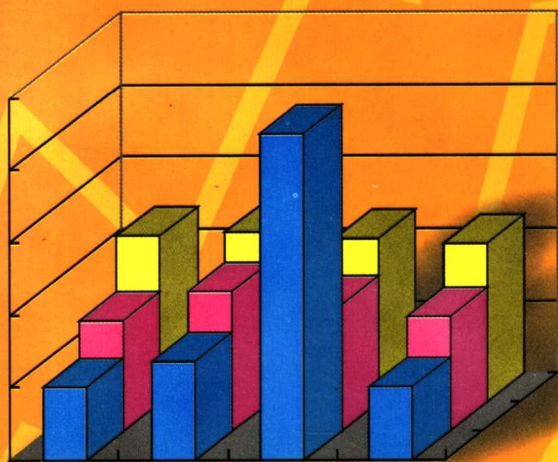


体育统计

史进 李旭芝 主编



西安地图出版社

体 育 统 计

史 进 李旭芝 主编

西安地图出版社

图书在版编目 (CIP) 数据

体育统计/史进, 李旭芝主编. —西安: 西安地图出版社, 2003. 9

ISBN 7-80670-453-1

I. 体... II. ①史... ②李... III. 体育统计

IV. G80-32

中国版本图书馆 CIP 数据核字 (2003) 第 072419 号

体育统计

史进 李旭芝 主编

西安地图出版社出版发行

(西安市友谊东路 334 号 邮政编码 710054)

新华书店经销 西安浪涛印刷有限责任公司印刷

850×1168 毫米·1/32 7 印张 169 千字

2003 年 9 月第 1 版 2003 年 9 月第 1 次印刷

印数 0001-1000

ISBN 7-80670-453-1/G·58

定价: 12.80 元

前 言

21 世纪是信息化时代,统计学是进行信息处理的有效方法,备受各界人士青睐。体育统计是统计方法在体育领域中的应用学科,已经作为我国体育科研的重要方法之一,被广泛应用于体育的各个领域。目前体育统计已成为体育院校重要的专业基础课。本书主要是供运动系、武术系、函授班及高职高专学生使用的体育统计课程教材。根据课时及学生的情况,结合多年来的教学实践经验,本书在着重介绍体育统计基本原理和方法的基础上,还增添了论文实例,突出教材的实用性,引导、启发学生参与体育科研的意识,提高学生独立认识、思考问题的能力。

本书共十二章,第一章到第九章是基本统计原理和方法的介绍,第十章是统计软件包的介绍和操作,第十一章是典型论文示例分析,第十二章是练习题汇编。

全书由史进、李旭芝主编,负责全书的编写及修改。在编写过程中承蒙程致屏教授、权德庆教授、雷福民副教授、张民朝副教授的指导,在此表示感谢。

本教材虽然经过多次修改,但是限于各种原因,错误难免,敬请读者指正。

编者

2003.5.21

目 录

第一章 绪论	(1)
第一节 概述	(1)
一、体育统计研究对象	(1)
二、随机现象、统计、统计规律	(2)
三、统计分析、变量	(2)
四、统计指标、统计数据	(3)
五、体育统计的主要内容	(3)
第二节 体育统计分析方法的介绍	(4)
一、统计描述方法	(4)
二、统计推断方法	(4)
三、高级统计方法	(4)
第三节 基本统计术语、统计符号	(5)
一、总体、样本、抽样研究	(5)
二、统计参数、统计量	(5)
三、常见统计符号	(6)
四、统计误差	(6)
第二章 统计资料的收集与整理	(8)
第一节 统计资料的收集	(8)
一、统计资料的来源	(8)
二、收集原始资料的方法	(9)
三、收集资料时应注意的问题	(10)
第二节 统计资料的整理	(11)

一、统计资料的检查	(11)
二、统计资料的分类与分组	(12)
三、编制频数分布表及频数直方图	(13)
第三节 研究设计	(17)
一、实验设计的内容	(17)
二、实验设计的原则	(17)
三、常见实验设计方法	(18)
四、实验设计示例	(20)
第三章 样本特征数	(22)
第一节 平均数	(22)
一、用直接法计算均数	(22)
二、加权平均数的算法	(23)
三、加权平均数的简捷算法	(25)
四、中位数的简捷算法	(27)
第二节 标准差	(29)
一、标准差的直接算法	(30)
二、标准差的加权算法	(32)
三、标准差的简捷算法	(33)
第三节 变异系数	(35)
第四节 相对数	(36)
一、常用相对数	(37)
二、相对数的标准化	(38)
第四章 概率基础知识	(42)
第一节 随机事件及随机事件的概率	(42)
一、随机事件	(42)
二、概率的定义	(43)
三、小概率事件	(45)

