



何江川 编著

体育统计 分析与应用

何江川 编著

体育统计 分析与应用

TIYU TONGJI FENXI YU YINGYONG

GUANGXI NORMAL UNIVERSITY PRESS
广西师范大学出版社

· 桂林 ·

图书在版编目(CIP)数据

体育统计分析与应用 / 何江川编著. —桂林: 广西师范大学出版社, 2013.2
ISBN 978-7-5495-3454-8

I. 体… II. 何… III. 体育统计—统计分析—高等学校—教材 IV. G80-32

中国版本图书馆 CIP 数据核字 (2013) 第 027633 号

广西师范大学出版社出版发行

(广西桂林市中华路 22 号 邮政编码: 541001)
网址: <http://www.bbtpress.com>

出版人: 何林夏

全国新华书店经销

衡阳顺地印务有限公司印刷

(湖南省衡阳市雁峰区园艺村 9 号 邮政编码: 421008)

开本: 787 mm × 1 092 mm 1/16

印张: 21.75 字数: 375 千字

2013 年 2 月第 1 版 2013 年 2 月第 1 次印刷

定价: 30.00 元

如发现印装质量问题, 影响阅读, 请与印刷厂联系调换。

前 言

体育教学、科研、训练已进入信息时代,由于体育领域中存在很多现象都是不确定的、随机性的,要在大量的调研或测试数据中获得有价值的科研信息,就要对数据进行有效的存储、整理、统计分析和绘制统计图表。如:竞技体育项目技术、战术结构特征、足球运动员技战术素质水平与之间的胜率、人群健康生命非遗传性影响因素、肥胖型和偏瘦型学生体质形态学特征成因等问题都需要对调研数据进行分析推理,从而达到解释与阐明现象特征客观规律的目的。因此,体育统计分析就是一项必不可少的科研工具。

目前全国各地高校开展体育科研及交流活动日趋频繁,中国体育科学研究迎来史上旺盛时期,以科研促教学成为很多高校教师的共识。广大的体育教育工作者、教练员、科研人员以及研究生越来越感受到掌握体育统计分析方法对于开展体育科学研究极为重要,迫切希望有一本通俗易懂而又能够结合体育科研、学术论文撰写的实际统计分析与应用的参考书,供平时开展科研活动进行数据处理时参考。

作者自 1992 年开始从事体育统计教学及科研,积累了丰富的教学经验以及科研成果。本书以学术论文为实例,以循序渐进的方式介绍体育统计分析与应用,利用 SPSS 17.0 for Windows 的多种统计模块对所采用的方法及得出的结果进行了详尽的解释,并对结果与图形进行统计学分析与推断。本书主要介绍体育科学研究中经常采用的描述性统计分析、均数比较分析、相关分析、回归分析、聚类分析、主成分分析、因子分析等分析方法及应用。

希望本书能成为一本对广大体育工作者有用的统计参考书。在编著过程中,参考了有关书籍,特在此谨向各书的作者和出版者表示深切的感谢。书中错误及不妥之处在所难免,敬请专家及读者给予批评和指正。



2012年10月

目 录

第一章 绪论

第一节 什么是体育统计	1
第二节 体育统计内容	2
第三节 体育统计分析的任务与思想	3

第二章 统计分析资料的收集与管理

第一节 资料的收集及案例	5
第二节 体育统计学的几个基本概念	11
第三节 SPSS 17.0 数据的录入与管理	14
第四节 频数分布 SPSS 统计分析方法	18
第五节 体育统计应用领域案例分析	21

