~30.5	3	0.15
合计	20	1.00



(2) 频率分布直方图如下.

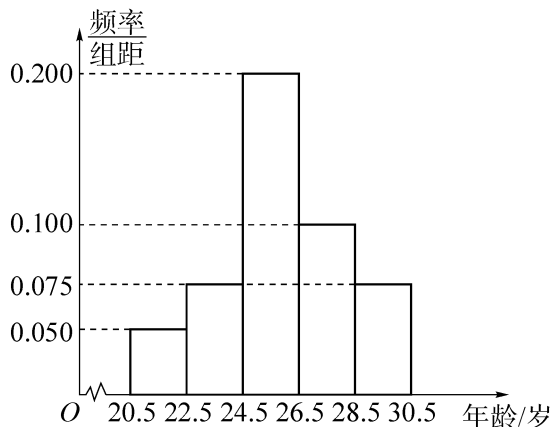


图 11-2

(3) 估计全体队员在 24.5~26.5 岁人数最多, 占总数的 40%.

【思维诊断】 频率分布表和频率分布直方图既可以揭示样本数据的分布类型, 也可以描述样本数据的分布特征, 便于进一步计算考察对象数值指标和统计分析.

例 3 从某小区抽取 100 户居民进行月用电量调查, 发现其用电量都在 50 至 350 度之间, 频率分布直方图如图所示.

(1) 直方图中 x 的值为_____;

(2) 在这些用户中, 用电量落在区间 $[100, 250)$ 内的户数为_____.

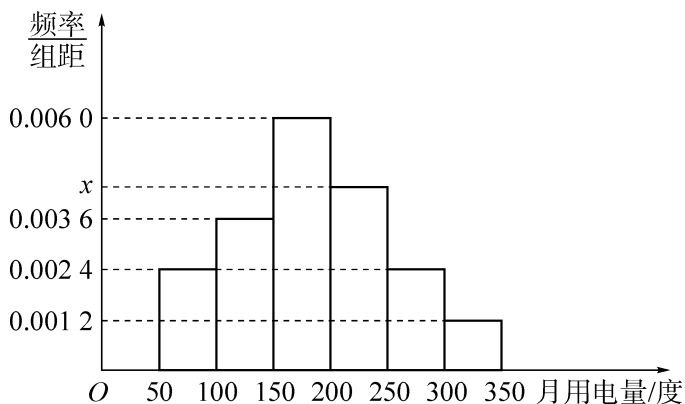


图 11-3

【思路分析】 (1) 根据“频率分布直方图中, 各组的频率之和为 1”, 可得一个关于 x 的方程, 解方程即可得到答案;

(2) 由已知中的频率分布直方图, 利用 $[100, 250)$ 内各小组的纵坐标 (矩形的高) 乘以组距得到 $[100, 250)$ 的频率, 用频率乘以样本容量即可求出频数.

解 (1) 依题意及频率分布直方图知,

$$0.0024 \times 50 + 0.0036 \times 50 + 0.0060 \times 50 + x \times 50 + 0.0024 \times 50 + 0.0012 \times 50 = 1,$$

解得 $x = 0.0044$.

(2) 样本数据落在 $[100,150)$ 内的频率为 $0.003\ 6 \times 50 = 0.18$,
 样本数据落在 $[150,200)$ 内的频率为 $0.006 \times 50 = 0.30$,
 样本数据落在 $[200,250)$ 内的频率为 $0.004\ 4 \times 50 = 0.22$,
 故在这些用户中,用电量落在区间 $[100,250)$ 内的户数为 $(0.18 + 0.30 + 0.22) \times 100 = 70$.

【思维诊断】 学习频率分布直方图的关键是学会画图、看图和用图.

达标练习

A 组

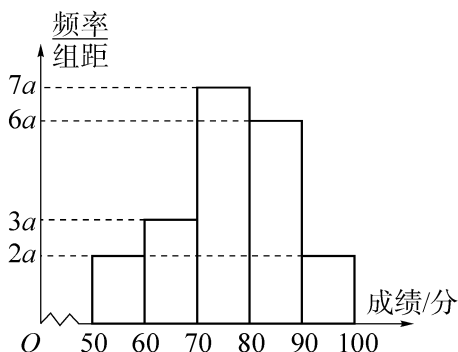
1. 随机抽取某工厂 20 名工人一天加工首饰的数量,获得数据如下(单位:件).

