

体育统计简明教程

(2009 年修订版)

祁国鹰 徐明 周星宇 编著

北京体育大学出版社

体育统计简明教程

(2009 年修订版)

祁国鹰 徐 明 周星宇 编著

北京体育大学出版社

策划编辑 张明庭 梁 林
责任编辑 梁 林
审稿编辑 李 飞
责任校对 木 凡
责任印制 陈 莎

图书在版编目(CIP)数据

体育统计简明教程/祁国鹰等编著. -3版. -北京:北京
体育大学出版社,2009.6
ISBN 978-7-81100-078-8

I. 体… II. 祁… III. 体育统计-成人教育:高等教育
-教材 IV. G80-32

中国版本图书馆 CIP 数据核字(2009)第 091637 号

体育统计简明教程

祁国鹰 等编著

出 版 北京体育大学出版社
地 址 北京海淀区信息路 48 号
邮 编 100084
邮 购 部 北京体育大学出版社读者服务部 010-62989432
发 行 部 010-62989320
网 址 www.bsup.cn
印 刷 北京市昌平阳坊精工印刷厂
开 本 787×960 毫米 1/16
印 张 13

2010 年 8 月第 3 版第 2 次印刷 印数 2501-5000 册

定 价: 26.00 元

(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应用学科。它是广大体育工作者都必须学习和掌握的一种重要的体育科研工具。

随着体育运动不断向现代化和科学化深入发展,体育统计已经日益普及至体育教学、训练、科研、管理等各方面并发挥着重要的作用。国内体育院校为适应当前形势发展的需要,已将体育统计列为必修课程。本书从体育专业学生的实际出发,在内容选择和编排上遵照“淡化系统,突出实用”的观点探索性地编写此书,目的是能够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。

本书重点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基本原理、应用条件和途径。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,从概念、计算、应用等各方面对体育统计中的一些比较典型问题进行分析,使学生们能通过对示范性例题的学习,进一步掌握基本概念、提高分析问题、解决问题的能力。

本书的第一、五、七、八、九章由徐明编写;祁国鹰编写第二、三、四、六章,并负责全书的组织、内容的选择、软件解题方法的确认和统稿工作;本书各章节中用 Excel 进行统计分析的内容由周星宇负责编写。在编写过程中,参考了有关书籍和教材,引用了其中的一些材料和统计图表,特在此谨向各书的作者和出版者表示深切地感谢。鉴于我们水平有限,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

2003 年 5 月

修 订 说 明

《体育统计简明教程》作为体育教育和运动训练专业本科常规体育统计方法课程的教材与成人体育教材以及研究生体育科研方法课程的辅助教材或教学参考书出版以来,受到了广大读者的深切关注和好评;尤其是对编写该书遵从“淡化系统,突出实用”的观点以及编写形式进行了肯定。

同行与读者在使用《体育统计简明教程》的时候指出了书中的一些错误之处,并且对编写内容也提出了积极性的建议。例如,在该书的内容中应当结合体育科研介绍一些有关体育评价方法和实验设计的知识等等。对同行与读者提出的批评和宝贵建议,在此一并表示衷心的感谢。

此次对《体育统计简明教程》再版,除了对原书做勘误之外,由祁国鹰负责新增补了两章内容,分别是“第十章:体育科研中常用的综合评价方法”和“第十一章:体育科研中常用的统计设计方法”。

祁国鹰

2006年5月18日于北京

再次修订说明

《体育统计简明教程》作为北京体育大学运动人体科学、体育教育、运动训练以及民族传统体育专业本科常规体育统计方法课程的教材与北京体育大学成人教育体育统计课程教材,自从 2004 年开始使用至今已有五年多了,其中在 2006 年应读者要求对书中的部分内容进行了勘误、增补和修订。