第三章 样本特征数

第一节 集中量数	46
第二节 离散量数	49
第三节 SPSS 统计分析方法	54
第四节 样本特征数在体育科研中应用	57

第四章 正态分布

第一节 正态分布	72
第二节 标准正态分布概率计算	74
第三节 非标准正态分布概率计算	75
第四节 正态分布在体育评价中应用	77

第五节 正态分布在体育科研中应用 89

第五章 体育评分

第一节 标准 Z 数 98

第二节 标准 T 分及 SPSS 统计分析方法 104

第三节 标准百分及 SPSS 统计分析方法 108

第四节 位置百分 113

第五节 累进评分法及 Microsoft Excel 分析方法 114

第六节 体育评分在体育科研中应用 121

第六章 抽样与参数估计

第一节 抽样调查方法 128

第二节 抽样误差 130

第三节 参数估计概述 133

第四节 一个总体均数的区间估计 135

第五节 两个总体均值之差的区间估计 137

第六节 总体率的区间估计 139

第七节 SPSS 统计分析方法 140

第八节 参数估计在体育科研中应用 143

第七章 假设检验

第一节 假设检验基本原理 159

第二节 一个正态总体的均数的假设检验 161

第三节 两个正态总体的均数的假设检验 168

第四节 样本率的假设检验 175

第五节 SPSS 统计分析方法 177

第六节 假设检验在体育科研中应用 186

第八章 方差分析

第一节 方差分析原理 207

第二节 单因素方差分析 208

第三节 SPSS 统计分析方法 211

第四节 方差分析在体育科研中应用 219

第九章 相关与回归

第一节 相关分析原理 227

第二节 线性相关系数 SPSS 统计分析方法 230

第三节 等级相关系数 SPSS 统计分析方法 234

第四节 一元线性回归及 SPSS 统计分析方法 240

第五节 相关与回归在体育科研中应用 243

第十章 主成分分析与因子分析

第一节 主成分分析概述 257

第二节 因子分析概述 258

第三节 SPSS 统计分析方法 259

第四节 因子分析确定权重方法 269

第五节 主成分分析与因子分析在体育科研中应用 273

第十一章 聚类分析

第一节 聚类分析概述 285

第二节 二阶段聚类分析及 SPSS 统计分析方法 286

第三节 系统聚类及 SPSS 统计分析方法 290

第四节 聚类分析在体育科研中应用 296

附表

附表 1 标准正态分布表 313

附表 2 *t* 值表 317

附表 3 *F* 值表(方差齐性检验用) 319

附表 4 *F* 值表(方差分析用) 321

附表 5 χ^2 值表 329

附表 6 相关系数界值表 331

参考文献 335

第一章

绪论

第一节 什么是体育统计

数理统计是现代应用数学的一个重要分支,是通用于研究自然现象和社会现象的方法体系。数理统计以概率论为基础,从如何收集、整理样本资料入手,以样本为依据,对资料进行科学分析和统计推断,得到科学的推论或为制定计划提供依据。

数理统计的研究对象主要是以不确定的随机现象为研究对象,运用数学模型来推断总体的一门科学。由于它所研究的从局部观察去推论整体的方法有着普遍的意义,所以它能与各种具体的专业研究结合起来,在实践中得到广泛的应用。它应用于经济学中称为经济统计,应用于工业生产中称为工业统计,应用于医学研究中称为医学统计,应用于体育研究中便称为体育统计等。(图 1-1-1)

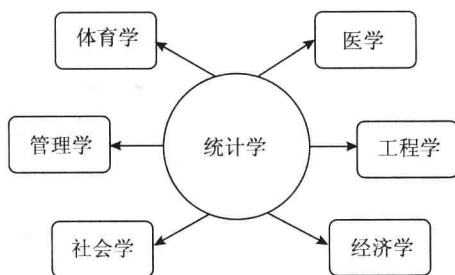


图 1-1-1

数理统计方法之所以能够运用于体育运动中,是由于体育运动中的大量现象都是不确定的、随机性的,它正是以概率论为基础的数理统计所研究的对象。由于体育运动中

的许多问题往往要依靠调查或实验才能解决,而调查或实验的人数总是有限的,所以要想从局部得来的资料推论整体的性质,得出合理科学的结论,就必须应用科学的统计方法。所以体育统计应运而生。

所谓体育统计就是在辩证唯物主义思想指导下运用统计的理论和方法,来研究体育教学、训练、科研和管理中的问题,探讨体育发展规律的一门学科。简言之,体育统计就是数理统计在体育中的应用。

第二节 体育统计内容

一、体育统计内容分类

体育统计大体上可分为描述统计、推断统计和应用统计三大类。(图 1-2-1)