53	60	41	49	59	51	47	46	51	45
57	58	54	50	40	52	51	55	53	48

列出频率分布表,并画出频率分布直方图.

2. 20 名学生某次数学考试成绩(单位:分)的频数分布直方图如下.

- (1) 求频数直方图中 a 的值;
- (2) 分别求出成绩落在 $[50,60)$ 与 $[60,70)$ 内的学生人数;
- (3) 从成绩在 $[50,70)$ 的学生中选 2 人,求此 2 人的成绩都在 $[60,70)$ 内的概率.



(第 2 题)

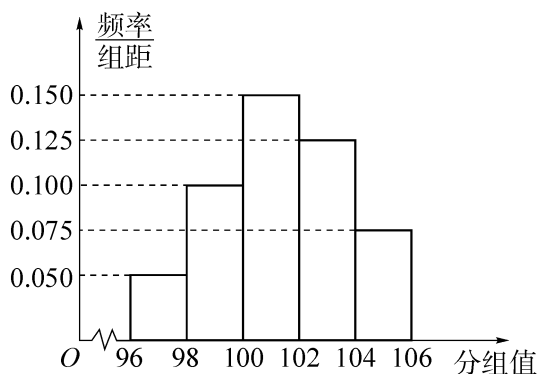
B 组

1. 在对 60 个数据进行整理的频率分布表中, 各组的频率之和与频数之和分别是 ().

- A. 60, 1 B. 60, 60 C. 1, 60 D. 1, 1

2. 某工厂对一批产品进行抽样检测, 下图是根据抽样检测后的产品净重数据(单位: 克)绘制的频率分布直方图, 其中产品净重小于 98 g 的共 60 个.

- (1) 求样本容量 n ;
(2) 求净重大于或等于 100 g 并且小于 104 g 的产品的个数.



(第 2 题)

11.5.3 统计图表的选用及数据分析

课标要求

了解频率分布表、频率分布直方图的适用条件及其特点, 能根据频率分布表、频率分布直方图所揭示样本数据的分布类型和分布特征, 对考察对象数值指标进行统计分析.

知识导读

1. 条形统计图能够清楚地表明各种数量指标的大小.
2. 扇形统计图能够比较清楚地反映出部分与部分、部分与整体之间的数量关系.
3. 折线统计图能够反映分析事物随时间或有序类别而变化的趋势.
4. 频率分布表能够显示各组频数分布的情况且易于显示各组之间频数的差别.
5. 频率分布直方图能够反映连续变量的频数分布.

学法建议

XUEFA JIANYI

本课时学习重点是根据频率分布表、频率分布直方图所揭示样本数据的分布类型和分布特征,对考察对象数值指标进行数据统计分析;难点是了解频率分布表、频率分布直方图的适用条件及其特点;关键是对考察对象数值指标进行数据统计分析.

学习中要注意以下几点:

1. 用图表来表现数据、传递信息,很重要的一个环节是根据所要表达信息的需要(占比、对比、趋势、分布等),选择恰当的统计图表,进而对数据进行分析.
2. 根据样本频率分布直方图进行数据分析时,要理解每个矩形的高与面积的意义.
3. 频率分布表(图)可以揭示样本数据的分布类型,如对称分布或偏峰分布,也可以描述样本数据的分布特征,即集中趋势和离散趋势,便于发现某些特大和特小的数值,便于进一步计算考察对象数值指标和统计分析.

例题导练

LITI DAOLIAN

例 1 某市为了节约能源,拟出台“阶梯电价”政策,即制定住户月用电量的临界值 a ,若某住户某月用电量不超过 a 度,则按平价计费;若某月用电量超过 a 度,则超出部分按议价计费,未超出部分按平价计费. 为确定 a 的值,随机调查了该市 100 户的月用电量,工作人员已将 90 户的月用电量填在了下面的频率分布表中,最后 10 户的月用电量(单位:度)为 18,63,43,119,65,77,29,97,52,100.