《体育统计简明教程》在使用中得到了授课教师和学生们的的好评;尤其是对该教材的“淡化系统,突出实用”编写理念,同行们给予了深切地关注和肯定。适用于体育专业本科教学的常规体育统计课程教材无论是从编写的理念到形式,还是从内容的选择到例题的设计都必须“与时俱进”。一本优秀的教材无不是经过了多次使用和反复修订逐步完成的。因此,对《体育统计简明教程》再做适时修订,逐步完善其内容及形式,使之成为一本优秀的适用于体育专业本科教学的体育统计教材是势在必行。

本次再版除了对《体育统计简明教程》(2006 年修订版)的原有内容进行勘误之外,还对该书中的附表使用 Microsoft Excel 软件中的内置函数进行了重新编制以便勘正印刷错误;再由原书的章节编写者对“第七章 假设检验”的部分内容进行了改写以及增补了“两个正态总体方差齐性的假设检验”和“方差不齐时,两个正态总体均数的假设检验”等相关内容。

祁国鹰

2009 年 5 月 18 日于北京

目 录

第一章 总 论

第一节	体育统计的概念	(2)
第二节	体育统计的作用	(2)
第三节	部分基本概念	(3)
第四节	体育统计的内容	(7)
第五节	常用统计资料的收集与整理	(8)
习题一		(12)

第二章 平均指标

第一节	算术平均数	(14)
第二节	中位数	(15)
第三节	众 数	(15)
第四节	百分位数	(16)
第五节	大样本资料的平均指标简捷计算	(16)
第六节	用 Excel 计算平均指标	(19)
习题二		(23)

第三章 标志变异指标

第一节	极 差	(25)
第二节	四分位差	(25)
第三节	平均差	(26)
第四节	方 差	(28)
第五节	标准差	(29)

第六节 变异系数	(31)
第七节 用 Excel 计算标志变异指标	(33)
习题三	(35)

第四章 正态分布

第一节 正态分布曲线的形式	(37)
第二节 正态分布曲线的特征	(38)
第三节 标准正态分布表	(39)
第四节 统计资料的正态性检验	(43)
第五节 可疑数据的舍、取	(45)
第六节 正态分布理论在体育中的两个应用	(46)
习题四	(49)

第五章 体育评分方法

第一节 标准分数	(51)
第二节 标准百分	(52)
第三节 累进评分法	(54)
习题五	(58)

第六章 总体参数估计

第一节 t 分布	(60)
第二节 t 值表	(60)
第三节 标准误	(62)
第四节 总体平均数的区间估计	(63)
第五节 总体平均数之差的区间估计	(66)
第六节 总体比例数的区间估计	(68)
第七节 用 Excel 做总体参数估计	(69)
习题六	(73)

第七章 假设检验

第一节	基本概念	(76)
第二节	一个正态总体均数的假设检验	(80)
第三节	两个正态总体均数的假设检验	(84)
第四节	关于比率之间差别的假设检验	(92)
第五节	两个总体方差齐性的假设检验	(95)
第六节	方差不齐时,两个正态总体均数的假设检验	(96)
第七节	假设检验注意事项	(98)
第八节	用 Excel 做假设检验	(100)
习题七		(114)

第八章 χ^2 检验

第一节	χ^2 分布与 χ^2 值表	(117)
第二节	四格表资料的 χ^2 检验	(119)
第三节	行 $\times$ 列表资料的 χ^2 检验	(121)
第四节	用 Excel 做 χ^2 检验	(123)
习题八		(126)

第九章 相关与回归

第一节	直线相关分析	(129)
第二节	一元线性回归分析	(136)
第三节	用 Excel 做线性相关与回归分析	(144)
习题九		(148)

第十章 体育科研中常用的综合评价方法

第一节	在评价指标体系的建立方面应充分注意的几个问题	(151)
第二节	一般综合评价方法	(152)

第三节 加权综合评分法	(154)
第四节 定性综合评分法	(158)

第十一章 体育科研中常用的统计设计方法

第一节 实验设计的基本要素和基本原则	(161)
第二节 常用统计设计方案	(163)

附 表:

1. 标准正态曲线下的面积表	(176)
2. t 值表	(178)
3. F 值表(方差齐性检验用)	(181)
4. F 值表(方差分析用)	(183)
5. χ^2 值表	(191)
6. 相关系数界值表	(193)

参考书目	(197)
------------	-------

第一章

总 论

内容摘要:

本章讲述体育统计的概念、作用,讲明体育统计分为统计设计、描述统计、统计推断三部分内容。对总体、样本、个体、误差、随机现象、随机事件的频率和概率等基本概念做出详述。另外讲述了收集、整理统计资料的部分常用方法和在具体实施工作的过程中应该加以注意的问题。

学习目标:

通过对本章的学习,应了解并掌握体育统计的若干基本概念,对概念正确理解才能在以后的学习过程中更好地掌握知识、掌握方法。要求学过本章之后,对基本概念和它们之间可能的联系要熟知;对常用统计资料的收集方法和整理方法要掌握;在实践工作中,要能结合体育实践加以应用。在统计工作中,对每一次统计调查都要有明确的研究目的,并要养成实事求是的科学态度。

自学指导:

对本章所给的基本概念在理解的基础上正确记忆,多参照体育实践去举例,达到完成记忆及搞清各种概念之间可能存在的联系的目的。可以进行小样本资料的收集和整理工作,一方面学习知识和方法;另一方面提高动手能力。

第一节 体育统计的概念

体育统计是运用数理统计的理论和方法,来研究体育教学、训练、科研和管理中的定量或定性的现象,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。

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订计划提供依据。

数理统计的研究对象主要是不确定的随机现象,在体育中的大量现象都是这种不确定的随机现象,所以体育统计应运而生。体育运动中的许多问题往往要依靠调查或实验才能解决,而调查实验人员总是有限的,因此要想从局部得来的资料去推论整体的性质,得出科学的结论来,必须应用科学的统计方法。

第二节 体育统计的作用

体育统计是一门工具学科,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的体育发展情况,并进行统计分析和预测,为制订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)为进行宣传教育和从事科学研究提供资料。

统计工作的任务,主要落实在提供资料上。这些资料,既包括对体育发展情况的描述,也包括对其规律性的揭示;既包括反映现实情况和历史情况的统计资料,也包括对现象的发展前景作出必要的预测;既包括“揭露矛盾”,也包括提出建议。资料提供得愈多、愈好,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。通过学习体育统计,有助于帮助我们分析资料,处理数据,科学地制订实验计划,有助于提高体育教学、训练和科研水平。同时,学习体育统计有助于我们学习国内外有关体育方面的先进经验,并有助于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。

体育统计方法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,多作练习,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,而不能创造出规律来。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。

第三节 部分基本概念

一、总体、个体、样本

根据研究目的而确定的研究对象的全体称为总体。总体中的每一个被研究的对象称为个体。总体所包含的个体数目用符号 N 表示,它可以是有限的,也可以是无限的。

在总体中用随机抽样方法抽出有限个个体的集合称为样本。所以样本是总体的“有代表性”的一部分。在总体较大,不可能了解每一个个体时,我们常常用随机抽样的方法对其样本进行观察研究,用以推断其总体的性质。样本所含的个体数目称为样本含量,用符号 n 表示。在体育统计中,我们常把样本含量超过 50 的样本称为大样本;把样本含量低于 50 的样本称为小样本。

例 1-1 要调查某年某市 12 岁健康男孩的身高水平,由于人力物力所限,不可能对该市每个 12 岁健康男孩都测量到,于是随机抽测了该市 120 名 12 岁健康男孩的身高。这里,该市全部 12 岁健康男孩的身高是总体,所测的这 120 名 12 岁健康男孩的身高是一个样本,样本含量 $n = 120$ 。

二、误 差

误差泛指测得值与真值之差,以及样本指标与总体指标之差,样本指标与样本指

标之差。主要有以下几种。

1. 系统误差

在收集资料时,测量系统(各种测量工具)不准或测试人员掌握的标准习惯性地偏高或偏低而造成的误差。系统误差经常是一个常量。在做调查工作之前,认真检查调校测量工具并对人员按同一标准培训,可避免系统误差发生。

