

SPORTS STATISTICS



体育统计

丛湖平 主 编

余群英 副主编

《体育与科学》杂志社出版

287334

前 言

“一切皆流”这是古希腊辩证法奠基人赫拉克利特的一句名言。如果说这句话在遥远的古代是一种真知灼见，那么它就更适用于当今体育的发展，与过去历史上的任何时期相比，当代体育正以空前的速度迅猛发展。也正是在这种快速发展的影响下，人们的认识水平和思维方式随之发生了深刻的变化。如所周知，当今的体育科学研究，不仅在内容和形式上有了很大拓展，而且在研究方法方面也有了突破性的改观，人们已不再拘于传统的定性研究，而将视野转向如何把客观事物的各种表面或内在的本质属性以量化形式予以分析和研究的问题上。《体育统计》是运用数理统计的原理、方法、研究体育领域里各种随机现象规律性的学科，也是高校体育院系的一门必修课。迄今，它正在体育领域的各个研究方面充当着“工具”的角色。

本书根据教学大纲以及编写者多年来的教学经验，同时考虑到学生的具体情况而编写的。其指导思想是：贯彻高师体育专业的培养目标，注重本学科的实用性，以及培养学生分析、解决体育中的实际问题的能力。在内容安排上，着重介绍在体育研究中常用的统计思想和方法，以及统计研究的实例，进一步突出应用性。

本书编写的组织工作在杭州大学体育系领导的大力支持和帮助下，由丛湖平同志牵头组织编写。参加本书编写工作的有：上海师范大学体育系孙利民（第一章的第一、二节、

第二章)、孙力(第一章的第三、四节、第十一章的第三节),杭州师范学院体育系朱建国(第三章、第六章)、鱼觉民(第四章),宁波师范学院体育系王承瑜(第五章、第十一章的第一、二节),湖北大学体育系余群英(第七章、第八章、第九章),杭州大学体育系丛湖平(第十章、第十一章的第四、五节、第十二章、第十三章、第十四章、附表整理),丛湖平同志对全书进行了统稿。

本书适用于高师体育专业本、专科学生学习和大、中学体育教师继续教育用书,也可作为体育专业工作人员和科研人员的参考用书。本书目录里注有“*”符号的内容为本科学生或继续教育使用材料。

由于编写时间较紧,且限于我们的水平,书中错误在所难免,敬请读者纠正。

编者

1992年4月于杭州

目 录

前言

第一章 绪论	(1)
第一节 体育统计及其研究对象.....	(1)
第二节 体育统计与体育科研.....	(3)
第三节 体育统计的发展简况.....	(6)
第四节 几个常用的基本概念.....	(7)
第二章 统计资料的搜集与整理	(12)
第一节 资料的搜集.....	(12)
第二节 资料的整理.....	(15)
第三章 统计表与统计图	(20)
第一节 统计表.....	(20)
第二节 统计图.....	(23)
第四章 样本特征数	(31)
第一节 集中位置量数.....	(31)
第二节 算术平均数的两种求法.....	(35)
第三节 离中位置量数.....	(38)
第四节 标准差的两种求法.....	(41)
第五节 平均数和标准差在体育中的应用.....	(44)
第五章 动态分析	(47)
第一节 几个基本概念.....	(47)
第二节 动态分析及其图表.....	(54)
第三节 动态分析在体育中的应用.....	(56)

第六章	正态分布	(62)
	正态分布的概念和性质.....	(62)
	正态分布表.....	(64)
	正态分布理论在体育中的应用.....	(67)
第七章	统计抽样与估计	(82)
	第一节 统计抽样的概念及其特点.....	(82)
	第二节 几种常用的抽样方法.....	(83)
	第三节 抽样误差及标准误.....	(86)
	• 第四节 参数估计.....	(88)
第八章	假设检验	(95)
	第一节 假设检验的基本概念、原理与方法.....	(95)
	第二节 均数的假设检验.....	(99)
	• 第三节 两个正态总体方差的检验.....	(110)
	第四节 总体率的假设检验.....	(111)
	• 第五节 分布的检验.....	(114)
	第六节 假设检验中的两类错误.....	(117)
	第七节 假设检验方法在体育中的应用.....	(119)
• 第九章	方差分析	(125)
	第一节 方差分析的基本概念.....	(125)
	第二节 单因素方差分析.....	(128)
	第三节 多重比较法.....	(135)
	第四节 双因素方差分析.....	(141)
	第五节 方差分析在体育中的应用.....	(158)
• 第十章	综合评价	(164)
	第一节 综合评价的概念及功能.....	(164)
	第二节 综合评价方法.....	(167)