表 11-10

组别	月用电量/度	频数统计	频数	频率
1	$[0,20)$	下		
2	$[20,40)$	正正一		
3	$[40,60)$	正正正正下		
4	$[60,80)$	正正正正正下		
5	$[80,100)$	正正正正下		
6	$[100,120)$	下		

- (1) 完成频率分布表并绘制频率分布直方图;
- (2) 根据已有信息,试估计全市住户的平均用电量(同一组数据用该区间的中点值作代表);
- (3) 若该市计划让全市 75% 的住户在“阶梯电价”政策出台前后缴纳的电费不变,试求临界值 a .



【思路分析】(1) 根据题目中的数据,补充完整频率分布表,绘制出频率分布直方图;
(2) 根据频率分布直方图,计算样本的平均数;
(3) 列出累计频率表,计算临界值 a 即可.

解 (1) 频率分布表如下:

表 11-11

组别	月用电量/度	频数统计	频数	频率
1	$[0,20)$	正	4	0.04
2	$[20,40)$	正正	12	0.12
3	$[40,60)$	正正正正	24	0.24
4	$[60,80)$	正正正正正正	30	0.30
5	$[80,100)$	正正正正正	25	0.25
6	$[100,120)$	正	5	0.05

频率分布直方图如下:

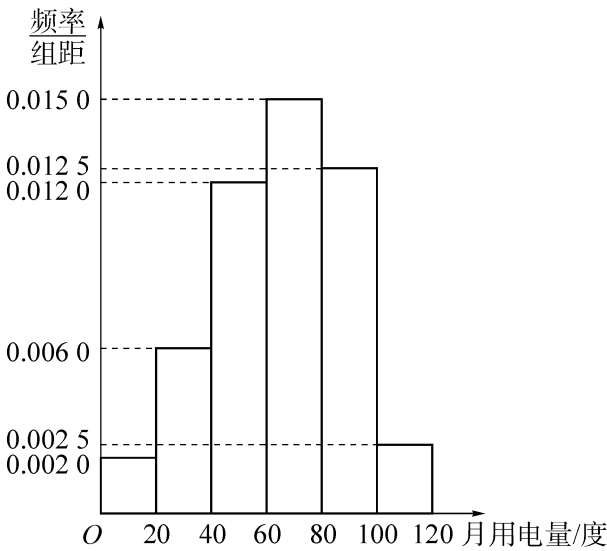


图 11-4

(2) 根据题意,估计全体住户的平均用电量为
 $10 \times 0.04 + 30 \times 0.12 + 50 \times 0.24 + 70 \times 0.30 + 90 \times 0.25 + 110 \times 0.05 = 65$ (度).
(3) 通过计算累计频率,可得下表:

表 11-12

区间	$[0,20)$	$[0,40)$	$[0,60)$	$[0,80)$	$[0,100)$	$[0,120)$
累计频数	4	16	40	70	95	100
累计频率	0.04	0.16	0.40	0.70	0.95	1.00

由此可知临界值 a 应在区间 $[80, 100)$ 内, 且频率分布直方图中, 在临界值 a 左侧的总面积(概率)为 0.75, 故有 $0.7 + (a - 80) \times 0.0125 = 0.75$, 解得 $a = 84$.

例 2 某校高职三年级财经专业有四个班. 在三年级数学调研测试后, 随机地从四个班抽取部分学生进行测试成绩统计, 各班被抽取的学生人数恰好成等差数列, 人数最少的班被抽取了 22 人. 抽取出来的所有学生测试成绩的频率分布直方图如图 11-5 所示, 其中 120~130(包括 120 分但不包括 130 分)分数段的人数为 5 人.