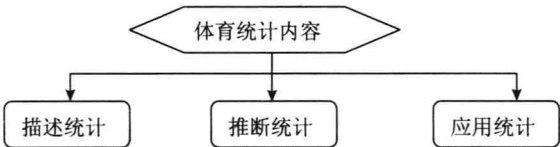


图 1-2-1

(一) 描述统计

描述统计主要研究如何整理实验或调查所得的大量数据,找出这些数据分布特征,计算出一些具有代表性的统计量(如平均数、标准差、相关系数),借助这些概括性的数据,我们就可能从杂乱无章的资料中获取有意义的信息,便于对不同的整体进行分析比较,以做出客观规律的结论。

(二) 推断统计

推断统计是在描述统计的基础上,利用数据所传递的信息,通过局部去对总体的情形加以推断。用统计语言说,就是根据样本的数据来推断总体的性质,并标明这种推断可能发生的误差的大小。

(三) 应用统计

应用统计主要研究如何应用统计分析方法去解决体育领域研究中的实际问题。

因此,体育统计是运用数学方法研究体育客观规律的一门学科。它是每一位体育工作者都应该学习和熟练掌握的工具。在体育实践中,使用描述统计的机会很多。这是基础,应该引起重视。

统计是数学的分支,有其严谨的理论依据,各种统计计算公式都是经过推导和证明得到的。体育工作者学习体育统计需要懂得一些基本的统计理论知识,但是重点应该是如何利用统计的核心思想正确地分析体育研究中的实际现象或问题。因此,描述统计、

推断统计、应用统计这三部分内容是互相联系、不可分割的。

二、描述统计、推断统计、应用统计的关系

利用统计学探索体育现象结构特征规律性的过程如图 1-2-2 所示。

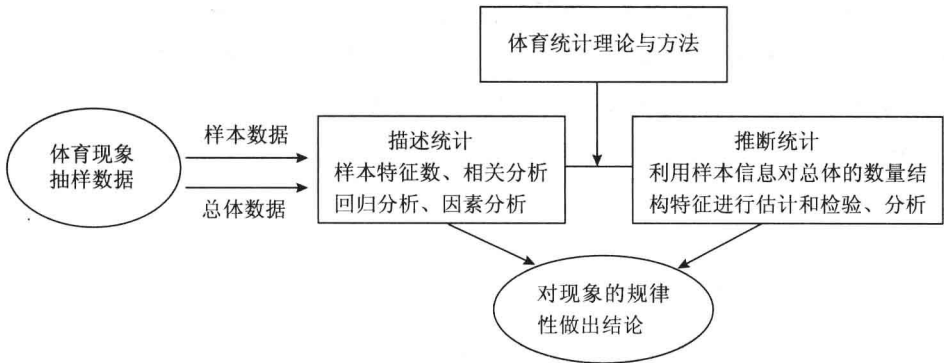


图 1-2-2

第三节 体育统计分析的任务与思想

一、体育统计工作任务

体育统计是一门应用学科,它的作用是通过统计工作任务的完成而发挥出来的。目前,我国体育统计的基本任务是为使我国全面成为世界体育强国,同时全面提高全民身体素质而服务。分开来说,应包括:(1)准确、及时、全面、系统地反映国际和国内体育发展的情况,并进行统计分析和预测,为制定政策和计划提供依据;(2)对政策和计划执行统计检查和监督;(3)为管理各项体育事业提供资料;(4)为进行体育教育、训练和科学研究提供资料。

体育统计工作的任务,主要落实在提供资料上。包括:(1)对体育发展情况的描述(体育教学、管理等);(2)对体育规律性的揭示(如运动选材);(3)包括反映现实情况和历史情况的统计资料(国民体质健康研究);(4)对现象(运动员)的发展前景作出必要的预测。这任务既包括“揭露矛盾”,也包括提出建议。资料提供得愈多,统计的作用也就发挥得愈好。

体育统计是广大体育工作者学习和研究现代体育科学技术不可缺少的基础知识和基本工具。学习体育统计,有助于我们分析资料、处理数据、科学地制定实验计划,有助于提高体育教学、训练和科研水平,学习国内外有关体育方面的先进经验,培养科学思维能力与实事求是的科学态度。