目 录

第二节 随机变量及其概率分布	(45)
一、随机变量	(45)
二、概率分布概念	(47)
三、离散型随机变量的概率分布	(48)
四、连续型随机变量的概率分布	(49)
第三节 二项分布	(50)
一、二项分布	(50)
二、二项分布在体育中的应用	(51)
第四节 正态分布及标准正态分布表	(52)
一、正态分布	(52)
二、标准正态分布	(53)
三、标准正态分布表	(54)
第五章 正态分布在体育中的应用	(57)
第一节 正态分布的通俗定义	(57)
第二节 利用正态分布估计实际分布情况	(59)
第三节 利用正态分布制定考核、考试标准	(62)
第四节 确定评价标准和考核标准	(64)
第五节 标准分(Z)	(66)
一、标准分	(66)
二、标准百分	(69)
第六节 百分位数	(69)
第七节 累进记分法	(72)
第六章 总体均数的区间估计	(81)
第一节 大样本的区间估计	(81)
一、样本均数的标准误、抽样误差	(82)
二、大样本资料总体平均数 μ 的区间估计方法	(83)
三、单侧区间估计	(85)

第二节 小样本资料的区间估计	(86)
一、 t 分布及 t 值表	(87)
二、小样本资料总体平均数 μ 的区间估计方法	(87)
第三节 率的区间估计	(88)
一、率的标准误	(88)
二、样本率的区间估计	(89)
第七章 假设检验	(91)
第一节 假设检验的概念	(91)
一、什么是假设检验	(91)
二、假设检验基本原理(小概率原则)的示例	(92)
三、假设检验的基本步骤	(94)
四、假设检验结论的认识	(95)
第二节 u 检验	(96)
一、单样本的 u 检验($\mu = \mu_0$)	(96)
二、两样本的 u 检验($\mu_1 = \mu_2$)	(98)
三、样本率的 u 检验($\mu_d = 0$)	(99)
第三节 t 检验	(101)
一、单样本 t 检验($\mu = \mu_0$)	(101)
二、两样本 t 检验($\mu_1 = \mu_2$)	(102)
三、配对资料 t 检验($\mu_d = 0$)	(104)
第四节 χ^2 检验	(107)
一、 χ^2 分布及 χ^2 表	(107)
二、单向分类 χ^2 检验	(109)
三、双向分类 χ^2 检验	(110)
四、正态性 χ^2 检验	(113)
第五节 方差齐性检验	(115)

第八章 方差分析	(119)
第一节 方差分析的意义	(119)
第二节 单因素多水平的方差分析	(120)
一、基本原理	(120)
二、基本步骤	(122)
三、具体举例([例 8.1]进行方差分析)	(126)
第三节 多重比较	(127)
第四节 单因素方差分析中应注意的问题	(129)
第五节 双因素方差分析	(130)
第九章 相关与回归	(136)
第一节 直线相关	(136)
一、直线相关的图像和类型	(136)
二、相关系数	(140)
三、相关系数 r 的显著性检验	(145)
四、运用相关系数 r 进行相关分析时应注意的问题	(146)
第二节 等级相关	(147)
第三节 一元直线回归	(150)
一、一元直线回归方程的建立	(152)
二、已知两样本统计量 $\bar{X}$, S 和 r , 建立直线回归方程	(154)
三、直线回归方程效果的检验	(155)
第十章 统计软件包在体育中的运用	(160)
一、统计软件包的介绍	(160)
二、运用 SPSS 绘制频数分布直方图	(160)
三、运用 Excel 进行配对样本的 t 检验	(162)
四、采用 SPSS 进行一元一次回归分析	(163)
第十一章 统计方法在体育实践中的应用示例	(166)

一、累进评分法、配对 t 检验在运动训练中的应用	(166)
二、 χ^2 检验在运动技术评价中的应用	(168)
三、回归分析在运动训练中的应用	(169)
四、方差分析在运动医学中的应用	(171)
第十二章 练习题汇编	(174)
附表 1 标准正态分布表	(181)
附表 2 t 标准正态分布表	(185)
附表 3 χ^2 值表	(188)
附表 4 F 值表(方差齐性检验用)	(190)
附表 5 F 值表(方差分析用)	(192)
附表 6 q 值表	(200)
附表 7 相关系数界值表	(202)
附表 8 等级相关系数界值表	(208)
附表 9 随机数字表	(210)
参考文献	(214)

第一章 绪 论

第一节 概 述

体育领域中有许多信息是通过大量的数据反映的,这些数据常常需要进行合理的处理,体育统计便是对数据进行处理的科学和有效方法之一。学习体育统计可以帮助我们更好地分析和认识问题,更好地解决体育实际问题。

