



普通高等教育“十二五”规划教材

教育统计与测评导论

(第二版)

刘新平 张运良 编著



科学出版社

普通高等教育“十二五”规划教材

教育统计与测评导论

(第二版)

刘新平 张运良 编著

科学出版社

北 京

内 容 简 介

本书是在第一版教材的基础上结合中学新课程标准的实施进行的修订。全书系统讲述教育统计、教育测量与评价的基础知识,部分章节进行了调整,增加了有关的案例分析及评价课题的研究,并收录了作者部分研究成果。主要包括:描述性统计,概率论基础,推断性统计,方差分析,回归分析,教育测量概述,考试质量的分析与评价,考试类别与考试设计,考试题型与命题技术,项目反应理论简介,教育评价概述,教学评价的基本方法,教育教学质量的模糊综合评价,教育评价中的多元分析方法及附录(理论与应用成果选编)等,为了便于教学对习题进行了调整并给出了解答。

本书可以作为高等学校教育类各个专业、学科教学论各个专业的研究生、教育硕士及高等师范院校本科生的教材;也可以作为教育行政管理干部、考试管理干部及大学、中学教师培训教材,本书还可供教育研究及管理工作参考。

图书在版编目(CIP)数据

教育统计与测评导论/刘新平,张运良编著. —2版. —北京:科学出版社,2013.11

普通高等教育“十二五”规划教材

ISBN 978-7-03-038549-9

I. ①教… II. ①刘… ②张… III. ①教育统计-高等学校-教材②教育测验-高等学校-教材 IV. ①G40-051②G449

中国版本图书馆 CIP 数据核字(2013)第 211356 号

责任编辑:张中兴 / 责任校对:宣 慧
责任印制:阎 磊 / 封面设计:迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京文林印务有限公司印刷

科学出版社发行 各地新华书店经销

*

2003 年 6 月第 一 版 开本:720×1000 B5

2013 年 11 月第 二 版 印张:23

2013 年 11 月第十三次印刷 字数:469 000

定价:39.00 元

(如有印装质量问题,我社负责调换)

前 言

教育统计与测评是当今教育科学研究三大领域之一,是数量化方法在教育科学研究中的直接渗透,是现代教育科学研究的方法论基础,是教育工作者必备的一种重要科学工具.随着教学改革及教育科学研究,尤其是教育教学实验研究及考试研究的不断深入,教育统计与测评已列为高等学校教育学科及师范院校各专业学生的主要基础课程,许多师范院校已将此列为全校公共限定选修课;学科教学论各专业,教育类各专业研究生以及教育硕士也将此列为学位课程;各级各类教师及教育管理人员,考试管理干部培训都将此列为必学课程之一,这对提高师范生、研究生以及教育管理、考试管理人员的综合素质,更好地从事教育科学及考试研究工作都有重要意义.为了满足不同层次、不同人员教学及研究需要,我们在积多年教学经验基础上,经过充分论证编写了这本书,内容共14章及附录,包括教育统计学、教育测量与评价以及作者应用研究成果选编等.本书读者对象为教育学、心理学各专业、学科教学论各专业研究生;教育硕士、师范院校本科生、专科生;教育行政管理干部、考试管理干部以及大、中、小学校教师等.如作为教材使用,可根据学时多少及教学大纲要求予以取舍.

本书有以下特点:

(1) 在结构上将教育统计、测量与评价三方面内容有机结合,组成一套完整的结构体系.同时各部分内容又相对独立,使读者既能掌握各部分基本原理和实用方法又能了解各部分联系.这样编排便于在学时不同情形下,有针对性地安排教学.

(2) 在内容上,针对不同层次需要,除选取读者必须掌握的统计与测评原理和方法外,并适当考虑研究生及教育管理和教学研究人员的需要,增加了多元统计分析方法以及作者的部分研究成果以供参考.

(3) 在方法上力求使叙述简明、通俗易懂,内容由浅入深,循序渐近,强调直观背景和应用背景,对有关数学定理、公式的证明尽可能加以详细推导,使读者对公式的来龙去脉有较深刻的理解,这也是本书不同于其他同类书之处.有些过于繁、难的证明和推导予以省略.重点叙述理论、公式和方法的应用.

