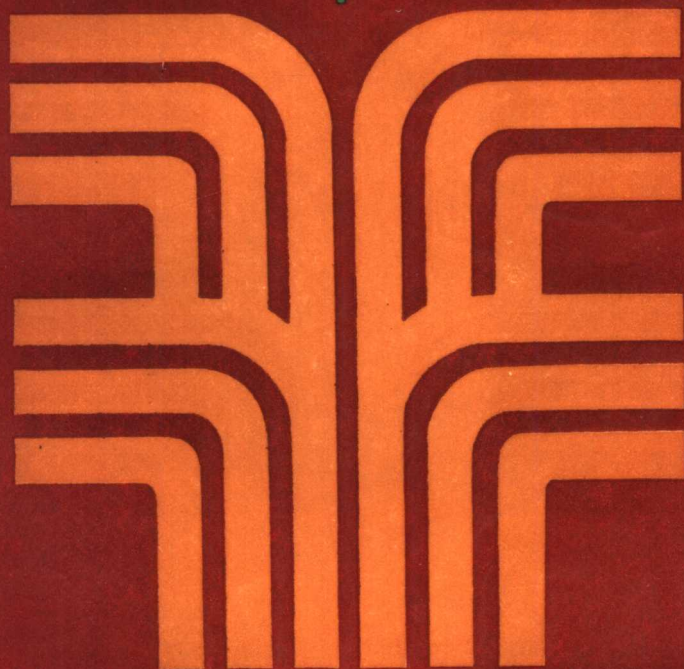


实用 体育统计

祁国鹰
徐明
张明立
编著



2

北京体育大学出版社

实 用 体 育 统 计

祁国鹰 徐 明 张明立 编著

北京体育大学出版社

责任编辑:佟 晖

审稿编辑:宇 奕

责任校对:未 茗

责任印制:长 立 张玉萍

图书在版编目(CIP)数据

实用体育统计/祁国鹰等编著. —北京:北京体育大学出版社,1995

ISBN 7-81003-927-X

I. 实… I. 祁… III. 体育统计—教材 IV. G80-32

中国版本图书馆 CIP 数据核字(94)第 16361 号

实用体育统计

祁国鹰 徐明 张明立 编著

北京体育大学出版社出版
(北京西郊圆明园东路 邮编:100084)

新华书店总店北京发行所发行
北京佳顺印刷厂印刷

开本:850×1168 毫米 1/32
1995 年 5 月第 1 版

印张:7.625
1997 年 8 月第 2 次印刷

定价:9.50 元
印数:6500 册

ISBN 7-81003-927-X/G·714

(本书因装订质量不合格本社发行部负责调换)

前 言

体育统计是用数理统计方法研究体育问题,探索体育运动规律的一门学科。它是广大体育工作者都必须掌握的一种重要的科学工具。

随着体育运动不断向现代化和科学化深入发展,体育统计已日益普及至体育教学、训练、科研、管理等各个方面,并正在发挥着重要的作用。国内各体育院校和师范院校的体育系(科),为适应当前这种形势发展的需要,已把体育统计列为必修课程,并编有适用于研究生和大学本科生学习的体育统计教材。但适用于体育大专生学习用的比较系统的体育统计教材,迄今尚不多见。本书即是针对这一需要而编写的,是一本带有探索性的教材,目的是想通过总结多年来的教学经验,编写适合于体育大专这个层次学生使用的体育统计教材,使学生们在学习时少遇到一些困难,多感到一些乐趣,并在学习后能初步掌握一些必要的体育统计方法,能独立分析资料和处理数据。

从国内各体育院校历届所招大专学生的组成和文化程度看,他们共同的特征是体育运动专业技术水平较高,运动经验比较丰富,求知心很迫切,但数理基础比较薄弱。他们来校学习的共同目的是想通过花费不太长的时间和精力,能基本掌握一些常用而必要的体育统计方法,能独立分析资料和处理数据,能科学地安排试验和制定抽样调查方案,并可进而能对他们所关心和研究的专业问题作出具有一定精确度的描述、判断和预测。针对这一实际情况,本书重点在于介绍一些常用而必要的体育统计方法,对每种方法又着重于通俗地阐明方法所依据的原理、应用的条件和途径。虽然书中有时也对一些定理和公式进行了证明与推导,但主要目的