学习体育统计,应着重理解统计的基本原理和基本概念,要掌握收集、整理和分析数据资料的基本知识和技能。在进行统计工作时,要重视原始资料的完整性、可靠性、代表性、一致性和可比性,对数据进行处理和分析时必须严肃认真、实事求是。

体育统计是实践性很强的应用学科,在学习这门课时,要理论联系实际,结合体育专业知识,多做练习,独立完成作业,这样才会更好地掌握这门知识。

体育统计不是万能的,它只能帮助我们认识客观事物的规律,而不能创造规律。正确应用体育统计方法,它会成为研究人员强有力的助手。但不分场合、不分条件地盲目搬用,则可能导致严重错误,在学习和应用体育统计方法时,这一点要引起我们足够的重视。

二、正确运用统计方法的观点与思想

体育科学研究在应用统计分析方法时,必须以唯物辩证法为指导,以体育科学为基础,有正确的观点与思想方法才能发挥它的作用。“统计无用”和“统计万能”的论点都是错误的。

有些人认为只凭常识判断,仅凭数字的表面值就可以看出实验研究对象之间的差别,不需要什么统计处理。这种没有考虑到抽样误差问题的判断往往是错误的。因此,正确运用统计方法能够帮助我们认识客观事物,阐明事物固有的规律,从而把感性认识提高到理性认识,才能合理地分析资料、做出具有很强说服力的结论,指导体育科研实践。

我们做任何工作,都必须掌握基本情况,做到心中有数,才能有的放矢地开展工作,从而提高工作质量。进行体育科学研究及实际应用工作也不例外。无论在运动训练、人体机能评价、运动素质特征研究、国民体质健康研究等各个方面的工作,都必须有计划地收集资料并进行合理的统计分析。例如:对全民健身计划阶段性评价,大学生体能评价,运动选材、训练和运动能力结构特征的分析,等等。因此,体育工作者必须学习和掌握统计分析方法的核心思想。在日常工作中要认真做好一切记录,有计划地进行调查研究。积累了一定的资料以后,就可以运用统计方法对资料进行科学的整理分析。透过众多的、偶然的、次要的因素,阐明事物客观存在的规律性,辨别事物间在数量上的差别是否仅是偶然现象,从而做出比较正确的结论应用于实际,并且针对工作中发现的问题,制定措施。

第二章

统计分析资料的收集与管理

体育统计方法的研究内容包括体育资料的收集、整理、分析及表达等。其方法:(1)收集资料之前必须根据研究目的制订研究设计,然后再按设计的要求去收集准确、完整的原始资料。这是进行统计工作的基础。(2)整理资料是对收集到的原始资料进行审核、归纳分组汇总,以便于统计计算。这是保证统计结果准确不可或缺的一步。(3)分析资料及表达是对经过整理的原始资料计算相应的指标和绘制必要的统计图表,结合专业知识进行分析、比较,找出其中规律,撰写研究分析论文。

以上三个基本方法是互相联系的,任何一步出现缺陷或差错都会影响研究结论的正确性。所以说,研究工作的整个过程都离不开统计方法。

第一节 资料的收集及案例

收集资料是体育统计工作的第一步,也是很重要的一步。如果收集资料的计划不妥善,收集的原始资料不正确或不完整,则无论怎样进行统计整理和分析,都不能弥补此缺陷,有时甚至使所得到的资料不能说明问题,失去应有的价值,从而浪费了不少的人力、物力和时间。收集资料的重要性由此可见一斑。因此,在统计工作中必须注意正确地、有计划地收集资料。

一、统计资料的主要来源

(一) 日常工作中积累的资料

平时进行的体质测试、上体育课、训练、开展群众性体育活动和组织运动竞赛中都可以积累许多宝贵的数据,如“国民体质测定标准”测试的数据、体育课考核的各项数据、运动员平时训练测试成绩及比赛成绩等。