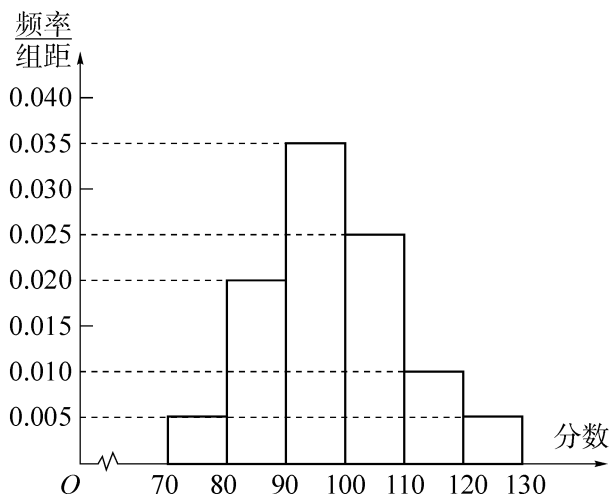


图 11-5

- (1) 各班被抽取的学生人数各为多少人?
- (2) 估计抽取出来的所有学生的平均成绩.
- (3) 在抽取的所有学生中任选一名学生, 其分数不低于 90 分的概率为多少?

【思路分析】 (1) 根据频率分布直方图, 求出抽取的学生总数, 由等差数列的知识, 求出各班被抽取的学生数;

(2) 同一组数据用该区间的中点值作代表, 可求出样本数据的平均数;

(3) 求出分数不小于 90(分) 的频率即可.

解 (1) 由频率分布直方图知, 抽取的学生总数为 $5 \div (0.005 \times 10) = 100$ (人).

因为各班被抽取的学生人数成等差数列, 设其公差为 d , 由 $4 \times 22 + 6d = 100$, 解得 $d = 2$, 所以各班被抽取的学生人数分别是 22 人、24 人、26 人、28 人.

(2) 样本数据的平均数是

$$75 \times 0.05 + 85 \times 0.20 + 95 \times 0.35 + 105 \times 0.25 + 115 \times 0.10 + 125 \times 0.05 = 98,$$

所以平均成绩为 98.

(3) 在抽取的学生中任选一名学生, 其分数不低于 90(分) 的概率为

$$0.35 + 0.25 + 0.10 + 0.05 = 0.75.$$

【思维诊断】 根据样本频率分布直方图进行数据分析, 要理解每个矩形的高与面积所表示的意义.



A 组

(40,50],5;(50,60],4;(60,70],2. 总体数据在区间(10,50]内的可能性为().

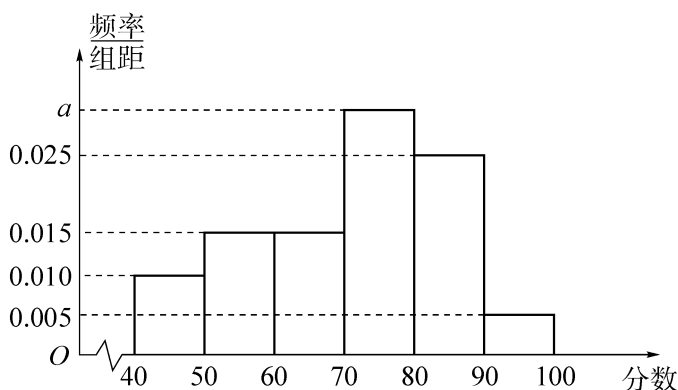
- A. 5% B. 25%
- C. 50% D. 70%

分布直方图.

- (1) 求 a 的值.

各抽取多少人的成绩?

的概率.



(第 2 题)

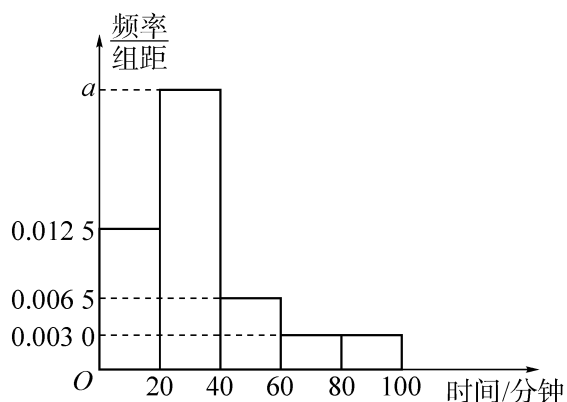
B 组

1. 下面是某职业学校 40 名学生一个月的零花钱数据(单位:元),请列出这些数据的频率分布表,画出频率分布直方图,并回答下列问题:

430	310	290	240	270	180	210	140	340	270
220	250	260	170	270	180	180	190	210	180
120	190	310	190	140	280	190	130	130	180
120	180	190	120	130	160	120	130	100	170

- (1) 每月零花钱不超过 150 元的学生占多少?
- (2) 90% 以上的学生每月零花钱在什么范围?
- (3) 零花钱数最集中的在哪个范围?