不仅在于要通过证明和推导得到有关结论的本身,更重要的是要通过问题的证明和推导,使学生们加深对概念与结论的理解。此外,在每一节中附以示范性例题,以例题的形式体现本节的内容与具体要求,并从概念、计算、证明、应用等各方面对体育统计中的一些比较典型问题进行剖析,使学生们能通过示范性例题的学习,进一步掌握基本概念,提高分析问题、解决问题的能力。

统计分析计算的工作量比较大。结合国内当前微机应用十分普及情况考虑,在全书的最后就体育统计中常用的电算程序,单独编写了一章,并附有算例。

本书是按规定的 54 学时的授课时数编写的,如果学时较紧,或设备条件暂缺,有些章节可以备讲或者是有选择地讲授。

本书的第七、九、十章由徐明编写,其余各章节由祁国鹰编写。张明立副教授审阅了全书,对部分章节提出了修改意见,并进行了修改。在编写过程中,曾参考了国内外有关书籍和教材,汲取了各书中的编写经验,引用了其中的一些材料和统计图表,特在此谨向各书的编著者和出版者表示深切的感谢。

切合体育院校大专学生使用的体育统计教材如何编写,在当时来说确实还是一个需要认真研究探索的问题。由于我们水平有限,加之时间仓促,书中错误或不妥之处在所难免,务请专家和读者给予批评和指正。

编 者

1994 年

目 录

第一章 绪 论	(1)
第一节 体育统计的性质和作用.....	(1)
第二节 体育统计中应注意的几个问题.....	(5)
第三节 有关观测数值的一些概念和规定.....	(6)
习题一	(12)
第二章 抽 样	(13)
第一节 概率的基本概念	(13)
第二节 随机变量	(17)
第三节 总体与样本	(18)
第四节 抽样方法	(20)
第五节 抽样误差	(21)
第六节 抽样数目的确定	(28)
第七节 样本资料的整理	(31)
习题二	(34)
第三章 统计参数	(35)
第一节 算术平均数	(35)
第二节 中位数	(37)
第三节 标准差	(37)
第四节 离差系数	(39)
第五节 大样本资料的统计参数简捷算法	(41)
习题三	(45)
第四章 正态分布	(47)
第一节 正态分布曲线的形式	(48)

第二节	正态分布曲线的特征	(49)
第三节	正态分布表	(50)
第四节	统计资料的正态性检验	(54)
第五节	可疑数据的舍取	(57)
第六节	正态分布理论在体育中的两个应用	(59)
习题四		(63)
第五章	体育中的评分方法	(65)
第一节	分布位置百分	(65)
第二节	名次百分	(68)
第三节	标准百分	(68)
第四节	累进计分	(71)
习题五		(75)
第六章	总体参数估计	(76)
第一节	t 分布	(76)
第二节	t 值表	(77)
第三节	标准误	(79)
第四节	总体平均数的区间估计	(80)
第五节	总体比例数的区间估计	(83)
习题六		(85)
第七章	假设检验	(86)
第一节	基本概念	(86)
第二节	一个正态总体均数的假设检验	(91)
第三节	两个正态总体均数的假设检验	(97)
第四节	关于比率之间差别的假设检验	(104)
第五节	关于方差齐性的假设检验	(107)
第六节	方差不齐时两总体均数差别的假设检验	(110)
第七节	假设检验注意事项及说明	(111)
习题七		(115)

第一章 绪 论

第一节 体育统计的性质和作用

一、什么是体育统计

在生产实践和科学实验中,我们经常会搜集到许多数据,这些数据都是科学研究的宝贵资料,能为人们提供许多有用的信息。但这些信息往往不是一目了然,而是蕴藏在大量的数据之中。如何从这些庞杂的数据中获得更多的信息呢?数理统计为我们提供了许多有效的方法。

数理统计是以概率论为基础的应用数学的一个重要分支,是专门研究关于数据的搜集、整理、分析和推断的一门数学学科。它主要包括以下内容:数据整理和样本统计量的研究;统计推断;方差分析;回归分析;抽样理论;质量控制;试验设计。

数理统计方法是以概率论为基础,以样本为依据,运用数学模型来推断总体的一门科学方法。由于这种方法具有普遍意义,因此它能同各具体专业研究结合起来,在实践中得到广泛的应用。随着生产迅速发展,科学技术进一步深化和精密化,数理统计方法正日甚一日地渗透到几乎一切自然科学、技术科学、生物科学以及工业经济管理等领域中,并分别形成了工业统计、医学统计、生物统

