

高等师范院校体育系科

体育统计

(试用教材)

高等师范院校体育系科
《体育统计》教材编写组



言

这本《体育统计》试用教材，系根据教育部先后在北京、安徽召开的全国高师体育系教材会议的精神和高等师范院校体育专业教学计划而编写的。1982年10月，教育部委托华中师范学院牵头，召开了高师体育专业本科《体育统计》教学大纲讨论会，继后，进行了教材编写的组织工作，曾经过教材编写细目讨论会、初稿讨论会、二稿讨论会，最后，整理成此书，作为试用。

本书力求贯彻高师体育专业培养目标，教学计划及大纲要求，并联系学校体育教学和科研实际，注意统计方法和理论的结合，以期达到培养和提高学生分析问题和解决问题的能力，在叙述上也力求简明。同时，为了照顾到目前体育专业学生数学水平和将来的实际需要，以及使教材做到相对稳定，书中安排了部分选用内容（以星号※标记）。供教学参考。本课程讲授42学时。

全书共分十章，参加编写的院校共八所，承担执笔任务的有：第一章第一节，昆明师大刘先祚；第二节，华中师范学院刘厚生；第二章，江西师院贾六本；第三章，华东师大孙昌荣；第四章，贾六本；第五章，苏州大学王文英；第六章，刘厚生；第七章，北京师大李哲；第八章及第九章，河南师大张庆林；第十章第一节，刘先祚；第二节青海师大何争流。在编写过程中，得到有关兄弟院校及部门、同志们热情支持和帮助，并由湖北省嘉鱼县印刷厂承印，在此表示感谢。

“体育统计”是一门新兴的应用学科，在高师体育系本科中，作为专业基础理论课开设的时间不长。限于我们的水平和实践，难免会有缺点和错误，诚恳地希望试用单位和同志们提出意见。以便作进一步的修改。

高师体育系《体育统计》

教材编写组

一九八四年五月

目 录

前 言

第一章 绪论	(1)
第一节 体育统计的研究对象、任务和方法	(1)
第二节 几个基本概念	(3)
第二章 资料的搜集与整理	(11)
第一节 资料的搜集	(11)
第二节 频数分布表	(14)
第三节 频数分布直方图	(18)
* 第四节 频数分布多边图的修匀	(19)
第三章 几个常用统计量	(23)
第一节 集中位置量数	(23)
第二节 离中量数	(35)
第三节 变异系数	(45)
第四章 相对数	(51)
第一节 率、构成比和相对比	(51)
第二节 动态数列	(58)
* 第三节 相对数的标准化	(69)
第五章 正态分布	(77)
第一节 正态分布的概念和性质	(77)
第二节 正态分布表	(79)
第三节 正态分布的应用	(83)
第四节 正态性检验	(102)

第六章 假设检验	(111)
第一节 假设检验的基本思想	(111)
第二节 显著性水平	(118)
第三节 u 检验	(120)
第四节 t 检验	(127)
第五节 χ^2 检验	(136)
第六节 符号检验和秩和检验	(147)
*第七节 参数估计	(156)
第七章 方差分析	(165)
第一节 方差分析的基本思想	(165)
第二节 单因素方差分析	(169)
第三节 样本均数的多重比较	(175)
* 第四节 双因素方差分析	(182)
第八章 相关与回归	(192)
第一节 相关与回归的基本概念	(192)
第二节 相关系数的计算与检验	(195)
第三节 等级相关	(203)
第四节 一元线性回归	(207)
第五节 两条回归直线的比较	(217)
* 第六节 多元线性回归	(220)
第九章 统计表与统计图	(243)
第一节 统计表	(243)
第二节 统计图	(248)
第十章 研究设计	(265)
第一节 调查设计	(266)
一、调查设计的基本内容	(266)
二、调查设计的基本方法	(269)

三、抽样调查设计中样本含量的估计.....	(275)
第二节 实验设计.....	(283)
一、实验设计的基本内容.....	(283)
二、实验设计中样本含量的估计.....	(287)
三、几种常用的实验方法.....	(292)
四、实验方案的制订.....	(298)