本书是编写过程中参考了多种书籍,吸取了某些好的经验和方法,引用了一些数据和资料,在此谨向各位作者表示诚挚的感谢.

本书是陕西省考试管理中心、陕西师范大学研究生院、西安文理学院科技处资助项目成果,本书各章编写分工为:第1~5章及附录由刘新平撰写;第6~10章由张运良撰写;第11~13章由何艳撰写;第14章由胡永宏撰写,习题解答由李翠撰

写,全书由刘新平统稿,参与部分工作的有李英、张幸梅、秦桂凤、杨东长、武娇、马玉芳等.本书的顺利完成,离不开陕西省考试管理中心的大力支持,也离不开陕西师范大学教务处、研究生院和西安文理学院科技处的大力支持.科学出版社的张中兴编辑为本书的出版付出了辛勤劳动,北京师范大学孟庆茂教授通读了全书,提出了许多宝贵意见并为本书作序,在本书出版之际作者一并致以真诚的谢意.

本书自2003年出版以来,得到了广大读者的支持和厚爱,2005年,本书被评为陕西省省级优秀教材.近年来,作者结合中学新的课程标准的实施,各级各类学校关于教育统计与测评工作的深入开展,结合承担的教育部考试中心关于教育考试及评价课题的研究,对本书进行了认真修订,部分章节进行了调整,增加了有关案例分析及理论应用成果.为便于教学并对习题进行了调整,给出了解答.

由于我们经验不足,水平有限,书中不妥之处一定存在,恳切希望专家及读者批评指正.

编 者

2013年5月

第一版序

教育统计、测量与评价是当今教育科学研究中的重要领域之一,同时也是教育科学研究方法的重要内容,是数量化方法在教育科学研究中的直接渗透,是现代教育科学研究的方法论基础,是教育工作者必备的一种重要科学工具.随着教育教学改革及教育科学研究,尤其是教育教学实验研究及考试研究的不断深入,教育统计、测量与评价已被列为高等学校教育学科及师范院校各专业学生的主要基础课程和各级各类教师及教育行政领导干部培训的主要科目.各级各类教育领导干部都充分认识到,所有的教师和教育行政领导掌握教育科学研究方法,对于提高科研能力有重要作用.而提高科研能力对于科研兴校、提高教师综合素质,落实素质教育、做到教育兴国的重大意义也日渐彰著.

为了满足不同层次、不同人员教学及研究需要,刘新平教授在积多年教学经验基础上,经过充分论证编写了该书,将自己多年的心得汇集与大家分享.该书在结构上将教育统计、测量与评价三方面内容有机结合,组成一套完整的结构体系,同时各部分内容又相对独立.在内容上,针对不同层次需要,除选取读者必需掌握的统计与测评原理和方法外,并适当考虑其他人员的需要,增加了多元统计分析方法以及作者的部分研究成果以供参考.在文字上叙述简明、通俗易懂,内容由浅入深,循序渐近,强调直观背景和应用背景,对有关数学定理、公式的证明尽可能加以详细推导,会使读者对公式的来龙去脉有较深刻的理解.

值此该书出版之际,除向作者表示祝贺之外,也向广大教育工作者予以推荐.