计等门类学科。

体育运动中出现的现象(事件)大都是随机的,这种现象正是数理统计所要研究的对象,因此在体育运动中应用数理统计方法研究体育问题是十分必要的,也是数理统计能应用于体育运动领域的根本原因。体育统计就是运用数理统计的理论和方法,来研究体育教学、训练、科研和管理中的问题,探讨体育运动发展规律的一门体育学科。它对进行体育教学和科研都有很大帮助,因此所有从事体育工作的人员,特别对于即将走上体育工作岗位的学生,都应学习和熟练运用这门应用科学。

在体育教学、科研中应用体育统计方法是必要的,但体育统计不是万能的,仅靠它是不可能完全解决复杂的体育问题的,只有和体育运动专项理论、运动实践紧密地结合,通过对体育现象进行综合和系统的研究,才能得出比较满意的结果。因为体育统计方法毕竟是一种数学工具,并不能说明或表达所研究对象的本质,只有对它进行正确运用时,才能成为研究人员的强有力的助手。如果在实际工作中不分场合地盲目搬用,就会产生严重错误,甚至可能歪曲事实真象,这一点应特别引起人们足够的重视。

二、体育统计的内容

体育统计的内容,是随着历史的发展而不断发展的,大体言之,可分为描述统计、推断统计和试验或调查研究设计三大部分。

(一)描述统计

描述统计主要是将通过试验或调查所获得的大量数据,经过归纳、压缩、简化和整编后,从中找出这些数据的分布特征,制成图表,或借助这些数据的分布特征,如集中趋势、离中趋势、相关程度等,计算出一些具有概括性的统计数字,如平均值、标准差、相关系数等。借助于这些概括性数字,即可从庞杂无章的数据中获得具有实用意义的信息,以便对不同总体进行分析比较,做出合乎客观规

律的结论。

(二)推断统计

推断统计,是在正确的描述统计的基础上,应用数据所传递的信息,通过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范围,这种工作方法,就叫做推断统计。在实际工作中,人们常需进行这种由已知推论未知,由样本的性质推论总体的性质的统计推断工作。推断统计的主要内容是参数估计和假设检验。

(三)试验或调查研究设计

在对体育运动中某一问题进行专题研究前必须先进行试验或调查研究。在试验或调查研究开始前,对被试验或调查对象的选择、测定项目、控制因素、全部工作进行步骤、观测数据如何处理等问题,都要进行周密的研究,作出妥善安排与详细设计。

试验或调查研究设计,是进行科研工作的重要依据,是取得良好科研成果的重要条件,是处理试验或调查结果的先决条件。试验或调查研究设计的错误,会使全部科研失去意义,并造成人力、物力和时间的浪费。因此试验或调查设计的主要任务,是提供一个通过较少次数试验或调查研究,从而能获得比较全面准确的结论的良好完整的工作方案。

体育统计所包括的以上三个内容,组成一个完整的体育统计体系,它们之间是相互关联不可分割的,都必须做到正确无误。

三、为什么要学习体育统计

我国体育统计起步较迟,因此在许多体育科研问题的研究上,常是仅凭经验或主观意向作出结论。要使体育科研工作科学化、现代化,大力提倡学习和掌握体育统计理论与方法是很必要的。其理由是:

(一)便于提高体育教学、训练和科研水平

体育科学是一门综合性科学,有其自身独特的发展规律。在体育实际工作中,无论是从事教学、科研工作,还是从事行政管理工
作,要想使自己的工作或决策切合实际,符合客观规律,就必须学
习运用体育统计方法。它可以帮助人们分析资料,处理数据,科学
地安排系统观察、体育调查和体育实验,为评估体育效率,改进体
育设施,改革体育体制,提高体育教学、训练和科研水平,提供有效
的途径。因此在实际工作中有效地运用体育统计方法,就能够对经
纬万端的复杂事物,做到用较简单的方法掌握其中的规律,洞察要
害,发现矛盾,解决问题,不断前进。

(二)能培养体育科学工作者的科学思维能力与科学态度

体育统计是以数理统计为基础发展起来的一门体育学科,具
有数学的严谨性等特点,它可以锻炼人们的科学思维,学会科学的
推理及思考方法,形成实事求是地对待一切事物的唯物主义态度。
无论是描述统计,或是推断统计都是依据客观事实去做去伪存真,
由表及里的工作,也即是用客观的科学方法研究体育,使体育研究
不仅能发现他人未发现的事实,而且更重要的是让他人能得到同
样的事实。因此,通过学习和应用体育统计,能培养体育工作者的
科学思维和科学态度。