（二）全面普查

如学校一年一度的“学生体质健康标准”测试。

（三）专题研究

这包括专题的调查研究 and 专题的实验研究。如,2000 年开始的每 5 年一次的全国国民体质监测,就是我国大规模的体质专题调查研究,其目的是为了摸清我国 3~69 岁国民的体质现状、特点及生长发育规律。有时为了研究某个问题,需要开展专题研究。又如,“东盟七国留学生体质及生命质量结构特征研究”(批准号:KT201101-35)获广西科学实验(中国—东盟研究)中心专题立项并进行实验研究。目的是了解东盟各国留学生体质健康与生命质量状况,为中国与东盟各高校文化交流和谐发展提供建议。

在进行专题研究前应该根据研究工作的目的和要求,结合统计原理的要求,对研究工作的全过程事先制定一个周密的工作计划,通常把制定这样的计划称为研究设计(或课题计划)。如果研究设计考虑得比较周密、完善,就能以较少的人力、物力、时间获得可靠的资料,使误差减小到最低限度,提高研究工作的质量和效率,并有助于得到科学的可靠的结论。反之,事先不做研究设计,随意收集一堆数据就进行统计计算,往往是不可能得到可靠的结论的。

收集资料的方法很多,按形式分,有常规性、专题性;按范围分,有全面调查和非全面调查,而非全面调查中又可分为典型调查、重点调查和抽样调查等。

二、研究设计的基本内容及案例

研究设计的基本内容:

(1)研究目的。一般来说,一项研究工作不宜涉及过多的问题,最好只要求解决一两个问题,否则容易影响工作的质量。

(2)研究对象和人数,包括抽样的方法及分组的办法等。如果本研究课题是为了了解东盟各国留学生体质健康与生命质量状况,就应该进行问卷调查和体质测试,而且这两组的人数在实验前后应该大体相仿。

(3)研究时间和方法。包括研究工作开始和预计结束的时间,在整个研究时间内是否划分成几个阶段,研究工作开始时搜集哪几项统计指标,研究过程中每隔多少时间再次搜集哪几项统计指标,对这几项指标用什么方法、什么仪器设备进行测试,也应同时定出详细的测试细则。应该特别重视指标的设计,要使每一项指标所测得的结果都能比较客观、真实地反映出测试对象某一方面的情况。

(4)各种原始记录表格的设计、制定。

(5)初步规定数据的整理、统计分析方法。

【案例 2-1-1】广西民族大学龙舟队是全国大学生龙舟协会发起会员之一。目前龙舟已从以技术取胜的民间节庆体育项目逐渐发展成为一项以肌肉工作能力为主的力量

型民族体育竞赛项目。为了从理论到实践对我国高校的高水平龙舟训练提供理论依据,笔者对广西民族大学女子龙舟队进行 8 个专项力量测试:2min 20kg 卧拉(次)、25kg 卧推(次)、俯卧撑(次)、800m 跑(s)、40kg 负重下蹲(次)、50kg 卧蹬腿(次)、30kg 拉铃高翻(次)、15kg 杠铃屈伸(次)。采用因子分析法从专项力量项目提取影响划水功率(运动成绩)的共性因子,解释共性因子的含义,建立划水功率和共性因子回归模型,以期达到对运动训练负荷的有效控制和比赛成绩的预测。其测试成绩及 SPSS 数据文件见表 2-1-1、图 2-1-1。详见第二章第五节“龙舟运动员陆上专项力量与划水功率因子回归分析”。