附 表

1. 正态曲线纵坐标表.....	(1—3)
2. 正态分布 (甲) 由 u 查 $\frac{\theta}{2}$ 值表.....	(4—6)
3. 正态分布 (乙) 由 $\frac{\theta}{2}$ 查 u 值表.....	(7—9)
4. 正态分布的双侧临界值 (u_{α}) 表.....	(10)
5. 正态性 D 检验界值表.....	(11—12)
6. χ^2 分布上侧临界值 (χ^2_{α}) 表.....	(13—14)
7. t 检验临界值 (t_{α}) 表.....	(15—17)
8. 符号检验表.....	(18)
9. 秩和检验表.....	(19)
10. 相关系数界值表.....	(20—22)
11. F 检验临界值 (F_{α}) 表.....	(23—32)
12. 多重比较中的 q 表.....	(33—36)
13. 等级相关系数界值表.....	(37—38)
14. 随机数表 (一)、(二).....	(39—42)
15. 两样本率比较时所需样本含量表 (单侧)、 (双侧).....	(43—44)
16. 配对比较 (t 检验) 时所需样本含量表.....	(45—46)
17. 两样本均数 t 检验时所需样本含量表.....	(47—48)

18. T₁值表(双侧)..... (49)

19. 正交表..... (50)

20. 本书统计符号表..... (51—54)

(T₁)..... 且是前是并本并中并前是并

(S₁)..... 一点何何何何何何何何何何

(S₂)..... 何何何何何何何何何何

表 1-1

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

(1—1)..... (附录)

(1—1)..... 且是前是并本并中并前是并

(1—1)..... 且是前是并本并中并前是并

第一章 绪 论

第一节 体育统计的研究对象、

任务和方法

随着体育科学的飞跃进步，特别是计算数学和电子计算技术在体育领域中的引进，为发展体育统计提供了条件。体育统计已逐渐成为体育领域中的一门新兴的基础应用学科，日益受到国家和广大体育工作者的重视，近年来，在理论提高和方法普及方面，都取得很大成绩。体育统计已列为我国高等学校体育专业的专业基础理论课。

体育统计学的范畴、内容，应从体育统计学科的广义上理解。体育统计的理论，应包含指导思想，统计的数理理论，结合专业知识；统计的方法，包含调查、综合、对比、归纳、推断等；方法论和认识论不可分开，统计的分析结论，必需受到体育实践的检验，那种把统计只看作是统计计算来理解体育统计学是狭义的和片面的。

体育统计是以辩证唯物主义思想为指导，以体育为对象，运用统计的理论和方法，通过科学实验、调查，数据的搜集、整理、分析和推断，从中揭示体育发展规律的应用科学。统计的分析方法，在体育的科学化、现代化，在指导体育实践中，均具有重要作用。

统计离不开计算，但它又不是纯数字概念的数学。因为体育统计的数字是对体育现象本质的数量描述，离开体育的性质、任务，单纯考虑数据的统计和计算，那将是毫无意义的。

本书从教学要求出发，安排十章，内容包含资料收集、整理，基本统计量的计算、统计图表等属于描述部分的内容；及假设检验、方差分析、相关回归，属于推断部分的内容；最后

是研究设计,介绍体育统计的基本研究设计方法。打※号的部分供讲授选用。由于教学时数对篇幅的限制,聚类、主成份分析以及模糊数学的统计方法等未列入本教材。

从阐述上,既注重基本方法的介绍和计算技能的培养,同时又注意原理的理解和方法的正确运用。

一、体育统计的研究对象

体育统计服务于体育的总目的。在体育领域中,凡能以数量表示的现象,包括体育教学、训练、科学实验和管理,都是体育统计的研究内容。因此,体育统计的研究对象很广泛。在体育领域中能用数量表现的方面很多,如:体质方面有体格、体能、有机体适应能力和心理状态方面的指标,这些已在我国目前实施的“国家体育锻炼标准”、“体育标准”、“健康卡片”中均已有明确规定。