(三)有助于学习国内外体育方面的先进经验

当前国内体育运动水平提高很快,为了赶超国际水平,并促进
我国体育事业现代化,体育工作者都必须随时汲取新的知识,不断
学习国内外成功经验。而这些经验大都建立在大量试验或调查研
究的基础上,所有的数据也都是经过严格的数理统计方法处理的,
因此在它的说明和结论中,常包含有许多数理统计方面的内容。如
不熟悉体育统计,当然也就无法理解和从中汲取对我们有用的东
西。更无法从中发现问题和通过统计推断以鉴别其真伪性、可靠
性。国内外经验交流都是相互对等的,都必须具备一定的学术水

平,如不熟悉体育统计,就不可能把自己所从事的工作或科研成果完满地总结出来,也就无法与他人通过高水平的学术交流达到取长补短,互相学习,促进提高的目的。当然也就必然会影晌我国体育事业的快速发展。就从这一点讲,学习体育统计也是极为重要的。

第二节 体育统计中应注意的几个问题

实测数据资料,是体育统计分析计算的根木依据,因此在数据搜集选取时,必须注意以下几个问题。

一、资料的一致性

体育统计要求所有的资料必须是属于同一类型,并且是在同一条件下产生的,严禁选用性质不同的资料混杂在一起进行统计分析。如男、女学生身高统计资料,由于性别不同,因此这些数据资料不能放在一起统计分析。又如某专项运动队的当前运动成绩,不能和它三年前的成绩放在一起统计,原因是这些运动成绩产生的时间不同。如一定要放在一起分析时,则需经过适当处理后方可。

二、资料的代表性

统计分析的目的,是要以过去资料的变化规律推论今后发展情况,或是以现有一定数量的资料去推断所要研究对象的总体。因此对统计分析所使用的资料,一定要具有足够的代表性,即是说数据资料的含量不能太少。过少的资料推得的结果可能偏大或偏小,甚至会失去实际使用的意义。根据经验,一般要有 20—30 个以上的数据资料,始具有一定的代表性。

三、资料的可靠性与独立性

只有准确可靠的资料,才能得出正确的结论。因此对已搜集到的资料,在统计分析前一定要进行可靠性检查。检查的办法是把前后的资料联系起来,或把有关现象(资料表现出的)放在一起来看,检查它们是否有矛盾或写错与伪造之处。

此外在统计分析中,还要求资料是相互独立的,切不可把互相有关的资料统计在一起使用,这一点在统计分析中非常重要,不能等闲视之。

第三节 有关观测数值的一些概念和规定

一、变 量

由观测所得的数据,具有离散性、变异性和规律性三个特点。它常以一个分散的形式出现,并在一定的时空范围内变动。但它们并不是杂乱无章的变化,而是在一定范围内,呈现出一定规律性。由于观测数据具有变异性,因而在统计学中把它称为变量。如测量一个班级男生的身高,由于每个人的身高不同,因此所测得的身高数值也就不同。该班级男生的身高指标就是变量,它是以各个男生参差不同的数据形式出现的。与变量相反的是常量,如圆周率 π 和自然对数的底 e 都是常量,其值分别是3.1416和2.7183。

变量依其性质不同,可分为:

(一)称名变量

称名变量只说明某一事物与其它事物在属性上不同或类别上的差异,而不说明两事物间的差异的大小。如学生依性别可分为男生与女生;衣服依其颜色可分为红、黄、白等色。以上所说的学生的性别、衣服的颜色都是称名变量。

(二)顺序变量

顺序变量是指可就事物的某一属性的多少或大小按次序将各事物加以排列的变量。这种变量既不具有相等单位,也不具有绝对零点。如将学生按身材高低由高到低排成一列,最高的定为“1”级,次高的定为“2”级,其余的可依次类推,这就是身高顺序变量。因为这种变量不具有相等单位,所以“1”级与“2”级的身高差,并不等于“2”级与“3”级的身高差。

(三)等距变量

等距变量除表明量的大小外,还具有相等单位。气温度数是一种等距变量,如星期一、二、三的气温分别是 20°C 、 30°C 、 40°C ,则人们可以说星期三的气温高于星期一、二,星期一、二的温差等于星期二、三的温差,原因是其单位“度”是相等的。但绝对不能说星期三的天气就比星期一“热一倍”,原因是等距变量无绝对零点,不能作比率性的陈述。温度计上的刻度仅是一相对数值,零刻度并非绝对零点,即使在摄氏计的绝对零点(-273°C)时,也并不表示一点温度都没有的自然原点。