表 2-2-1 广西民族大学高水平龙舟队 23 名女队员专项力量指标一览表

序号	姓名	性别	卧拉 (次)	卧推 (次)	俯卧撑 (次)	800m 跑 (s)	划水功率 (W)	负重下蹲 (次)	卧蹬腿 (次)	拉铃高翻 (次)	杠铃屈伸 (次)
1	胡春艳	女	65	19	19	165	85	16	25	16	15
2	龙珍香	女	56	10	17	164	83	17	25	15	16
3	陈家惠	女	48	8	14	160	75	14	24	14	14
4	卓丽娟	女	69	24	20	158	85	18	28	15	16
5	张 艺	女	44	14	13	164	82	14	23	14	12
6	谭晓媚	女	58	12	16	163	84	15	27	15	16
7	梁凤兰	女	59	13	19	153	82	16	25	16	16
8	骆理慧	女	50	18	24	152	86	19	29	17	20
9	张 丽	女	46	16	20	162	82	15	25	15	17
10	朱洁梅	女	60	12	14	167	81	17	24	15	17
11	粟纬纬	女	56	10	13	160	80	17	29	17	14
12	陈声丽	女	60	12	16	162	86	19	29	16	17
13	满丹丽	女	54	14	15	155	83	18	29	17	15
14	黄美莲	女	58	25	18	155	84	15	27	15	16
15	农素娟	女	50	10	15	156	80	15	25	15	15
16	吕 梅	女	54	19	15	160	79	15	26	15	15
17	李连囊	女	61	15	14	159	80	17	27	17	15
18	钟海韵	女	50	9	14	161	77	15	25	15	14
19	陆 丹	女	59	23	17	149	85	19	29	16	16
20	刘金惠	女	40	15	12	165	79	14	26	12	12
21	劳纯丹	女	52	10	14	155	75	13	26	13	14
22	莫兴梅	女	41	17	12	171	76	14	24	12	12
23	何东兰	女	40	12	12	178	75	12	22	10	12

女龙丹队.sav [数据源5] - SPSS Statistics 数据编辑视图											
文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 图形(G) 宏程序(M) 附加内容(O) 窗口(W) 帮助											
	名称	类型	宽度	小数	标签	值	缺失	列	对齐	度量标准	
1	姓名	字符串	8	0	女子龙丹队	无	无	6	居 中	名义	
2	性别	字符串	6	0	女	无	无	4	居 中	名义	
3	测试时间	日期	10	0		无	无	10	右(F)	度量(S)	
4	卧推	数值(N)	6	0	2 min 20kg卧推...	无	无	4	居 中	度量(S)	
5	卧推	数值(N)	8	0	25kg卧推(次)	无	无	4	居 中	度量(S)	
6	俯卧撑	数值(N)	8	0	俯卧撑(次)	无	无	8	居 中	度量(S)	
7	800 m	数值(N)	8	0	800 m(秒)	无	无	6	居 中	度量(S)	
8	划水功率	数值(N)	8	0	每划平均划水功	无	无	8	居 中	度量(S)	
9	负重下蹲	数值(N)	8	0	40kg负重下蹲...	无	无	8	居 中	度量(S)	
10	卧蹬腿	数值(N)	8	0	50kg卧蹬腿(...)	无	无	8	居 中	度量(S)	
11	拉铃高翻	数值(N)	8	0	30kg拉铃高翻...	无	无	8	居 中	度量(S)	
12	杠铃屈伸	数值(N)	8	0	15kg前屈伸(...)	无	无	8	居 中	度量(S)	

图 2-1-1

【案例 2-1-2】当前,中国足球“赌球”现象日趋严重,如何通过统计学原理来分析各队是否有打假球现象?针对 2009 年赛季中超联赛 16 支参赛队:北京国安、长春亚泰、河南四五老窖(建业)、山东鲁能、上海申花、天津泰达(康师傅)、成都建工地产(谢菲联)、大连实德(万达)、广药白云山、江苏舜天、深圳红钻(健力宝)、陕西中新、青岛中能、长沙金德、杭州绿城、重庆诗仙太白(力帆),通过中国足球超级联赛官网数据库,收集 2009 年赛季中超联赛 16 支参赛队的进球数 X_1 、场均控球率 X_2 、射门数 X_3 、角球数 X_4 、前场30m任意球数 X_5 、界外球数 X_6 、犯规数 X_7 、黄牌数 X_8 、攻入前场30m数 X_9 、传球成功率 X_{10} 、抢断数 X_{11} ,共 11 项数据指标,并建立 SPSS 数据文件,见图 2-1-2、图 2-1-3。运用原指标对各队技、战术结构进行评价、排名。通过统计学处理各队真实实力,建立评价体系,运用统计学对赛前训练、比赛成绩等因素裁定运动队是否打“假球”。详见第二章第五节“2009 年赛季中国足球超级联赛 16 支参赛队技、战术结构差异多元统计分析”。