运动训练方面有力量、速度、耐力、柔韧、灵敏等运动素质和专项运动素质的指标。

体育教学有课的密度、负荷、成绩考核以及体育锻炼、竞赛、评定等指标。其他运动生理、生化方面的指标也都可以是体育统计的研究内容。

二、体育统计的任务

在我国,体育统计的任务,是为迅速发展我国体育事业、增强人民体质,提高运动水平、实现四化的总目标服务的,体育统计所承担的任务和体育统计的职能作用也是紧密相连的。

人体运动因素极为复杂,大多属于多因素影响的不确定性现象,纯定性的研究已经完全不适应形势的需要,从而必须借助数理统计的理论和方法,从大量观测的数据中,科学的推断和预测,以揭示客观事物的统计规律性,指导实践。

一般地说,体育统计的任务有以下四个方面:

(一)运用有效手段,通过抽样调查和研究,及时提供反映我国人民体质状况、健康水平和体育运动规模等有关指标,作为改善和提高体育管理水平,制定计划、政策的依据。

(二)在体育教学方面,通过经常性的资料积累及实验数

据的统计处理，及时获得青少年体质健康状况的多项指标，必要时还可建立追踪观测，为考察教学效果、改进教学方法、指导体育锻炼提出建议，提供依据。

（三）在体育训练中，可以通过统计观测，进行多因素分析，确定训练效果；对因素控制和预测，综合评定、评价等。

（四）体育统计的任务，当然也包括对学科本身的研究，如理论的提高和方法的完善、普及等方面。

三、正确运用体育统计方法

体育统计的方法论是和辩证法的认识论分不开的，统计的方法很多，有描述统计、推断统计及研究设计等。但任何一种统计方法的运用，必须是从体育的质和量、因素的相互约束、发展、变化等观点出发，去研究体育对象的数量特征。离开体育的本质，任何统计计算都将失去意义；其次，作为研究随机现象的数理统计方法，有着本身所要求的数理前提，离开这些前提，统计结论将失掉意义，如我们若利用典型调查的极端数据去替代随机样本，那么，任何推断结论都将是不真实的。所以，在应用统计方法时，必须强调应持严谨的科学态度。体育统计必须从体育的本质出发，提出课题，明确要求，严格统计程序、计算步骤，保证数据的有效性、真实性、可靠性。例如我们不应对两项实际无关的指标求相关回归，更不应带着结论去选择数据，装饰论文。而必须保证统计工作的严肃性，只有具有科学的态度，才能取得科学的统计结论。

体育统计方法的学习和运用，必须紧密联系实际，方法应用于实践，服务于实践。讲授本课时，还须注意培养学生分析问题和解决问题的能力。

第二节 几个基本概念

体育统计方法的应用，自始至终贯穿着数理统计的几个基本概念，如总体、样本、随机变量、概率与分布等，掌握这些基本概念，有助于方法的正确应用。

一、总体、样本

总体指统计研究对象的全体；样本指从总体中随机抽取作为代表进行直接观测的部分。例如我们要研究全国13岁儿童身高，那么，全国13岁儿童身高，便构成统计研究的指标总体 X ，其中每个13岁儿童身高数据 X_i 便称为统计总体中的个体，也是统计观测的基本单位。总体平均数、标准差等习惯称为总体参数。在实际的统计研究中，由于大范围的直接调查困难，有时甚至不可能进行，故统计研究大多采用抽样方法，即按随机原则在总体中抽取一定数量为 n 的个体（ n 称为样本含量）， n 个个体构成总体中的一个统计样本。对样本进行直接的统计调查（实验），然后用样本资料对总体情况进行推断和分析。直接计算得出的样本平均数，样本标准差等称为样本统计量。上例若随机地抽测一万名青少年，则一万名青少年的身高数据 X_i （ $i=1, 2, 3, \dots, 10,000$ ），便构成一个样本。