第三节	综合评价的研究设计及要注意的几个问题·····	(183)
第四节	综合评价方法在体育中的应用·····	(187)
第十一章	相关分析·····	(194)
第一节	相关分析的概念及性质·····	(194)
第二节	相关系数的计算与检验·····	(197)
第三节	等级相关·····	(202)
• 第四节	偏相关与复相关·····	(204)
第五节	相关分析在体育中的应用·····	(210)
第十二章	一元回归线性分析·····	(214)
第一节	回归分析及其功能·····	(214)
第二节	一元线性回归方程·····	(216)
第三节	一元线性回归分析在体育中的应用·····	(227)
• 第十三章	多元线性回归分析·····	(230)
第一节	多元线性回归方程·····	(230)
第二节	多元线性回归分析在体育中的应用·····	(244)
• 第十四章	研究设计·····	(252)
第一节	研究设计的基本内容和结构·····	(252)
第二节	调查设计·····	(255)
第三节	实验设计·····	(263)
附表	·····	(275)
参考文献	·····	(296)

第一章 绪 论

随着体育科学的发展，体育统计这门新兴学科正以前所未有的速度蓬勃发展。无论是统计知识的普及，还是统计研究的深度和广度，均达到了一定的水平，这也是发展之必然。

统计学是一门研究社会现象和自然现象规律的一门学科。当我们在应用统计学的原理、方法来研究体育领域里的各类问题时，就会面临一个如何将统计思想和方法与实际研究内容相联系的问题。由于体育现象的多重性和复杂性，以及多学科的综合性，在很大程度上决定了体育统计方法有其自身的特点。

本书主要介绍一些常用的、适用于体育科学研究的统计原理和方法。

第一节 体育统计及其研究对象

一、什么是体育统计

统计，是一种特定方法的计算。人们进行统计工作的目的是研究大量事物、现象数量方面（包括数量多少、现象之间的数量关系、数量的分布特征以及质与量互变的数量界限等）的某些规律。统计可分为两类：一类描述性统计；另一类是推断性统计。前者主要是对事物的某些特征的状况进行实际的数量描述，后者则是通过样本的数量特征以一定方式估

计、推断总体的某些状况。体育统计从其学科体系来看，属推断性统计。

由于体育统计的发展时间较短，有关体育统计的定义尚未形成共识。有人认为体育统计是以适应体育事业的发展，训练手段科学化、现代化而发展起来的学科，且由于体育研究中涉及到很多随机现象，许多问题往往需要进行调查和实验。所以，体育统计的任务是揭示体育科学中大量随机现象的规律性。另有一些人认为体育统计是在辩证唯物论的指导下，运用数理统计和其他数学方法来研究体育教学、训练、科研和管理中的各种问题的学科。目前，对体育统计较一致的观点是：体育统计是应用数理统计的原理和方法对体育领域里各种随机现象规律性进行研究的一门基础应用学科。

二、体育统计的研究对象

正因为体育研究的内容在不断扩大，体育统计的研究对象也没有明晰的界面。迄今为止，有关体育统计的研究对象也有不同的论述。一种观点认为：体育统计的研究对象应是体育领域里的所有随机现象。只要能量化且又能进行大量观察的事物，都可作为体育统计的研究对象；另一种观点认为：体育统计的研究对象除了体育领域里的随机现象外，还应包括非体育领域但对体育有一定影响的随机现象。

尽管体育中的各种事物千差万别，但无不具有质和量两个方面的特征，换言之，体育中任何事物的规定性均表现为质和量的规定性。通过对事物进行量的规定性的研究，能更准确地认识事物发展的规律。体育统计就是从量的角度认识体育的各种随机现象的发生规律。由于体育系统并不是孤立存在着的，它是与其他系统，诸如社会系统、经济系统、自然