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 图形(G) 宏程序(M) 附加内容(O) 窗口(W) 帮助											
	名称	类型	宽度	小数	标签	值	缺失	列	对齐	度量标准	
1	参赛队	字符串	10	0	参赛队	无	无	10	靠左(L)	名义	
2	赛季名次	数值(N)	10	0	赛季名次	无	无	11	靠右(R)	名义	
3	得分	数值(N)	10	0	得分	无	无	11	靠右(R)	名义	
4	进球	数值(N)	10	0	进球	无	无	11	靠右(R)	名义	
5	失球	数值(N)	10	0	失球	无	无	11	靠右(R)	名义	
6	净胜球	数值(N)	10	0	净胜球	无	无	11	靠右(R)	名义	
7	场均控球率	数值(N)	10	4	场均控球率	无	无	11	靠右(R)	名义	
8	射门	数值(N)	10	0	射门	无	无	11	靠右(R)	名义	
9	角球	数值(N)	10	0	角球	无	无	11	靠右(R)	名义	
10	前场30米任意球	数值(N)	10	0	前场30米任意球	无	无	11	靠右(R)	名义	
11	界外球	数值(N)	10	0	界外球	无	无	11	靠右(R)	名义	
12	犯规	数值(N)	10	0	犯规	无	无	11	靠右(R)	名义	
13	越位	数值(N)	10	0	越位	无	无	11	靠右(R)	名义	
14	黄牌	数值(N)	10	0	黄牌	无	无	11	靠右(R)	名义	
15	红牌	数值(N)	10	0	红牌	无	无	11	靠右(R)	名义	
16	攻入前场30	数值(N)	10	0	攻入前场30米数	无	无	11	靠右(R)	名义	
17	传球成功率	数值(N)	10	4	传球成功率	无	无	11	靠右(R)	名义	
18	传球助攻球数	数值(N)	10	0	传球助攻球数	无	无	11	靠右(R)	名义	
19	头球助攻率	数值(N)	10	4	头球助攻率	无	无	11	靠右(R)	名义	
20	抢断数	数值(N)	10	0	抢断数	无	无	11	靠右(R)	名义	

图 2-1-2

Table with 20 columns: Team, Player Name, Points, Rebounds, Assists, Steals, Blocks, Turnovers, Free Throws, 3-Point Shots, Field Goals, Points in Paint, Points Outside Paint, Fouls, Double Doubles, Red Cards, Points in the 3rd Quarter, Points in the 4th Quarter, Points in the 5th Quarter, Points in the 6th Quarter, Points in the 7th Quarter. Rows 4-16 show player statistics for various teams.

图 2-1-3

【案例 2-1-3】为了研究广西边远聚居少数民族学生体质与健康发展状况,从广西民族大学学生体质研究中心调取 2000 ~ 2010 年 13.6 万多名各民族大学生(包括汉族)的监测数据,并从中调取 26399 名来自壮族、侗族、瑶族、苗族、毛南族、京族、水族、仫佬族、仡佬族和土家族等少数民族大学生体质测试资料,建立 SPSS 数据文件。结果见图 2-1-4 和图 2-1-5。详见第二章第五节“广西边远聚居少数民族学生体质特征因子分析”。

Table with 20 columns: Item, Unit, Type, Height, Weight, Chest, Arm, Leg, Hip, Waist, Neck, Shoulder, Elbow, Wrist, Ankle, Heel, Arch, Ball, Toe, Claw. Rows 4-25 show various physical fitness test items and their units.

图 2-1-4

Table with 20 columns: Year, Semester, Class, Student ID, Ethnicity, Gender, Birth Date, Height, Weight, Chest, Arm, Leg, Hip, Waist, Neck, Shoulder, Elbow, Wrist, Ankle, Heel, Arch, Ball, Toe, Claw. Rows 1-20 show student physical fitness test results.