北京师范大学心理学院 孟庆茂

2003年5月2日

目 录

前言

第一版序

第 1 章 描述性统计 1

1.1 数据的收集 1

一、简单随机抽样 2

二、分层抽样 3

三、等距抽样 3

1.2 数据的整理 4

一、数据的基本类型 4

二、频数分布表 5

三、累积频数分布表 6

四、频数分布图 7

五、累积频数分布曲线图 9

1.3 集中量数 10

一、平均数 10

二、众数 14

三、中数 15

四、三种集中量数的比较 16

1.4 差异量数 17

一、极差 18

二、四分位差 18

三、平均差 19

四、标准差 20

五、差异系数 21

六、几种差异量数的比较 23

思考与练习 23

第 2 章 概率论基础 25

2.1 事件与概率 25

一、事件 25

二、概率 27

2.2 随机变量与概率分布 29

一、随机变量 29

二、二项分布 35

三、正态分布 37

2.3 统计量及抽样分布 43

一、总体和样本 43

二、样本分布函数 43

三、统计量 44

四、三种重要的理论分布 χ^2 , t , F 46

五、统计量的分布(抽样分布) 47

思考与练习 49

第 3 章 推断性统计 51

3.1 参数估计 51

一、点估计 51

二、区间估计 54

3.2 参数假设检验 60

一、假设检验的概念 61

二、总体均值的检验 64

三、总体方差的检验 70

四、二项分布总体参数检验 73

3.3 非参数假设检验 76

一、总体分布的统计检验 76

二、两个样本是否来自同一总体的

检验 81

案例分析 87

思考与练习 88

第 4 章 方差分析 90

4.1 单因素方差分析 90

4.2 双因素方差分析 98

一、无交互作用的双因素方差

分析 98

二、有交互作用的双因素方差

分析 101

4.3 方差秩分析·····	105	四、测验的实施方法·····	136
一、独立样本的方差秩分析·····	105	6.3 测量的误差·····	136
二、相关样本的方差秩分析·····	106	一、误差及其种类·····	137
案例分析·····	108	二、真分数与相对误差·····	137
思考与练习·····	110	三、测量误差的来源·····	138
第5章 回归分析 ·····	112	思考与练习·····	139
5.1 一元线性回归·····	112	第7章 考试质量的分析与评价 ·····	140
一、如何配回归直线·····	112	7.1 考试的信度·····	140
二、样本相关系数与平方和分解		一、信度的定义·····	140
公式·····	114	二、信度系数的计算·····	141
三、相关性检验·····	117	三、信度的应用·····	146
四、预测·····	118	四、提高测验信度的途径·····	148
五、应注意的几个问题·····	119	7.2 考试的效度·····	148
5.2 多元线性回归·····	120	一、效度的定义·····	148
一、用最小二乘法确定 a, b_1, b_2		二、效度与信度的关系·····	149
·····	120	三、效度的估算·····	150
二、相关性检验·····	122	四、提高测验效度的途径·····	152
三、偏相关系数·····	124	7.3 试题的难度·····	153
5.3 相关系数的其他表示法		一、难度及其计算·····	153
·····	126	二、难度对测验的影响·····	156
一、斯皮尔曼秩相关系数(秩相关)		三、难度的转换·····	157
·····	126	7.4 试题的区分度·····	159
二、点二列相关·····	127	一、区分度及其计算·····	159
案例分析·····	129	二、区分度对测验的影响·····	162
思考与练习·····	129	7.5 试卷分析报告案例·····	164
第6章 教育测量概述 ·····	131	一、命题思路、依据·····	164
6.1 教育测量的概念·····	131	二、考试质量分析·····	165
一、什么是教育测量·····	131	三、试卷分析·····	167
二、测量的基本要素·····	131	四、教学建议·····	168
三、教育测量的特点·····	132	思考与练习·····	168
四、教育测量的有关量表简介		第8章 考试类别与考试设计 ·····	171
·····	133	8.1 考试类别·····	171
6.2 测验·····	134	一、国家级考试·····	171
一、测验的定义·····	134	二、学业成就阶段考试·····	171
二、测验的分类·····	134	三、其他考试·····	172
三、教育测验的作用·····	136		