在实践中，有时作为研究对象的总体范围无法明确确定，可能是因总体很大，其中个体数数不尽，或是总体不很大，但个体数数不清，也可能根本无法确定范围。如一个篮球队或一个篮球队员的投篮命中率，每测一次得一个样本率，样本率是测出可知的，但总体率却说不清楚。又如推断未来时期一个总体指标，该指标的总体范围现在也无法具体明确，我们称这类总体为无限总体。须说明，对这类总体的范围，虽无法确定，但对它的参数特征值，仍可用样本作出估计，仍可用统计推断的方法进行总体比较。

由总体和样本的定义可知：

总体和样本的概念是相对的，如当选定的观测指标是从一个县范围推断一个省范围时，省范围为总体，县范围为样本；当选定的观测指标是从一个乡范围推断一个县范围时，这时县范围便成了总体。样本是对总体而言的，总体随研究课题而定。

此外，统计的任务是直接观测样本推断总体，故一般地说，总体参数未知，直接测知并计算的均系样本值。样本统计量反映了样本的数量特征，虽然它能推断总体情况，但它不等于总体的参数。样本质量如何，将直接影响统计结论，所谓样

本的质量，一方面指样本数据的真实性、可靠性，更重要的是由于统计推断是以随机理论为基础，故必须保证抽样的随机性，离开这个前提，样本便失去推断意义。

二、随机事件、随机变量

随机事件，指在某种条件下，可能发生也可能不发生的事件。如果，随机事件能用数量指标 X （ X 是一种符号）来表示，那么数量指标 X 便是一个随机变量，统计中的样本指标都具有这种性质都可看作随机变量。

例如，对某市某一年龄组男生身高进行抽样调查，抽取 $n=100$ 人的样本。属于该总体中的任何一个人都可能被抽，也有同等可能不被抽，抽测是一个随机事件。样本指标身高 X ，碰到个子高的 X 值较大，碰到个子矮的 X 值便较小。指标 X 随抽测对象不同而随机地变化，故 X 是一个随机变量。再如抽测一个篮球队员的投篮命中率 P ，由于随机因素的影响，一次成绩并不稳定， P 可在某一个水平上随机波动，样本命中率 P 也是一个随机变量，测10次，可能10次的成绩 P 均不一样。应注意的是：统计指标的随机性只有当作样本时，才体现出来，若对某市某年龄组男生身高进行全面调查，这时，测录的身高指标 X ，虽然也随测试对象不同而变化；但是，它不再是随机的，因属于该总体的每个青年男生，必定要受测，也就必定要在测录中出现，不存在什么可能出现或不可能出现的问题，也就不存在随机性，故总体指标、总体参数不是随机变量，它是客观存在的已确定的。

其次，样本指标的随机性主要是由抽样的随机性质决定的。样本指标 X 是一个随机变量，由样本数据演算得出的样本统计量，如样本平均数、标准差以及 $\bar{X} \pm K$ （ K 为常数），或两个样本数据相加减等，也是一个随机变量。

三、概率、概率分布

随机事件的规律，主要体现在它的概率及其分布两个方面，概率和概率分布在数理统计中都有严格的定义。

（一）概率

随机事件的概率含义系指它在某种条件下出现可能性（或称出现机会）的大小，概率大，即表示出现的可能性大。定义公式分为古典的和统计的两种形式。

1、概率的古典定义公式

设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，则在此实验中，称事件A出现的概率P等于M与N之比，即 $P(A) = \frac{M}{N} \dots\dots\dots (1-1)$

这里分别解释等可能、独立的、基本的三个概念。如球类比赛中抛挑边器，挑边器只有两个面，若一为兰面，一为白面。抛落一次，要么是兰面，要么是白面，没有另外的可能，既不可能出现非兰非白的，也不可能出现又兰又白的。因此，出现兰面的是一种基本结果，出现白面又是另一种基本结果，而且只有这两种基本结果。所谓独立的是指基本事件（出现兰面和出现白面）之间，均无因果或相联关系。不存在第一次出现兰面导致第二次出现白面，也不存在第一次出现白面导致第二次出现兰面，同色各次也无联系。各次之间及出现兰面和出现白面两不同基本事件间均是相互独立的。所谓等可能的是指在一次抛落中，两事件出现的可能性大小程度相等，显然，满足等可能条件，要求挑边器的制作，结构和质量均匀，对两面不存在倚重倚轻现象。