环境等系统并存，且互为联系，互相影响，互相制约的。因此，我们认为，体育统计的研究对象不能仅仅限于体育系统内部的随机事物，而应将外延扩大至凡与体育有联系和影响的其他系统的随机事物。例如，当要考察国家或地区的经济发展水平，人们的文化素质对竞技体育发展的影响时，就不得不要我们采用体育统计的手段、方法研究经济与文化素质等方面的内容。可以设想，随着体育“软科学”的进一步发展，所涉及的研究领域将会有更大的延伸，相应的研究对象的范围也会越来越大。

基于此，我们认为，体育统计的研究对象除了体育领域的随机变量外，还包括非体育领域但与体育有着一定联系的其他系统的随机变量。

第二节 体育统计与体育科研

体育统计作为一门“工具”学科，在当今的体育科学研究中有着广泛的应用。

一、体育科研的基本过程

尽管体育的研究类型很多，涉及的学科也不尽相同，但整个研究工作基本上具有相同的过程。

体育科研的基本研究过程为：

- (一) 选择课题或建立假设
- (二) 制订研究设计
- (三) 搜集资料
- (四) 整理和分析资料
- (五) 解释结果或检验假设

这些步骤中的每一环都与其他各环相依赖。明显的是，人们在搜集资料之前不可能分析资料。当然，也许不太明显的是，若一个研究人员对其所研究的对象茫无所知，也不知道如何分析资料，那就不可能制定假设，无力制订研究设计，因而无法进行研究工作。在上述5个步骤中，研究者必须具有研究方法方面的知识，这是至关重要的。

二、体育统计在体育研究中的作用

(一) 帮助研究者有效地搜集资料

体育科学工作者需随时获取新的知识，在进行研究工作时尤其如此。可以想象，如果一个研究人员缺乏必要的统计知识，就难以理解统计学的专门术语以及各种符号所表示的意义。这样，也就无法理解和引用以统计方法为主要研究手段的一些研究成果，自然也就难以获取部分新知识。

当今社会乃信息社会，各种论文、专著浩瀚如海，在任何一项研究中，研究者都可寻到大量的参考资料，由于这些研究资料可能是以不同的实验方案或调查方法、不同的仪器、不同的对象予以研究的，所以在研究结果方面可能不尽相同。这样我们必然会面对选择什么资料，如何选择资料等问题，那么，怎样鉴别研究结论的信度和效度、参考价值的大小。除了相应专业的知识判定外，体育统计原理和方法是判别真伪的有效工具。

首先，体育统计是建立在严密推理和数学分析的基础上的，它是发现事物内在规律的基本方法，各种调查或实验所获的数据是否经过统计学处理，已成为其结论可靠程度大小的重要标志之一。

其二，体育统计方法对实验或调查等研究方式有一定要

求。如果存在抽样方法、对照方法等不符要求，或样本含量过少的情况时，就可判定其结论的可靠性程度不高，参考价值有限。

其三，任何一种统计方法均基于相应的概率分布，也就是说，只有在基本认定总体数据的分布符合某种特征的前提下，样本资料才能采用某种统计方法予以处理，如若不然，盲目套用公式，则很可能导致错误结论。我们也可从这个角度来判定参考资料的价值高低。

总之，较好地掌握统计的原理和方法，能帮助我们有效地搜集资料。

（二）帮助研究者制定研究设计

发现问题（或提出问题）和解决问题是研究工作中最为重要的两个方面。发现问题的能力往往体现研究者掌握基础知识的面和深度上，而解决问题的能力往往表现在研究者掌握各种研究手段、方法以及应用方面的能力上。在制定研究设计时，若已对所要研究的问题进行了界定，此时整个研究设计的重点则是整体研究的框架和相应的研究方法的设计问题，这里包括以何种方式搜集资料，有哪些要求必须满足，变量的性质，变量的分布、以及资料的处理和资料处理结果的解释等。这一切都与研究者掌握定量研究方法的程度有密切关系。

体育统计作为定量研究的“工具”，不仅为研究者提供了各种类型、解决各种实际问题的研究方法，而且还为研究者提供了非常有效的研究思想，诸如，以样本推断总体、统计检验等，且能使研究者最大限度地减少人、财、物的投入量，使研究工作经济有效。