图 2-1-5

【案例 2-1-4】为了探讨东盟各国留学生体质健康与生命质量状况,为中国与东盟各高校文化交流和发展提供建议,通过专题立项对东盟七国留学生体质及生命质量结构特征进行实验研究,对东盟各国 800 名留学生进行心境量表问卷调查,调查 6 个一级指标、40 个问题(二级指标),并建立问卷 SPSS 数据文件,见图 2-1-6。可见《广西大学生与东盟留学生心理结构特征分析》,中国学校卫生,2012(2)。

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 图形(G) 实用程序(U) 附加内容(O) 窗口(W) 帮助(H)												
	名称	类型	宽度	小数	标签	值	缺失	列	对齐	度量标准		
1	姓名	字符串	8	0	姓名	无	无	8	居中	名义		
2	国籍	字符串	8	0	国籍	(1, 中国)	无	8	居中	名义		
3	民族	数值(N)	8	0	民族	(1, 汉族)	无	8	居中	名义		
4	性别	数值(N)	8	0	性别	(1, 男)	无	8	居中	度量(S)		
5	年龄	数值(N)	8	0	年龄	无	无	8	居中	度量(S)		
6	紧张的	数值(N)	8	0	紧张的	(0, 几乎没有)	无	8	居中	度量(S)		
7	生气的	数值(N)	8	0	生气的	(0, 几乎没有)	无	8	居中	度量(S)		
8	无精打采	数值(N)	8	0	无精打采	(0, 几乎没有)	无	8	居中	度量(S)		
9	不快乐的	数值(N)	8	0	不快乐的	(0, 几乎没有)	无	8	居中	度量(S)		
10	轻松愉快	数值(N)	8	0	轻松愉快	(0, 几乎没有)	无	8	居中	度量(S)		
11	慌乱的	数值(N)	8	0	慌乱的	(0, 几乎没有)	无	9	居中	度量(S)		
12	为等地	数值(N)	8	0	为等地	(0, 几乎没有)	无	8	居中	度量(S)		
13	心烦意乱	数值(N)	8	0	心烦意乱	(0, 几乎没有)	无	8	居中	度量(S)		
14	气坏的	数值(N)	8	0	气坏的	(0, 几乎没有)	无	8	居中	度量(S)		
15	劳累的	数值(N)	8	0	劳累的	(0, 几乎没有)	无	8	居中	度量(S)		
16	悲伤的	数值(N)	8	0	悲伤的	(0, 几乎没有)	无	8	居中	度量(S)		
17	精神饱满的	数值(N)	8	0	精神饱满的	(0, 几乎没有)	无	8	居中	度量(S)		
18	集中不了注	数值(N)	8	0	集中不了注意力	(0, 几乎没有)	无	8	居中	度量(S)		
19	自信的	数值(N)	8	0	自信的	(0, 几乎没有)	无	8	居中	度量(S)		
20	内心不安的	数值(N)	8	0	内心不安的	(0, 几乎没有)	无	8	居中	度量(S)		
21	气恼的	数值(N)	8	0	气恼的	(0, 几乎没有)	无	8	居中	度量(S)		
22	筋疲力尽的	数值(N)	8	0	筋疲力尽的	(0, 几乎没有)	无	8	居中	度量(S)		
23	沮丧的	数值(N)	8	0	沮丧的	(0, 几乎没有)	无	8	居中	度量(S)		
24	主动积极的	数值(N)	8	0	主动积极的	(0, 几乎没有)	无	8	居中	度量(S)		

图 2-1-6

三、抽样原则

为了获得对总体有较好代表性的样本,并排除选定样本时的主观偏见,一定要按“随机化”原则抽样。常用的随机抽样方法有以下四种。

(一) 简单随机抽样(纯随机抽样)

即从总体单位中不加任何分组、排队,完全随机地抽取调查单位。

随机抽选可有各种不同的具体做法,如直接抽选法,抽签法(抓阄法),随机数码表法。

简单随机抽样的优点:在抽样过程中完全排除主观因素干扰,简单易行。其缺点:只适应总体样本数量不大的情况。如全校有 800 个男学生,要从其中抽出 100 人,就把男学生编成 1~800 号,然后抽签,把抽出的 100 个号所代表的学生组成样本。

(二) 类型抽样(分层抽样)

先对总体各单位按一定标志加以分类(层),然后再从各类(层)中按随机原则抽取样本,组成一个样本。

分层抽样的好处:(1)样本代表性高,抽样误差小,抽样调查成本较低。如果抽样误差的要求相同的话则抽样数目可以减少。(2)分层抽样除了可以对总体进行估计外,还