2. 某工厂随机抽取部分工人调查其上班路上所需时间(单位:分钟),并将所得数据绘成频率分布直方图(如图),上班路上所需时间的范围是 $[0, 100]$, 样本数据分组为 $[0, 20)$, $[20, 40)$, $[40, 60)$, $[60, 80)$, $[80, 100]$.



(第 2 题)

- (1) 求直方图中 a 的值;
- (2) 如果上班路上所需时间不少于 1 小时的工人可申请在工厂住宿,若招工 2 400 人,请估计所招工人中有多少名工人可以申请住宿.



11.6

样本均值与标准差

11.6.1 样本均值



课标要求

KEBIAO YAOQIU

1. 理解样本均值的意义,会根据公式计算一个样本的均值.
2. 会用样本均值估计总体数据的集中趋势.



知识导读

ZHISHI DAOYU

1. 称 $\frac{a_1 + a_2 + \cdots + a_n}{n}$ 为 n 个数据 $a_1, a_2, \cdots, a_n$ 的算术平均数或均值. 记为 $\bar{a} = \frac{a_1 + a_2 + \cdots + a_n}{n} = \frac{1}{n} \sum_{i=1}^n a_i$. 如果这 n 个数据是从总体中抽出来的一个样本,那么 $\bar{a}$ 就称为样本均值.

2. 若一组数据中有 $x_1, x_2, \cdots, x_n$ 共 n 种取值,且取值为 $x_1, x_2, \cdots, x_n$ 的频率分别为 $p_1, p_2, \cdots, p_n$, 则其均值为

$$\bar{x} = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n = \sum_{i=1}^n x_i p_i.$$

这个公式称为加权平均数计算公式.

3. 一般地,将样本 $a_1, a_2, \cdots, a_m$ (均值为 $\bar{x}_1$) 和样本 $b_1, b_2, \cdots, b_n$ (均值为 $\bar{x}_2$) 合并成一个新样本,则这个新样本的均值为

$$\begin{aligned} \frac{a_1 + a_2 + \cdots + a_m + b_1 + b_2 + \cdots + b_n}{m+n} &= \frac{m}{m+n} \cdot \frac{a_1 + a_2 + \cdots + a_m}{m} + \frac{n}{m+n} \cdot \frac{b_1 + b_2 + \cdots + b_n}{n} \\ &= \frac{m}{m+n} \bar{x}_1 + \frac{n}{m+n} \bar{x}_2. \end{aligned}$$

4. 分层抽样的均值: 设样本中不同层的均值和相应权重分别为 $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n$ 和 $\omega_1, \omega_2, \cdots, \omega_n$, 则这个样本的均值为 $\omega_1 \bar{x}_1 + \omega_2 \bar{x}_2 + \cdots + \omega_n \bar{x}_n = \sum_{i=1}^n \omega_i \bar{x}_i$.

学法建议

XUEFA JIANYI

本课时学习重点是理解样本均值的意义,能用样本均值描述总体水平;难点是理解均值刻画数据集中趋势的特征值;关键是注意到在有些实际问题中并不需要知道考察对象的整体情况,而只需要了解数据的集中趋势即可.

学习中要注意以下几点:

1. 本节内容实际上是对初中有关内容的复习和提高,重在从理论上说明其依据,培养理性思维.
2. 均值能对数据的集中趋势作出简要的描述.在有些实际问题中,并不需要知道考察对象的整体情况,而只需要了解它的某些数字特征即可.

例题导练

LITI DAOLIAN

例 1 一部电梯的最大载重量是 1 000 kg,现有 13 位“重量级”的乘客要搭乘电梯,已知其中 11 位男士的平均体重是 80 kg,2 位女士的平均体重是 70 kg,请问他们能否一起安全地搭乘这部电梯? 他们的平均体重是多少?