8.2 考试标准化·····	173	二、项目反应理论的产生和发展	199
一、传统考试与标准化考试	173	三、潜在特质理论 ·····	200
二、标准化考试的主要环节	174	10.2 项目反应理论的数学模型	200
8.3 考试设计·····	175	一、项目特性曲线 ·····	200
一、考试目标的设计 ·····	175	二、项目反应函数 ·····	202
二、考试大纲的制定 ·····	177	10.3 项目反应理论的应用	203
三、双向细目表及其要素 ·····	178	一、题库建设与题目参数等值	203
四、试题设计 ·····	180	二、自适应测验的编制 ·····	203
五、试卷的编制 ·····	181	三、目标参照性测验的编制	204
思考与练习 ·····	182	思考与练习 ·····	205
第9章 考试题型与命题技术 ···	183	第11章 教育评价概述·····	206
9.1 客观性试题的特点与编制	183	11.1 教育评价的基本概念	206
一、选择题的编制 ·····	184	一、教育评价的意义和对象	206
二、是非题的编制 ·····	185	二、教育评价的作用和功能	206
三、匹配题的编制 ·····	186	三、教育评价的类型 ·····	207
四、填空题的编制 ·····	187	四、教育评价的基本原则 ·····	207
五、简答题的编制 ·····	188	五、教育评价的实施 ·····	208
9.2 主观性试题的特点和编制	188	11.2 教育评价的指标体系	209
一、论述题的编制 ·····	189	一、教育评价指标体系的基本要素	209
二、作文题的编制 ·····	191	二、教育评价指标体系的设计	210
三、应用题编制 ·····	191	11.3 教育评价的方法和工具	213
9.3 主观题与客观题的差异	193	一、教育评价中的数学方法	213
一、试题的呈现方式和解答方式	193	二、教育评价中常见的数量化方法	214
二、考查的内容层次不同 ·····	193	三、其他评价方法 ·····	216
三、评分过程的差异 ·····	193		
9.4 题库建设·····	194		
一、什么是题库 ·····	194		
二、如何建立题库 ·····	195		
思考与练习 ·····	198		
第10章 项目反应理论简介·····	199		
10.1 项目反应理论产生 ···	199		
一、经典测量理论的缺陷 ·····	199		

四、教育评价的工具	217	一、评价教学质量的指标体系	248
思考与练习	217	二、模糊信息的获得	249
第 12 章 教学评价的基本方法	218	三、模糊综合评价的方法和步骤	250
12.1 考试结果的评价	218	13.4 模糊综合评价在学生德育评价中的应用	254
一、不同班级同学科的考试成绩评价	218	一、德育评价指标体系	254
二、同班(年)级不同学科的考试成绩评价	220	二、评价的方法和步骤	255
三、不同测验分数等值问题中的评价	222	案例分析	257
12.2 学生学习进步幅度的评价	225	思考与练习	262
一、双向表评价法	225	第 14 章 教育评价中的多元分析方法	263
二、进步有效率的差异分析	226	14.1 主成分分析	263
三、N 评价法	227	一、主成分分析的基本原理	263
12.3 标准分数在对学生个体成绩评价中的应用	231	二、主成分的计算步骤	264
一、学生各科成绩比较的评价	231	三、用主成分进行综合评价	266
二、学生某科成绩进步程度的评价	232	14.2 因子分析	269
案例分析	234	一、因子分析的基本原理	270
思考与练习	239	二、用因子分析方法进行教育评价的步骤	272
第 13 章 教育教学质量的模糊综合评价	241	14.3 聚类分析	275
13.1 模糊综合评价的意义	241	一、聚类分析的基本步骤	276
13.2 模糊综合评价的数学模型	241	二、普系聚类分析	280
一、有关数学基础	241	14.4 判别分析	287
二、模糊综合评价的数学模型	246	一、距离判别法的基本思想	287
13.3 模糊综合评价在教学质量评价中的应用	248	二、两个总体的距离判别法	287
		思考与练习	293
		附录	294
		一、理论与应用成果选编	294
		二、附表	331
		部分参考答案	352
		参考文献	358

第1章 描述性统计

统计学是一门研究数据的搜集、整理、分析与推断方法的科学,是一门数据的艺术和科学,单纯对一组数据的面貌特征进行分析研究称为描述性统计。例如,某中学要了解初一学生期末数学成绩,通常是先取得数据,对数据归纳分类,列表画图,再计算出这组数据的代表值以及衡量这组数据分散程度的数量指标,用以描述这组数据的特征。然而,在实际工作中,我们往往需要利用一组数据获得信息作决策去推断更一般的情形。例如,我们要了解某区初一学生期末数学成绩,事实上,不必要对全区所有初一学生进行统计调查,而是选取样本,通过对样本的描述来推断整体的特性,统计学上称此为推断性统计,简称为统计推断。这种基本的统计方法可以通俗地理解为“坐井观天”。