2. 过失误差

由于测试人员口误、笔误等粗心大意造成的误差。一定要注意消除过失误差。无法纠正时,要剔除错误数据。

3. 随机测量误差

由于各种偶然因素而造成的不可避免的一种误差。如对同一对象多次测定的结果不完全一致。这时应尽量保持仪器性能和操作方法的稳定并将误差控制在一定的、允许的范围之内,在必要时可做统计处理。

4. 抽样误差

由于随机抽样而造成的样本与总体之间的差别,及样本与样本之间的差别。抽样误差是由个体间的差异造成的,是不可避免的。抽样误差有一定的规律。一般地,样本越大,抽样误差越小,因为这时样本越和总体接近,样本对总体的代表性越好。反之,样本越小,抽样误差越大,样本对总体的代表性越差。这里可明显看出保持一定的样本含量 n ,即注意保持样本的代表性的重要意义。

三、随机事件

在一定的条件下可能发生也可能不发生的事件,称为随机事件。

在一定的条件下必然发生的事件,称为必然事件,记作 U ;在一定的条件下不可能发生的事件,称为不可能事件,记作 V 。它们是随机事件的两个特殊情形。

体育运动中有大量的随机事件。随机事件一般用大写的英文字母 A 、 B 、 C 等来表示。

例 1-2 一名篮球运动员在 3 米线处投篮,投出的球“投进篮筐”是一个随机事件。

例 1-3 一名射手打靶,“命中 10 环”“命中 9 环”……“脱靶”都是随机事件。

例 1-4 在海平面,一个大气压下,把水加热到 100°C ,水必然沸腾。在上述条件下,“水沸腾”是一个必然事件,“水不沸腾”是不可能事件。

四、频率和概率

对于随机事件,在一次试验中是否发生,我们虽然不能预先知道,但随机事件在一次试验中发生的可能性是有大小之分的。例如,一个优秀射手在距靶 20 米处手枪立射一发子弹“命中 10 环”;一个初学者在同样条件下射击“命中 10 环”;这两个随机事件发生的可能性是明显不同的。然而,对事件发生的可能性只停留在基本上是定性的了解与描述上,实在太不够了。我们希望对它给出客观的定量的描述。

【频率的定义】在不变的一组条件 S 下,若随机事件 A 在 n 次试验中,共发生了 m 次,则比值 $\frac{m}{n}$ 称为 A 发生的频率。即

$$A \text{ 发生的频率} = \frac{\text{频数}}{\text{试验次数}} = \frac{m}{n}$$

经验表明,在 n 足够大时,频率常呈现稳定状态,即向某一数值趋近,该数值可称为频率稳定的中心。

【概率的统计定义】在不变的一组条件 S 下,重复作 n 次试验,当试验的次数 n 很大时,如果频率 $\frac{m}{n}$ 稳定地在某一数值 p 的附近摆动;而且一般说来随着试验次数的增多,这种摆动的幅度愈来愈小,则称数值 p 为随机事件 A 在条件组 S 下发生的概率,记作

$$P(A) = p$$

上述定义把概率概念规定为频率稳定的中心。明显地,数值 p 就成为 A 在 S 下发生的可能性大小的数量刻划,所以人们又称概率是可能性的度量。

前述定义可简单地说是:“频率具有稳定性的事件叫做随机事件,频率的稳定值叫做该随机事件的概率”。

由于频率 $\frac{m}{n}$ 总介于 0,1 之间,因而由概率的定义知,对任何随机事件 A ,有

$$0 \leq P(A) \leq 1$$

而对必然事件 U 及不可能事件 V ,显然有

$$P(U) = 1, \quad P(V) = 0$$

最后要指出,随机事件的频率和概率有区别也有联系,区别在于频率明显与试验次数有关,是一个随机变量;而概率与试验次数无关,是理论上客观存在的常量。至于联系,由概率的统计定义即知。