(四)比率变量

比率变量除表明量的大小、单位相等之外,还具有绝对零点。身高、体重等都属于这种变量,并允许作比率性陈述。如一成人体重 80 千克,一儿童体重 20 千克,如果说成人体重是儿童体重的四倍,这是完全正确的。

变量依其相互关系,可分为因变量与自变量,两者间关系可以 $y=f(x)$ 表示,意即 y 是 x 的函数。当两个变量的函数关系式已知时,即可以从自变量 x 求得相应的因变量 y 。自变量 x 有连续型变量与离散型变量之分。连续型变量是指在量尺上任何两点间,均可加以细分,并取得无限多大小不同的数值。如以千克做单位去称重,即使在 10 与 10.1 千克之间,也可读出无限多个微差值,如 10.03、10.036、10.039 千克等。离散型变量是指在量尺上任何两

点间,只能读取有限个数值,如在 2 与 4 间,只能读取一个整数 3,不可能再有其它整数。如一个班级的学生中的田径运动员人数可能是 1、2、3 或更多个,但绝不可能是 1.3 或 2.7 个。

二、数据简化与尾数舍入

凡原数据的准确度超过所要研究问题的需要者,都可以加以简化与舍入尾数。如时间测记值准确到十分之一秒(如 12.3 秒和 12.4 秒),而要研究的问题只要求准确到秒,这时上述括号中两个数值,即可以四舍五入方法舍去尾数,而改用秒的约整值 12 秒,这种处理方法就称为数的约整化。

设一数为 8.6354,如欲将它分别约整成两位小数、一位小数、最接近的整数,则该数即可分别写成 8.64、8.6、9。舍入尾数的规则是:

- (一)若需舍入的尾数大于 5 时,将前一位数加 1;
 - (二)若需舍入的尾数小于 5 时,前一位数字保持不变;
 - (三)若需舍入的尾数恰为 5,则前一位数字为奇数时,将前一位数字加 1;若前一位数字为偶数时,则前一位数字保持不变。
- 循此规则,相应于各记录值的约整值如下,供读者参考。

记 录 值	保留小数位数的约整值		
	三位小数	二位小数	一位小数
6.5648	6.565	6.56	6.6
7.1650	7.165	7.16	7.2
7.1644	7.164	7.16	7.2
0.4978	0.498	0.50	0.5

三、有效数字

仪表上的读数部分,总是分刻成许多小格,测量时只能读到这些小格中的分数,包括这个分数在内的读数值,一般认为是有效的,称为有效数字。

如量尺上最小一格代表 1 毫米,而读得的数是 12.5 毫米,则 12 是可靠的,因有格子刻度标志。后面的 0.5 是估计的,虽不很准确,但仍有意义,称为估计数字。如认为估计数字的估计偏差不大于 0.2,则上述读数可写成: 12.5 ± 0.2 毫米,意即此测定值的真值是在 12.3 到 12.7 毫米之间,其变动范围在 0.3 到 0.7 毫米之间,变幅是 0.4 毫米。一般规定记录测定值时,只保留一位估计数字。

在观测值的前面或末尾出现零时,若不特别注意,容易产生错误。如在上述条件下读得的测定值恰好是 12 毫米,这时记录时仍应记为 12.0 毫米,不应记为 12 毫米。前者表示小数点前两位是准确值,后一位“0”是估计值。而后者 12 毫米则会被人误认为数字 2 是估计值。因此必须记住,最后一位数字“0”在测定值的表示中不能略去,否则会导致发生错误。对于 0.0120 米这个数值,前面两个零都不是有效数字,是因单位改变(由毫米改为米)而产生的,后面一个 0 仍是有效数字。如测得两地的距离为 25720 米,若由于仪器原因只能准确到百位数,而十位数是可疑数,此时用 25720 这种数字形式来表示观测结果是不可以的,它会使人误解为十位数是准确的,个位数才是可疑的,改写成 25.72×10^3 米,则整个数字形式既表达了两地的距离, 10^3 前面的数值 25.72 也表明了有效数字的位数。下面的说明有助于掌握有效数字的涵义。

测定值读数	356	3560	0.356	0.3560	4.003050
有效数字位数	3	4	3	4	7