（三）帮助研究者有效地发现事物的规律

体育统计是基于概率论数理统计的，具有严谨的科学性，它是认识世界、寻找事物客观规律的有效工具。

例如，判定教学和训练计划的合理性，探索新的教学方法和教学模式、新的训练方法和训练模式；运动员的选材；研究青少年发展时期各方面的敏感期；从众多因素中选择主要研究因素等课题都可采用统计方法进行研究，寻找规律。

第三节 体育统计的发展简况

在原始社会末期，已有了统计的萌芽，当时仅限于人口、土地、等方面的统计。从现代统计分类的角度来看，上述统计属于社会统计学范畴。概率论作为一门学科，最早出现在十七世纪中期的欧洲，由于当时的概率论仅研究等可能的事件的规律，故很长的时间未能较好的发展。自1925年左右。英国的农业技师、统计学家费希尔（Fisher）的《研究工作应用数理统计方法》一书问世后，数理统计的原理和方法才得以广泛的应用，分别形成了很多统计学的分支，诸如教育统计学、生物统计学、卫生统计学和体育统计学等。我国最早在本世纪30年代后期有几所学校开设了体育统计学，由于当时国内处在混乱状态中，整个教育体系都不稳定，因此，体育统计学没有得到很好的发展。

我国体育统计学在70年代末期才得以真正发展。80年代初，国家教委明确规定将体育统计学列为高校体育院系的必修课，这标志着体育统计学科的发展进入了一个新的阶段。在此以后，为了适应教学工作的需要，全国举办了多次体育统计学培训班，同时又有几十本体育统计学教材、讲义相继

编写、问世（有的正式出版、有的是油印本）。1984年经中国体育科学学会批准，成立了全国体育统计专业委员会，真正打下了体育统计学进一步发展的基础。迄今为止，全国体育统计专业委员会已举办了八次全国学术性学术会议，有效地促进了体育统计学科研究的发展。可以预计，体育统计学科不仅在学科体系，还是学科本身的研究上将会有更大的发展。

第四节 几个常用的基本概念

为了以后各章节的需要，本节介绍几个最常用的基本概念，读者要很好地掌握它们，以便今后的学习。

一、总体、样本和个体

根据研究目的而确定的同质对象的全体为总体。组成总体的每个基本单位为个体。根据需要从总体中抽取的部分个体所形成的子集为样本。样本中所包含的个体数量为样本含量，通常用字母 n 表示。

总体可分为假想总体和现存总体两种形式。本节仅介绍现存总体。现存总体又可分为有限总体和无限总体，如上海市初三年级男生是一个有限总体，它的边界是有限、明晰的，类似这样的总体为有限总体，若要研究全人类的体质，由于每天都有死亡和新出生的人，这样的总体可视为是无限总体。

样本又可分为随机样本和非随机样本两种形式。所谓随机样本是指采用随机取样方法获取的样本，样本中的每个个体在抽样前所抽取的可能性是均等的。如若要在上海师大体育系里随机抽取20名男生作为研究样本，那么可通过抽签法或其他方法获取该样本，这个样本是随机样本。所谓非随机

- 、样本是指研究者根据研究的需要，寻找具备一定条件的对象为研究样本。如要进行配对实验研究，就要求在实验前保证实验组和对照组的对象水平相同（或基本相同），这时所获取的样本不是随机的，而是按照一定要求选择的。

在任何抽样研究中，对样本含量是有一定的要求的。一般认为：

$n \geq 45$ 为大样本

$n < 45$ 为小样本

若一项研究的样本含量小于30，一般情况下要对样本含量的推断效果进行估计，具体方法见第十四章。

二、随机事件

客观世界中有许多现象，在一定条件下人们是可以事先预知它的出现与否。例如，重物抛起后必然要落到地面上；在标准大气压下，水温达 100° 时会沸腾，这是一类必然事件。除此之外，还有大量的事件事先是无法明确确定的，例如在投篮之前，我们无法明确确定是投中还是投不中；再如抛掷一枚均匀的硬币，结果可能是正面，也可能是反面，在抛掷前是无法确定哪一面定会发生。对这类事件我们称它为随机事件。随机事件的定义可以这样表示：在一定的实验条件下，有可能发生也有可能不发生的事件为随机事件。