【思路分析】 (1) 他们能否一起安全地搭乘这架电梯,需要计算这 13 位乘客体重之和; (2) 计算这 13 位乘客的平均体重,可以通过他们体重之和除以总人数求解.

解 因为 11 位男士体重之和为 $80 \times 11 = 880$ (kg), 2 位女士体重之和为 $70 \times 2 = 140$ (kg), 所以 13 个人体重之和为 $880 + 140 = 1\,020$ (kg).

因为 13 个人体重之和超过电梯最大载重量,所以他们不能一起安全地搭乘这部电梯.

他们的平均体重 $= 1\,020 \div 13 \approx 78.46$ (kg), 或他们的平均体重 $= \frac{11}{11+2} \times 80 + \frac{2}{11+2} \times 70 \approx 78.46$ (kg).

例 2 某企业有 70% 的人年收入 10 万元, 25% 的人年收入 12 万元, 5% 的人年收入 18 万元, 求该企业的人均年收入.

【思路分析】 要求该企业的人均年收入, 可先计算该企业员工的年收入总和, 然后除以总人数. 该企业总人数未知, 可设为 m 人, 则总收入为

$$m \times 70\% \times 10 + m \times 25\% \times 12 + m \times 5\% \times 18 = m \times (70\% \times 10 + 25\% \times 12 + 5\% \times 18).$$

可见, 知道数据的频率, 可直接求均值.

解 该企业的人均年收入为 $70\% \times 10 + 25\% \times 12 + 5\% \times 18 = 10.9$ (万元).



例 3 某制造商 3 月生产了一批乒乓球,随机抽取 100 个进行检查,测得每个球的直径(单位 mm),将数据分组如下:

表 11-13

分组	频数	频率
$[39.95, 39.97)$	10	0.1
$[39.97, 39.99)$	20	0.2
$[39.99, 40.01)$	50	0.5
$[40.01, 40.03]$	20	0.2
合计	100	1.0

统计方法中,同一组数据经常用该组区间的中点值(例如,区间 $[39.99, 40.01)$ 的中点值是 40.00)作为代表.据此,估计这批乒乓球直径的均值(结果保留两位小数).

解 方法 1: 各组数据所在区间的中点值分别为 39.96, 39.98, 40.00, 40.02.

这批乒乓球直径的总和为

$$39.96 \times 10 + 39.98 \times 20 + 40.00 \times 50 + 40.02 \times 20 = 3\,999.6(\text{mm}),$$

所以这批乒乓球直径的均值为 $3\,999.6 \div 100 \approx 40.00(\text{mm})$.

方法 2: 各组数据所在区间的中点值分别为 39.96, 39.98, 40.00, 40.02, 对应的频率分别为 0.1, 0.2, 0.5, 0.2, 根据加权平均数计算公式可得这批乒乓球直径的均值为

$$39.96 \times 0.1 + 39.98 \times 0.2 + 40.00 \times 0.5 + 40.02 \times 0.2 \approx 40.00(\text{mm}).$$

达标练习

A 组

1. 10 名工人某天生产同一种零件,生产的件数分别是 15, 17, 13, 17, 18, 14, 17, 16, 14, 19, 这组数据的均值为().

- A. 14 B. 15 C. 16 D. 17

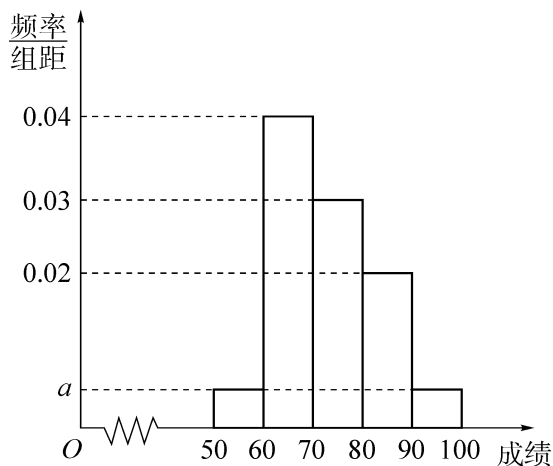
2. 在一次选秀比赛中,五位评委为一名表演者打分,若去掉一个最低分,则平均分为 90 分;若去掉一个最高分,则平均分为 86 分.由此可知,最高分比最低分高_____分.