当我们根据样本信息进行统计推断时,势必要冒导致错误结论的风险,因为样本并非总是与总体一致。例如,可能正好抽到数学成绩都好的初一学生组成样本去推断全区初一学生的数学成绩显然存在不确定因素。但是通常情况下,我们宁愿担风险也不愿考虑所有对象,因为考虑所有对象工作量太大而且有时是不可能的。这样我们就需要研究如何抽取样本、什么样的样本较为合适;同时,还要考虑如何有效地处理和分析数据、如何设计最佳试验方案以减少导致错误结论的风险,并应用数学理论计算出出现这种风险可能性的大小。描述这种可能性大小的数量指标称为概率。运用概率论来研究统计学的学科称为数理统计。随着科学理论的发展,数理统计已成为数学的一个重要分支,目前已广泛应用于科学和社会中一切领域。

教育统计学是运用数理统计方法去研究教育现象的一门应用学科。教育学与心理学中许多问题借助于统计学都可以予以量化,从而揭示教育规律和心理规律。例如,学生学习成绩的分布,教学方法实验的比较,不同年龄组幼儿智力差异的比较,刺激与反应联结的关系等。每个教育工作者都应该掌握这门科学方法,这对改进教育管理水平、培养教育科研能力、正确开展教育与心理实验、提高教学质量和实际工作效率都将是十分必要的。

1.1 数据的收集

统计离不开数据,数据也称为资料,我们把搜集记录下来的数量依据称为数据。实际工作中,一般采用抽样调查的方法来取得数据,我们以一个例子来说明这种方法。

“某地区初中一年级学生每星期约看几小时电视?”是个需要统计的问题.就某一个学校来考虑,获得这些数据并不困难,只要对每个学生调查一下就行了.但是,某地区拥有众多学校,不可能一一调查,因此我们只能从全地区所有初中一年级学生中抽出部分学生展开调查来获取数据,统计学上称此为抽样调查.我们所考虑对象的全体统计学上称为总体或母体,其中每一个对象称为个体,而从总体中抽取的一部分个体称为样本或子样,样本中所含个体的数目称为样本容量,通常用字母 n 表示.样本分大样本($n \geq 30$)与小样本($n < 30$),样本容量的选取取决于实验的条件和精度,样本越大,反映总体的信息越充足,但计算量也越大,故样本容量最好适当.我们以后会看到,对于不同的样本采用的统计方法也不尽相同.

采用抽样调查获取数据要受到一些条件的制约.例如,这个地区初中一年级学生每星期看电视的小时数受学生有无看电视的嗜好、对电视节目喜好程度的影响,可能得到不同的调查数据.如果抽到的学生大多喜欢看电视,则相应看电视的时数可能多一些.这就是说,随着样本选取的不同,结果也可能不同.因此,在抽样调查中要求样本:①能充分反映总体的信息,也就是说每个个体被抽到的可能性相同,个体与个体之间互不影响,数学上称这样的个体互相独立;②每个个体具有和总体相同的本质特性,即样本具有某种代表性,数学上称此为与总体同分布.满足以上两条的样本称为随机样本,也可以通俗地说,随机样本是指凭机会抽取的样本.从总体中抽出一个随机样本,统计学上称为随机抽样.

随机抽样的方法很多,常用的有以下几种.

一、简单随机抽样

1. 随机数表法

随机数表是根据数理统计的原理,由许多随机数学排列起来的数字表(附表1),表中数字的构造方法是:利用计算机使 $0, 1, \dots, 9$ 十个数字号码中每个自动出现一个号码,用这种方式得到一串数,编排成组(一般四个数为一组).我们仍以“某地区初中一年级学生每星期约看几个小时电视”为例,说明使用随机数表的方法.例如,某地区有 3000 名初中一年级学生,需抽出容量为 40 的一个样本.先将 3000 人从 0001 号编到 3000 号,第一步,闭上眼睛用铅笔在表上任意划一个点,规定如点到奇数则查第一页,如点到偶数则查第二页;第二步,在选定的那一页上再点一次,由点中的数字决定从哪一行开始;最后,再点一次决定从哪一个数起,然后以此为起点,开始以四位数字为一节读下去,小于等于 3000 的选中,大于 3000 的舍去,直到取满 40 个数据为止,这 40 个数据对应的学生即为选中的随机样本.