三、随机变量

随机事件中，有很大一部分与数值有关，我们把随机事件因结果不同而变化的数值称为随机变量，换言之，随机事件所对应的量，就是随机变量。例如，我们测定某小学六年级学生的身高，那么，每个具体的学生的身高数值是不相同的，即135、137、140等，这些数就可认为是随机变量；又

如，打靶的“打中”与“打不中”，“打”、“打不中”是随机事件，若以1表示“打中”，以0表示“打不中”，这样，就将随机事件转换成了随机变量，上述1和0就是随机变量。

随机变量有两种形式：一是连续型变量，二是离散型变量。前者在一定的范围里，变量个数是无限不可列的，如上述的身高变量，在130cm—140cm范围里，变量可以是无限多个，且不可列出的（在具体测试中此类变量只是取了整数或取一个近似值），后者则是变量范围明确，变量是有限可列的或无限可列的，如打靶，变量范围只含1和0，且是有限可列的。

在统计分析中，变量的形式是非常重要的，它决定了我们采用不同统计方法来研究问题。这是因为不同形式的变量，有其不同的分布特征，这就要求我们采用相应的统计方法。这一点读者务必要注意。

四、随机变量的概率及性质

随机变量的规律，主要体现在它的概率及其分布两个方面。本节主要介绍概率及其性质。

（一）概率的古典定义

设在实验中全部等可能的独立的基本结果有 n 个，其中有 m 个属于事件 A ，则在实验中，称事件 A 出现的概率 P 等于 m 与 n 之比，即

$$P(A) = \frac{m}{n} \quad (1.1)$$

例1.1，某校学生共300人，其中患近视眼的有58人，若随机抽取一个学生，抽取患近视眼的概率是多少？

此时, $n = 300$, $m = 58$

$$\text{则 } P(\text{近视眼}) = \frac{m}{n} = \frac{58}{300} = 0.193$$

可知在该校进行一次抽样, 抽到患视眼的对象的概率为 0.193。

(二) 概率的统计定义

在统计的实际研究中, 由于总体状况往往不清, 古典概率定义不再适用。因而产生了概率的统计定义:

设在一定条件下, 重复进行某实验, 如果该实验重复进行 n 次, 事件 A 出现 m 次, 则称 m/n 为事件 A 在 n 次实验中的频率 $f(A)$, 记成 $f(A) = m/n$ 。当 n 很大时, 频率 $f(A)$ 逐渐稳定在一个常数 P 附近摆动, 则称事件 A 有概率, 且定义为:

$$P(A) = \frac{m}{n} \quad (1.2)$$

这就是概率的统计定义。

如球类比赛中抛挑边器, 挑边器只有两个面, 若一边为蓝面, 另一边为白面。抛一次挑边器只有两个可能结果, 当我们连续抛掷 n 次, 且 n 很大时, 出现蓝面或白面的次数非常接近, 若设蓝面在 n 次实验中出现 m 次, 则出现蓝面的概率为:

$$P(\text{蓝}) = \frac{m}{n} = 0.5 \quad (1.3)$$

同理, 白面出现的概率为:

$$P(\text{白}) = 0.5 \quad (1.4)$$

上述结果很容易在实验中得以验证。

(三) 概率的主要性质

概率的主要性质有:

性质 I: 概率 P 为非负值, 因 $m \geq 0$, 故任何随机事件的概率 $P \geq 0$

性质 II: 当 $m = n$ 时, $P(A) = 1$, 事件 A 为必然事件。

当 $m = 0$ 时, $P(A) = 0$, 事件 A 为不可能发生事件。

性质 III: 当 $m < n$ 时, $P(A)$ 值小于 1, 且各事件 A 、 B 、 C ... 的概率和等于 1。即

$$P(A) + P(B) + P(C) + \dots = 1 \quad (1.5)$$

性质 IV: 概率具有可加性, 若 A 、 B 两事件互相独立, 则有 $P(A) + P(B) = P(A + B)$ (1.6)

习 题

- 1、试述体育统计的研究对象?
- 2、用实例说明总体, 样本、个体的涵义。
- 3、随机事件、随机变量是什么? 它们之间有什么关系? 为什么在一些情况中要将随机事件转换成随机变量?
- 4、什么是概率, 并用例子说明其涵义。