3. 某同学使用计算器求 30 个数据的均值时,错将其中一个数据 105 输入为 15,那么由此求出的均值与实际均值的差是().

- A. 3.5 B. -3 C. 3 D. -0.5

4. 如图所示为某校 100 名学生期中考试语文成绩的频率分布直方图,其中成绩分组区间是 $[50, 60)$, $[60, 70)$, $[70, 80)$, $[80, 90)$, $[90, 100]$. 根据频率分布直方图,估计这 100

名学生语文成绩的平均分.



(第4题)

B 组

1. 已知样本数据 $x_1, x_2, \dots, x_{10}$, 其中 x_1, x_2, x_3 的均值为 a ; $x_4, x_5, x_6, \dots, x_{10}$ 的均值为 b . 该样本数据的均值为().

- A. $\frac{a+b}{2}$ B. $\frac{3a+7b}{10}$ C. $\frac{7a+3b}{10}$ D. $\frac{a+b}{10}$

2. 一组数据的和为 87, 均值为 3, 则这组数据的个数为().

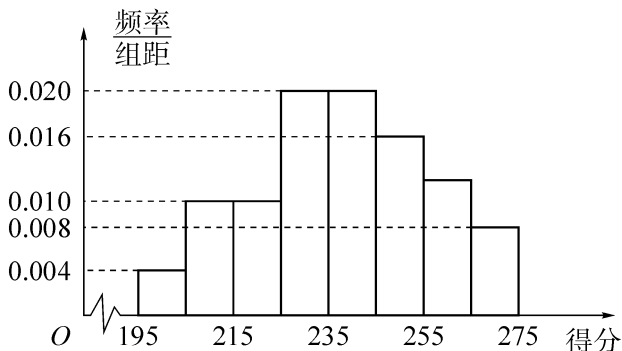
- A. 87 B. 3 C. 29 D. 90

3. 某班有 48 名学生, 在一次数学考试后统计出平均分为 70 分, 后来发现有 2 名学生的成绩有误, 甲实得 80 分却记为 50 分, 乙实得 70 分却记为 100 分, 更正后平均分是().

- A. 70 B. 72 C. 68 D. 65

4. 某校组织自主招生考试, 其有 2 000 名学生报名参加了笔试, 成绩均介于 195 分到 275 分之间, 从中随机抽取 50 名学生的成绩进行统计, 将统计结果按如下方式分成八组: 第一组 $[195, 205)$, 第二组 $[205, 215)$ 第八组 $[265, 275)$. 如图是按上述分组方法得到的频率分布直方图.

- (1) 从这 2 000 名学生中任选 1 人, 这个人的分数在 255~265 之间的概率约为多少?
(2) 求这 2 000 名学生的平均分.



(第4题)



11.6.2 样本方差与标准差



课标要求

KEBIAO YAOQIU

1. 理解样本方差的意义,会根据公式计算一个样本的方差.
2. 会用样本方差估计总体数据的离散程度.



知识导读

ZHISHI DAODU

1. 一组数据中的每一个数据与这组数据均值的差的平方和除以数据的个数所得的值称为这组数据的方差.
2. 因为方差与原始数据的单位不同,我们将方差的算术平方根称为这组数据的标准差. 标准差也可以刻画数据的稳定程度.
3. 一般地,设一组样本数据 $x_1, x_2, \dots, x_n$ 的均值为 $\bar{x}$, 则称

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

为这个样本的方差,称其算术平方根 $s = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ 为样本的标准差.



学法建议

XUEFA JIANYI

本课时学习重点是理解样本方差的意义,能用样本方差描述总体水平;难点是理解方差是刻画数据离散程度的特征值;关键是理解在有些实际问题中并不需要知道考察对象的整体情况,而只需要了解数据的离散程度即可.