2. 抽签法

抽签法是将所有个体编号打乱次序用类似于抽签的方法从中来获取随机样

本. 例如, 要从某班 50 名学生中抽取容量为 5 的一个样本. 先把全班学生的编号写在不同的卡片上, 然后将卡片放进一只盒子里, 把盒子摇几摇使卡片混杂, 再从盒中抽出 5 张卡片, 卡片所对应学生的编号即为选中的随机样本.

二、分层抽样

在抽样时由于总体内各个个体之间差异较大, 因此, 在样本容量一定的条件下进行分层抽样可以提高调查结果的精度.

分层抽样是按一定标准把总体内的每个个体划分为若干层, 使相互差异小的个体集中在一层内, 从而可以缩小各层内个体之间的差异程度, 使样本中各个个体在总体中散布更均匀.

分层抽样时, 从各层抽取的样本个数可以与各层个体数成比例. 具体做法如下.

把总体中 N 个个体划分为 L 个不相重叠的部分, 使每一部分包含的个体数分别为 $N_1, \dots, N_L$, 且

$$N_1 + N_2 + \dots + N_L = N,$$

则第 h 层所含的样本个体数 n_h 为

$$n_h = n \cdot \frac{N_h}{N}, \quad h=1, \dots, L, \quad (1.1.1)$$

其中 n 为样本容量, $\frac{N_h}{N}$ 为第 h 层的层权数.

例如, 要从某校 210 名 7~9 岁儿童中抽出三分之一进行智力测验. 已知该校 7 岁儿童有 63 人, 8 岁儿童有 112 人, 9 岁儿童有 35 人, 现在用分层抽样法确定各年龄组儿童人数. 由式(1.1.1)得

$$n_1(7 \text{ 岁组}) = (210/3) \times (63/210) = 21(\text{人}),$$

$$n_2(8 \text{ 岁组}) = (210/3) \times (112/210) = 37(\text{人}),$$

$$n_3(9 \text{ 岁组}) = (210/3) \times (35/210) = 12(\text{人}).$$

三、等距抽样

实际应用中为了保证样本在总体中分布的均匀性, 有时也采用等距抽样法. 所谓等距抽样法是把所有个体按顺序排列起来, 然后以确定的相等距离抽取随机样本. 例如, 某大学抽查部分学生作业进行检查, 先确定学生学号中的一个数, 如末位是 3 的学号, 然后依次取各系各班学号末位是 3 的学生作业为样本. 又如, 要从 100 名儿童中抽出 20 人进行智商测定. 先对 100 名儿童编号, 然后从 1~20 中任抽一个号, 如果抽到 4, 则每隔 5 名取一人, 从而得到容量为 20 的一个样本. 显然, 这种方法人为地规定了距离, 故抽样的随机性有所减弱.

抽样分有放回抽样(从总体中抽出一个个体记下其特征后再放回总体, 然后进

行第二次抽样)和无放回抽样(从总体中抽出一个个体后不再放回去,再抽第二次)两种方式.当总体内个体数目较多时,这两种抽样方式没有本质区别.教育统计中一般采用无放回抽样,但由于有放回抽样能简化某些计算,故当总体内个体数目较多时,可以看成是有放回抽样.

通过抽样获取数据离不开求实的科学态度和认真的工作作风,数据如果不准确、不完整、或有遗漏,不仅数据本身失去价值,而且以此进行分析推断还会导致错误的结论.

作为本节的结束,我们回顾一个收集数据要注意的几个问题:

(1) 是否需要抽取样本? 如果需要,样本多大为宜? 样本是随机抽取的吗?

(2) 样本结论是否代替总体结论? 也就是说,样本是否来自一个总体? 因为如果把样本结论用到另一个总体上去,可能会出现错误的结论.例如,根据一个学校抽取的样本预测有 80% 的中学生能考上大学而断言本地区也有 80% 的中学生能考上大学显然是不切合实际的.