根据以上解释例示，便知针对抛挑边器这一随机现象，全部等可能的、独立的、基本结果N=2。其中出现兰面的基本结果M=1，同理出现白面的基本结果M=1。代入概率的古典定义式便有

出现兰面的概率

$$P(A = \text{兰}) = \frac{M}{N} = \frac{1}{2} = 0.5$$

同理出现白面的概率

$$P(A = \text{白}) = \frac{M}{N} = \frac{1}{2} = 0.5$$

又如在一个人数N=50的教学班中，体重X=30公斤的有

4人。若在该班随机地抽测一个学生的体重 X ，样本指标 X 便是一个随机变量，试求对该班的一次随机抽测中得出体重为30公斤的概率为多少？即求 $P(X=30)=?$

分析：在随机抽测前提下，测量指标体重 X 是一个随机变量，用 $X_i (i=1, 2, 3, \dots, 50)$ 代表该班50个人的体重值，则 X_i 中任何一个值都有同等可能被抽测出现， $X=X_i$ 中任何一个都是等可能事件。其次， $X=X_i$ 中任何一个值的事件，都具有相互独立性，都是独立事件。第三， $X=X_i$ 中任何一个值的事件又是基本事件，不可能把任何一个事件再分解，也不可能 X 同时等于 X_i 中多个值。所以，全部等可能的独立的基本结果 $N=50$ 。 X_i 中有4个值相等等于30公斤。4个属于 $X=30$ 。这时，代入概率的古典定义式便有

$$P(X=30) = \frac{M}{N} = \frac{4}{50} = 0.08$$

知在该班进行一次随机抽测中， $X=30$ 的可能性有8%。

2、概率的统计定义

在统计的实际研究中，由于总体情况往往并不知道， M 和 N 具体数值不明，运用古典定义出现了困难，因而产生了概率的统计定义式。

设在一定条件下，我们完全重复地进行抽样调查或实验，如果我们重复 n 次，事件 A 出现 m 次，则称 $\frac{m}{n}$ 为事件 A 在 n 次出现的频率 f 。当 n 不断扩大时，频率 f 的取值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：

$$P(A) = \frac{m}{n} = P \quad \dots\dots (1-2)$$

这就是概率的统计定义式

定义式(1-1)和定义式(1-2)将有同一的结论，如抛落挑边器，当我们抛落次数逐渐增大时，很容易发现出现兰面和出现白面的频率均趋近于0.5，并在0.5附近左右摆动。如上例在该50人的教学班中，大量重复的抽查，很容易发现 $X=30$

的频率也将趋向于 $\frac{m}{n} = \frac{4}{50} = 0.08$, 且在 0.08 附近摆动。

在运用概率定义式解题时, 须注意分析条件。例如, 在挑边器质量不均匀一面为铜质一面为铝质的条件下, 虽然仍属两个基本事件且相互独立, 但不满足等可能, 这时就不能代入概率的古典定义式。一个运动员的一次投篮, 虽然投中或未投中只两个基本事件, 但由于人的技术结构不一样, 也不能套入古典定义式得出, 投中或投不中的概率均为 $\frac{1}{2} = 0.5$ 。在以上两

种情况下, 便只能用实验结果, 代入统计定义式求解。又如上例体重的抽测中, 如果, 抽测并非随机, 抽测人员偏爱测某种体型的人, 那么, $X = X_i (i=1, 2, 3, \dots, 50)$ 中任何一个的事件也非等可能事件, 也不能用古典定义式求解。因此, 随机性是样本质量的重要条件。概率对随机而言, 不随机就难定义概率。