例 1-5 用一枚匀称的分币,用一定的动作向上抛,让分币自由落在桌面上,“正面向上”是一个随机事件。历史上,有人成千上万次地进行过试验,表 1-1 列出几个试验记录:

表 1-1 抛币试验记录

实验者	投掷次数 n	“正面向上”次数 m	频率 = $\frac{m}{n}$
迪摩甘	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

可看到,投掷次数越多,频率越接近 0.5。

0.5 是本例随机事件频率稳定的中心,按概率的统计定义,可知“正面向上”这个随机事件 A 的概率为 $P(A) = 0.5$ 。

【中心极限定理】(a) 无论数据的原始分布是何状态,只要当样本含量 n 无限增大时,样本均数的总体越来越接近于正态分布。

$$(b) \mu_{\bar{x}} = \mu$$

$$(c) \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

五、统计资料的类型

统计资料按性质可分为计量资料和计数资料两类。

计量资料是对每个个体进行测量所得的数据,一般用度量衡等单位表示。如身高(cm)、体重(kg)、血压(mmHg)、100 米跑成绩(秒)等。这些数据的性质是连续的可不断细分的,所以又称为连续型数据。

计数资料是先把个体按性质类别分组,同质同类的进入同一组,然后数一下各组中有多少个数据。如做某项运动成绩的测验,成绩可分为优、良、中、差四组,测验后清点每组人数。这种数据是整数,是间断的不可再细分的,所以又称为间断型数据。

六、统计量与参数

统计工作中,一般把由样本计算得到的指标称为样本统计量,简称“统计量”;而

把由总体计算得到的指标称为总体参数,简称“参数”。总体参数是客观存在的数值,因而是常量。样本统计量是随抽样而变的,是随机变量。在大量情况下,总体参数是我们难以取得的数,所以由样本统计量去估计总体参数是统计工作的重要任务之一。

第四节 体育统计的内容

统计工作是对社会进行调查研究以认识其本质和规律性的一种工作,这种调查研究的过程是我们对客观事物的一种认识过程。统计认识活动是一个不断深化的无止境的长过程,随着客观事物的不断变化,统计认识活动也要不断地进行。但是就一次体育统计活动来讲,一个完整的过程可为统计设计、描述统计、推断统计三部分内容。

体育统计是一项通过对事物的量的研究来认识其质的工作,因此,贯穿于整个统计认识过程的中心问题,就是如何正确对待量和质的关系问题。没有数量就没有质量;没有质量也就没有数量。例如,某单位篮球运动员人数众多而运动水平不高,则明显使用价值不高;反之,另一单位的篮球运动员水平虽高,但人数太少,也是满足不了需要的。

在统计工作中应怎样贯彻这个质与量的辩证关系呢?统计是研究量的,但却不是从定量开始的,而是从定性开始的,即在搜集统计资料(统计调查)之前,在统计设计阶段,就要确定调查对象的范围,规定分析这些对象的统计指标、分组方法、研究步骤、操纵控制的因素,对提出的假设、观察数据如何处理等等进行妥善的安排和设计。这种工作是下面定量工作的必要准备。统计设计是进行统计工作的依据,是处理实验或调查研究结果的先决条件。统计设计的主要任务是提供一套科学的方法,用较少的人力物力,通过较少次数的实验或调查研究而得出比较全面准确的结论。

描述统计主要研究如何整理实验或调查所得的大量数据,对它进行加工整理,列成分组统计表。并找出这些数据的分布特征,计算得到若干具有代表性的统计量。我们借助于这些统计量,可从杂乱无章的数据中得到有意义的信息,便于我们对不同的总体进行分析比较。描述统计是推断统计的基础,应重视它,注意方法的正确使用和统计量的准确得出。

推断统计是在描述统计的基础上,利用各种统计方法和所算出的统计量指标,对所掌握的资料加以分析和评价,认识事物的本质和规律性,通过局部(样本)去推断