1.2 数据的整理

一组数据受获取方法的限制,在未经整理前多是分散和零乱的,为了揭示和发现一组数据的内在规律,我们首先要做的工作是把数据列在表上,绘制成图形进行归纳分类.

一、数据的基本类型

教育统计中常见的数据类型有以下两种.

1. 离散型数据

离散型数据一般指取整数值数量指标.这类数据是计数性的,数据之间不能再划分为更小的单位.例如,学校的个数、学生的人数;有些教育现象的指标是按属性来划分的,如学生能力分为优、良、中、差.但我们可以使其量化,用 5 代表优、4 代表良、3 代表中、2 代表差,这样得到的数据仍然是离散型的.

2. 连续型数据

连续型数据一般指经过度量和测定而得到的数量指标.这类数据取值可以连续变化,尽管数据本身仍然是数轴上的点,但数据与数据之间可以无限细分,也就是数据的取值范围可以充满一个区间.例如,学生的考试成绩、儿童的体重等.连续型数据通常以小数形式出现,虽然有时也会以整数形式出现,但当提高精度后总会出现小数.例如,某学生期末语文成绩为 90 分,可以记为 90.0 分.

确定了数据类型,可以进一步利用频数分布表和频数分布图来研究数据的变化规律.

二、频数分布表

一组数据中每个数据出现的次数称为这个数据的频数.按频数分类列出的一览表称为频数分布表.

1. 离散型数据的频数分布表

例 1 某幼儿园测定 5 岁组儿童智力,共 7 个项目,全园 30 名 5 岁儿童中有 1 人答对 1 项、3 人答对 2 项、4 人答对 3 项、8 人答对 4 项、7 人答对 5 项、5 人答对 6 项、2 人答对 7 项.列出频数分布见表 1.1.

表 1.1 30 名儿童智力测定分布

答对题数	频数	答对题数	频数
1	1	5	7
2	3	6	5
3	4	7	2
4	8	总和	30

2. 连续型数据的频数分布表

例 2 附中初二年级实验班 40 名学生期末数学统考测验得分如下:

68.0	79.0	80.0	83.0	62.0	59.0	80.0	81.0
61.5	83.0	67.0	97.0	63.5	93.0	76.0	97.5
84.5	64.0	75.0	88.0	91.0	78.5	67.0	94.0
81.0	70.0	86.5	72.0	85.0	94.0	78.0	91.0
60.0	75.5	82.0	91.0	95.0	52.0	76.5	80.0

这一组数据中最大值是 97.5,最小值是 52.0,可见数据分布很散,项数较多.因此我们将它们分组,组的范围称为组区间,每组的起止分别称为组下限和组上限,每组的大小称为组距,各组组距一般是相同的.分组的原则是:100 个以上的数据分为 12~20 组,数据较小则分为 8~10 组,分组多则计算量大,分组少则可能将不同性质的数据归在一起产生较大的误差.组距为便于计算一般取为 3,5,10 较为合适,本例分为 10 组,组距取 5.将组号放在表的第一列,组区间放在第二列,组中值记为 m .放在放三列, $m=\frac{1}{2}(\text{上限}+\text{下限})$,然后用类似于选举唱票、对号入座的方法数出各组的频数放在第四列.细心的读者可能会出现,如果第二组为(56.5~

61.5),则第二组的上限 61.5 正好是第三组的下限,那么 61.5 放在哪一组? 对此,我们可以提高一位分点或降低一位分点,本例取两位小数,这样第二组为(56.55~61.55),因此 61.5 应放在第二组,有了组频数,当然也可以算出组频率(每组组频数与总频数之比),为了以后的方便,把频率放在表的最后一列(表 1.2),有了频数分布表,大致可以看出数据所呈现的统计规律性.