学习中要注意以下几点:

1. 统计里有两类特征值,一类反映数据的集中趋势,常见的有均值、中位数、众数等;一类反映数据的离散程度(波动大小),常见的有极差、方差、标准差等.
2. 在有关数据的离散程度的特征值中,最重要的是方差与标准差. 标准差是方差的算术平方根,在很多情况下采用方差与采用标准差实际上是等价的,但采用方差反映离散程度的不足之处在于其与原始数据单位不一致,并可能夸大偏差的程度.



例题导练

LITI DAOLIAN

例 1 已知一个样本 $1, 2, 3, x, 5$ 的均值为 3 , 则这个样本的极差为 _____, 方差为 _____.

【思路分析】 根据极差和方差的定义计算即可.

解 因为 $(1+2+3+x+5) \div 5 = 3$, 所以 $x=4$, 故极差 $= 5-1=4$.

$$\text{方差 } s^2 = \frac{1}{5} [(1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2] = 2.$$

【思维诊断】 熟练掌握样本均值、方差的计算公式. 本题在未确定 x 的值的条件下, 是不能直接求极差的.

例 2 已知某 7 个数的均值为 4 , 方差为 2 , 现加入一个新数 4 , 此时这 8 个数的均值为 $\bar{x}$, 方差为 s^2 , 则 ().

A. $\bar{x} = 4, s^2 < 2$

B. $\bar{x} = 4, s^2 > 2$

C. $\bar{x} > 4, s^2 < 2$

D. $\bar{x} > 4, s^2 > 2$

【思路分析】 依据均值与方差的定义求解.

解 因为 $\bar{x} = (7 \times 4 + 4) \div 8 = 4$, 所以 C, D 错误.

又加入一个新数据后的平均数与原七个数据的平均数相同, 所以 8 个数据每一数据与均值差的平方是原七个数据每一数据与均值差的平方和与 0 之和, 即

$$s^2 = (7 \times 2 + 0) \div 8 < 2.$$

所以 A 正确.

【思维诊断】 根据条件直接求解是解选择题的基本方法.

例 3 在一次歌唱大奖赛上, 七位评委为某歌手打出的分数如下: $9.4, 8.4, 9.4, 9.9, 9.6, 9.4, 9.7$, 去掉一个最高分和一个最低分后, 余下数据的均值和方差分别为 _____、_____.

解 去掉一个最高分 9.9 和一个最低分 8.4 后, 余下数据为 $9.4, 9.4, 9.6, 9.4, 9.7$. 余下数据的均值为

$$(9.4 + 9.4 + 9.6 + 9.4 + 9.7) \div 5 = 9.5,$$

方差为

$$[(9.4 - 9.5)^2 + (9.4 - 9.5)^2 + (9.6 - 9.5)^2 + (9.4 - 9.5)^2 + (9.7 - 9.5)^2] \div 5 = 0.016.$$



达标练习

A 组

1. 抽样统计甲、乙两名射击运动员的 5 次训练成绩(单位:环),结果如下,则成绩较为稳定(方差较小)的那位运动员成绩的方差为_____.

运动员	第 1 次	第 2 次	第 3 次	第 4 次	第 5 次
甲	87	91	90	89	93
乙	89	90	91	88	92

2. 在某项体育比赛中,七位评委为一名选手打出的分数如下:

90 89 90 95 93 94 93

去掉一个最高分和一个最低分后,余下数据的均值和方差分别为().

- A. 92,2 B. 92,2.8 C. 93,2 D. 93,2.8

B 组

1. 若样本 $9, 10, 13, x, y$ 的均值为 9, 方差为 6.8, 则 $x^2 + y^2 =$ _____.

2. 如果 $x_1, x_2, \dots, x_n$ 的均值为 a , 标准差为 s , 则 $x_1 + 2, x_2 + 2, \dots, x_n + 2$ 的平均数和标准差分别为().

- A. a, s B. $2+a, s$ C. $2+a, 2s$ D. $2+a, 4s$

3. 已知一组数据 $x_1, x_2, \dots, x_n$ 的方差是 2, 则另一组数据 $3x_1, 3x_2, \dots, 3x_n$ 的方差是().

- A. 2 B. 18 C. 12 D. 4