一、体育统计研究对象

体育统计已成为体育科研不可缺少的重要工具,在体育领域中被广泛应用。体育统计是数理统计方法在体育领域中的应用,它以体育运动中随机现象的规律性为研究对象,为我们提供实验设计方法和收集、整理、分析数据资料的方法。

在体育领域存在着大量的随机现象,需要我们去认识。例如,我们经常进行在校学生运动成绩的测试,要评判他们的身体运动能力,必须依据合理的等级标准。运动成绩是随机现象,学生的运动成绩存在个体间的差别,但通过对大量成绩(数据)的观察,我们认为学生成绩的分布是有一定规律的,根据概率分布理论,能够制定出合理的等级评判标准,通过等级标准客观准确地反映学生的运动能力。

二、随机现象、统计、统计规律

客观现象可以分为确定性现象、非确定性现象。非确定性现象就属于随机现象,是指在某种条件下可能发生也可能不发生的现象。例如,观察天气的变化、运动成绩、新出生婴儿的性别等等,这些都是随机现象。这些现象都有以下特征:它们的结果具有多样性,且哪种结果会出现事先具有不确定性。例如,对新生儿性别的观察,结果多样性分为男性或女性;不确定性表现为婴儿出生前不能肯定是男性或是女性。在统计研究中,我们把对随机现象的观察就称为随机实验,观察取得的数据称为观察值或测试值。

统计是一个把信息变成知识,使知识有利于社会、科学进步的过程。这里的信息就是根据观察取得大量随机现象的统计数据得来的,可以讲统计也就是通过大量数据揭示随机现象的内在变化规律,增进对现象的深入认识。这种规律就是统计规律,具有普遍意义。

例如,抛一枚硬币的随机实验中,结果有两种,即出现“正面朝上”或“反面朝上”。每次结果的出现是不确定的,可能出现“正面朝上”,也可能出现“反面朝上”。但通过成千上万次的抛掷,我们发现“正面朝上”、“反面朝上”出现的可能性一样,即抛掷一枚硬币出现“正面朝上”或“反面朝上”结果的可能性均为 $\frac{1}{2}$ 的统计规律。那么在我们抛一枚硬币前就可以认为,出现“正面朝上”或“反面朝上”机会是均等的。

三、统计分析、变量

统计分析是现代科学的组成部分,是采取统计方法对问题进行认识的过程。统计分析的最基本要素是变量,即观察研究对象的

特征或测量计数的结果。变量可分为连续型、离散型。

例如,在对运动比赛的研究中,运动成绩及运动员的特性是我们感兴趣的变量,可能是跳高、百米、铅球、赢输、性别、种族等。跳高、百米、铅球等是连续型变量,这些测试结果,都带有度量衡单位,赛前无法一一列举,只能在区间取值。赢输、性别、种族是离散型变量,是具体的观察结果,没有度量衡单位,但比赛前可以一一列举,是点的取值。

四、统计指标、统计数据

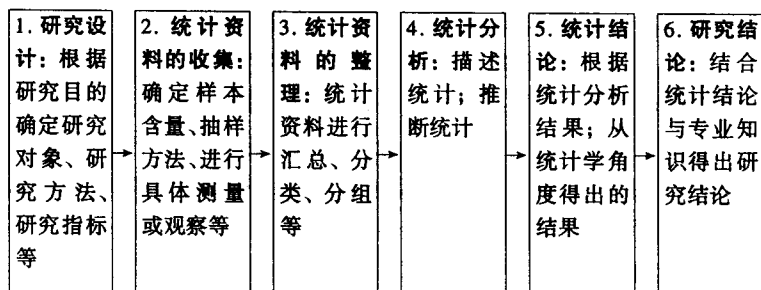
统计指标是反映现象数量特征的概念和具体数值。例如,1982年7月1日零时我国大陆人口为1008175288人;1957年我国已拥有体育场地18191个;2000年我国人均GDP已达806美元。这些都是统计指标,具有时间、地点、对象及数值的特征。

统计数据是研究对象同变量的测量值或观察值的集合,是非个人的,常常可以形成一个分布,进行统计处理。例如,研究某地区中学生的耐力运动水平,选择1500米项目进行测试,那么1500米项目的统计数据是指该地区所有或部分中学生的1500米成绩的集合,它们可以形成一个变量分布;而某一个人的成绩并不构成统计数据。

五、体育统计的主要内容

体育统计就是运用统计分析的方法研究、解决体育领域中的实际问题,为体育科研服务。迄今体育统计方法已成为重要的体育科研方法,广泛应用于体育教学、运动训练、人体科学、体育社会学、运动心理等方面的研究。

体育统计的主要内容有:研究设计,资料收集、整理、分析预测。体育统计的具体运用过程概括如下图:



第二节 体育统计分析方法的介绍

统计分析方法是进行统计处理的手段,统计描述、统计推断是最基本的统计分析方法,多元回归是高级的统计分析方法。

一、统计描述方法

统计描述是对统计资料进行整理的统计分析方法,包括分类、分组、特征量及图表的形式。统计描述简单易行,特征明显,直观易懂。本书第二、三章就是统计描述的内容。

二、统计推断方法

统计推断是通过样本估计推断总体的统计分析方法。概率论是推断理论的基础。本书中统计推断包括参数估计、假设检验及相关回归的内容。它是统计分析的常见方式,也是重点难点内容。本书第四、五、六、七、八、九章就是统计推断的内容。

三、高级统计方法

随着统计方法的不断丰富、发展,数学方法的应用日益增多,例

如线性代数、线性规划及模糊数学等高级计算方法大量应用于统计学,形成线性回归分析、聚类分析、判别分析、因子分析等高级统计方法。鉴于教学对象及课时的限制,本教材不对该内容进行介绍。

第三节 基本统计术语、统计符号

体育统计有其独特的研究问题的方法,在统计处理过程中采用了大量的统计符号和统计术语,它们是进行统计研究的基本概念。

一、总体、样本、抽样研究

总体是指根据研究目的确定的同质研究对象的特征全体。通常总体很难得到。

样本是按一定的原则,随机地从总体中抽出部分个体所构成的。

抽样研究是通过样本特征研究总体特征的方法。通常总体难以取得,样本容易得到,客观上要求我们经常用样本反映总体。

例如,根据国家需要,要求了解西安市青少年的体质状况,我们常常进行的就是抽样研究。该研究对象为西安市青少年,特征为他们的体质指标,即研究的总体为西安市所有青少年的体质特征数据,由于时间和条件限制我们很难取到总体,但可以从总体中随机抽取 3000 人作为研究对象的样本,观察测量他们的体质状况,通过 3000 人的体质样本数据反映全市总体情况。

二、统计参数、统计量

统计参数是描述总体数量特征的指标。总体通常难以取得,因

而统计参数也不易获得。统计参数是确定的数值,但常常是未知数。统计量是描述样本数量特征的指标。样本可以随机获取,不同的样本会得到不同的统计量,因而统计量是不确定的变量。例如,假定某年级学生总体身高平均数为 172 厘米,属于参数,从中第一次抽取 30 人,得到身高样本平均数为 173 厘米,属于统计量;第二次再抽取 30 人,得到的样本平均数是 171 厘米,也属于统计量,它与第一次的统计量不同。

三、常见统计符号

1. 求和符号 Σ (读作西格玛)

$$X_1 + X_2 + X_3 + \cdots + X_n = \sum_{i=1}^n X_i \text{ 或简写为: } \sum X$$

例如, $X: 10, 20, 30, 40, 50$, 在这里 $i = 1, 2, 3, 4, 5$, 则

$$\sum_{i=1}^n X_i = 10 + 20 + 30 + 40 + 50 = 150。$$

2. 常见统计参数、统计量

μ —— 总体平均数,	$\bar{X}$ —— 样本平均数,
σ —— 总体标准差,	S —— 样本标准差,
π —— 总体率,	P —— 样本率,
ρ —— 总体相关系数,	r —— 样本相关系数。

四、统计误差

观察值与真值的差就是统计误差。根据误差形成的原因,统计误差可以分为:

1. 随机误差

(1) 随机测量误差 由于各种偶然因素的影响造成同一对象

在相同条件下多次测量结果却不完全一致,这种差异不固定,有时高有时低,这就是随机测量误差。在测量样本较大时,随机测量误差会相互抵消。通常我们不考虑此误差。

(2) 抽样误差 在一定条件下,由于个体之间存在差异,抽取的样本与总体之间就会产生误差,这种误差就是抽样误差。要进行抽样研究,抽样误差就会存在。随着样本数量的增大,抽样误差就会减小。抽样误差通常表现为统计量与统计参数的差异。

2. 系统误差

在收集资料时,由于测量工具、观察环境等外在条件的变化,造成测量值的偏大或偏小,与真值产生误差,这种误差就叫系统误差。系统误差产生的原因是可以观察、分析出来的。

3. 非系统误差

在登记、记录数据时,由于记录者的大意或过失,造成观察值与真值之间的差异,为非系统误差。只要工作认真,这类误差是可以避免的,通常我们不去考虑此类误差。

练习题(1)

1. 注意观察了解社会、生活、学习中遇到的统计处理方法。
2. 试述统计分析方法的内容。
3. 什么是总体?什么是样本?
4. 什么是抽样研究?
5. 什么是统计量?什么是统计参数?