表 1.2 附初中二年级实验班期末数学统考测验成绩分布

组号	组区间	组中值	频数	频率
1	51.55~56.55	54.05	1	1/40
2	56.55~61.55	59.05	3	3/40
3	61.55~66.55	64.05	3	3/40
4	66.55~71.55	69.05	4	4/40
5	71.55~76.55	74.05	5	5/40
6	76.55~81.55	79.05	8	8/40
7	81.55~86.55	84.05	6	6/40
8	86.55~91.55	89.05	4	4/40
9	91.55~96.55	94.05	4	4/40
10	96.55~100.55	98.28	2	2/40
总和			40	1.00

综上,对于分组数据编制频数分布表归纳为以下几步:

第一步 找极差 R , R =最大值-最小值,由 R 大致了解数据的差异范围.

第二步 定组距 d ,一般为便于计算, d 多取为 2,3,5,10 等.

第三步 定组数 r ,一般数据在 100 个以上,分为 12~20 组,数据较少则分为 8~10 组,也可以借用下面公式确定近似组数.

组数 $r = \left[\frac{R}{d} \right] + 1$,其中方括号为 $\frac{R}{d}$ 的整数部分, d 为组距.

例 2 中, $r = \left[\frac{45.5}{5} \right] + 1 = 10$.

第四步 定分点,通常使分点比原测量精度多一位或少一位,要注意的是最低组的上、下限应能包括最小值,最高组的上、下限应能包括最大值.

第五步 数频数,根据组限归类,数出全体数据落入每一组的个数.

频数分布表也有其缺点,这是因为分组后原数据不见了,我们在 1.3 节会看到计算描述一组数据特征的数依赖于各组的组中值,因而出现了误差.但是在理论上一般假定各组内频数分布是均匀的,因而各组的误差会相互抵偿,使总误差减少.

三、累积频数分布表

频数分布表给出了各组的频数分布情况,然而有时需要知道某一数据之上或

之下的频数有多少. 如例2中需要知道成绩低于71.5分的人数, 则把前四组数据累加起来, 因而给出累积频数分布表有一定的实际意义.

累积频数分布表的列法很简单, 只需在频数分布表上添加一列累积频数即可. 具体方法是: 从数值最小的一组开始, 逐组累加频数至数值最大的一组, 最后累加的频数与总频数相等. 如果把累积频数除以总频数, 则得到相应的累积频率, 再把累积频率乘以100, 则得到相应的累积百分比. 表1.3给出了例2中数据的累积频数、累积频率、累积百分比分布表.

表 1.3 附中初二年级实验班期末数学统考成绩累积频数分布表

组号	组区间	组中值	频数	频率	累积频数	累积频率	累积百分比
1	51.55~56.55	54.05	1	0.025	1	0.025	2.5
2	56.55~61.55	59.05	3	0.075	4	0.100	10.0
3	61.55~66.55	64.05	3	0.075	7	0.175	17.5
4	66.55~71.55	68.05	4	0.100	11	0.275	27.5
5	71.55~76.55	74.05	5	0.125	16	0.400	40.0
6	76.55~81.55	79.05	8	0.200	24	0.600	60.0
7	81.55~86.55	84.05	6	0.150	30	0.750	75.0
8	86.55~91.55	89.05	4	0.100	34	0.850	85.0
9	91.55~96.55	94.05	4	0.100	38	0.850	95.0
10	96.55~100.00	98.28	2	0.050	40	1.00	100.00
总和			40	1.000			

四、频数分布图

利用频数分布表我们已经看到了频数的一些分布规律, 如果借助于图形来表示这些规律, 将显得更为直观、形象, 通常我们用频数分布图来表示这种规律性. 常见的频数分布图有以下几种.

1. 直方图

直方图是在横轴上标出组距, 纵轴上标出频率与组距之比, 然后以每组组距为底边, 相应的频率与组距之比为高作矩形. 显然, 每个矩形的面积恰好等于数据落在该矩形所对应组内的频率, 这样所有矩形面积总和为总频率1. 直方图是利用各个矩形的高低来描绘频数分布情况的. 图1.1给出了例2中数据的直方图, 图中断裂号表示由0~51.55组区间的距离是缩短了. 有时为了简单, 横轴上只标出组中值, 包括组中值在内的区间即为本组组距.