3. 概率的重要性质

从以上两定义公式, 可得出概率的重要性质。

性质 (I) 概率 P 不可能出现负值, 因 $m > 0$, 故任何随机事件的概率 P , $P > 0$ 。

性质 (II) 当 $m=n$ 时, $P(A) = 1$ 。概率为 100% 的事件, 说明次次必定出现, 故称必然事件。
当 $m=0$ 时, $P(A) = 0$ 。概率为零的事件, 说明一次也不会出现, 故称它为不可能事件。

性质 (III) 当 $m < n$ 时, $P(A)$ 值小于 1。因 m 不可能大于 n , 故任何随机事件的概率不会大于 1, 而处于零与 1 之间, 即 $0 < P(A) < 1$ 。统计所研究的正是这类随机事件。

(二) 概率分布

举例说明, 例如在某教学班中, 决定随机抽测一个学生的体重, 则体重的样本指标 X 便是一个随机变量。设该教学班全部学生体重已知, 总人数 $N=50$ 人, 体重的具体数据为 27, 27, 30, 29, 35……(公斤)。经整理有下表:

表 1—1

(1) 体重 X_i	27	29	30	31	32	33	34	35	36	37
(2) 人数 M_i	2	3	4	7	9	8	7	6	2	2
(3) M_i/n	0.04	0.06	0.08	0.14	0.18	0.16	0.14	0.12	0.04	0.04

第(1)行表示该班不同的10个体重数据 X_i ($i=1, 2, \dots, 10$)。

第(2)行表示不同体重值 X_i 对应的人数 M_i ($i=1, 2, \dots, 10$)。第(3)行 M_i/N 表示不同体重数据与全班总人数之比。

根据概率的古典定义公式(1—1)和已知总体的数据,我们便很容易地得出样本指标 X 的取值范围,最小为27,最大为37,共10个可能的取值。故可算出每个取值的概率。如 $X=27$ 的概率 $P_1(X=27) = \frac{M_1}{N} = \frac{2}{50} = 0.04$,同理有

$$P_2(X=29) = \frac{3}{50} = 0.06, \quad P_3(X=30) = \frac{4}{50} = 0.08.$$

这正是表中第(3)行的数值,相应展开成为一个概率数列,概率的总和为1,故称这数列为概率分布或称分布列。随机变量的概率分布,不尽相同,本例的分布 P_i ($i=1, 2, \dots, 10$)仅只反映本例随机变量的具体规律。

由上例已知,随机变量并非没有规律,它的规律,具体体现于它的概率分布:即存在一个确定的取值范围,且每个取值对应一个确定的概率,故随机变量不可理解为完全没有规律的随意变量,由人意识决定的变量。恰恰相反,应是排除人的主观意识支配的变量。

概率及其分布是统计推断中的重要概念,几乎所有推断理论都是建立在这两个概念的基础上,包括参数估计,两总体的比较、因素分析等,我们虽无法确定具体答案的数值,如差值

究竟多少？但是，却可以根据随机变量的概率分布，给出一个差值的数值范围，通常叫区间，并对这种答案提出概率保证，如提出95%的把握。这种解答方式，是针对随机事物的特点而提出的。

随机变量的规律，具体体现在它的概率分布上，但也应说明，在统计的实践中，并不需要对每个样本变量，给出具体的分布，掌握所有取值的概率；而只需要掌握几种典型的分布类型，掌握少数几个特征参数，能对样本加以判别。本书第五章、第六章将作介绍。

习 题

- 1、试述体育统计的研究对象？
- 2、体育统计为什么要求正确的指导思想和严谨的科学态度？须举例说明。
- 3、试述总体、样本、随机事件、随机变量的概念。
- 4、什么叫随机事件的概率，设已知某班达标测验成绩：优秀5人，良好8人，及格10人，不及格27人，全班总人数 $N=50$ 人。若在该班再随机复测一个学生，求测试结果为优秀、良好、及格、不及格的概率各为多少？
- 5、用上例说明随机变量概率分布的概念。