

实用体育卫生统计方法



沈阳体育学院解剖生理教研室

实用体育卫生统计方法

朱培才 黄庆兴 于 葆 编

朴仁铉 核校

沈阳体育学院解剖生理教研室

前 言

本书是沈阳体育学院《运动医学》补充教材。对广大中小学体育教师、卫生人员、保健教师也适用。

教育部、国家体委、卫生部《关于加强学校体育、卫生工作的通知》中指出：“一个学校体育、卫生工作搞的好不好，最根本的一条是看学生的体质是否有所增强”。衡量学生体质是否增强的指标很多，如身体素质、运动能力、生理、生化和身体发育指标等，但不论用哪一种指标最后都需要经过统计处理。本书就是从这种基本认识出发，为了帮助读者掌握统计方法，更好地开展学校体育、卫生工作而写的。

根据读者的特点，本书的重点放在统计方法的应用上。对于统计原理的阐述，尽量不用数理统计理论上的严格推导，而是采用多举例子讲明道理的办法，做到简明易懂。

本书第六章部分数据由任家华、罗尚文、何正芳老师提供，特此表示感谢。

由于我们的知识和经验有限，书中定有不当之处，欢迎批评指正。

编 者

一九八〇年六月

目 录

第一章 概述.....	1
第一节 统计方法在体育卫生工作中的作用.....	1
第二节 几个基本概念.....	1
一、概率.....	1
二、总体和样本.....	2
三、随机化.....	3
四、计数资料和计量资料.....	3
第二章 统计资料的收集和整理.....	4
第一节 统计资料的收集.....	4
第二节 统计资料的整理.....	6
一、资料检查.....	6
二、资料分组.....	7
三、拟订整理表.....	12
四、归组汇总.....	13
第三章 平均数与标准差.....	15
第一节 平均数.....	15
一、平均数的意义.....	15
二、平均数的计算方法.....	16

第二节 标准差	19
一、标准差的意义	19
二、标准差的计算方法	19
三、标准差的应用	21
第三节 正态分布及检验	23
一、正态分布	23
二、正态分布的检验	26
第四章 抽样误差及其应用	31
第一节 抽样误差的概念和意义	31
第二节 标准误的计算和应用	31
第五章 平均数的显著性检验	33
第一节 显著性检验的意义	33
第二节 各类资料的显著性检验	36
一、样本均数与总体均数差异的显著性检验	36
二、配对资料或同一批对象实验前后的比较	37
三、两个样本平均数显著性检验	39
四、两个大样本平均数的比较	41
第三节 显著性检验应注意的几个问题	42
第六章 相对数	44
第一节 相对数的意义	44
第二节 比(或百分比)	45

一、比的意义	15
二、比的计算与应用	15
第三节 率(或百分率)	52
一、率的意义	52
二、率的计算与应用	52
第四节 注意合理的应用与正确的计算相对数	55
第五节 率的差别显著性检验	57
一、率的标准误	57
二、总体率的可信限	58
三、两个样本率差别的显著性检验	59
第七章 身体发育相关评价表的制定	62
第一节 大样本相关评价表的制作方法步骤	62
第二节 小样本($N < 40$)回归方程的计算方法	75
第三节 回归系数相关系数和的显著性检验	79
附: 辽宁省中小学生身体发育相关评价表和使用方法	83

第一章 概 述

第一节 统计方法在体育 卫生工作中的作用

统计种类很多，有工业统计、农业统计、商业统计、医学统计等。这些统计尽管方法不完全一致，但统计上的基本方法是相同的。体育卫生统计方法，就是这种统计的基本方法在体育卫生科研和实际工作中的应用。

体育卫生工作，特别是学校体育卫生研究工作，需要定期进行体格检查，评定体育教学、训练效果，研究学校常见病多发病和运动性伤病的防治问题等。这些研究就是要透过许许多多的偶然性数据的整理分析，阐明学校体育卫生工作的规律性，从而更好地为增进学生的健康服务。体育卫生统计方法就是解决这个问题的重要工具。因此，体育、卫生工作者必须学习和掌握统计方法。

第二节 几个基本概念

一、概率（机率、然率）

概率是统计中的一个重要概念。它是反映某一事件发生的可能性大小的量。还可以这样说，某事物可能出现的机会与全部机会的比例，叫做概率。现用投掷硬币的例子来说明。投掷均匀的硬币，每投掷一次就会出现两种可能的结

果：正面（带国徽的面）出现或反面（带字的面）出现。我们把每一个可能的结果叫做一个事件。概率研究结果证明，当投币次数增加到无穷多的时候，正反两面出现的机会（概率）各占 $1/2$ 。投币不是出现正面就是出现反面，这是必然发生的事件。因此，必然发生的事件概率为 $1/2 + 1/2 = 1$ 。

不可能发生的事件概率为0。概率用符号P表示，其值的大小用分数、小数或百分数表示。概率用小数表示时，取值范围是0到1。概率越接近1，表明发生的可能性越大。概率接近0，发生的可能性小。我们常在科研论著中见到 $P \leq 0.05$ ，或 $P \leq 0.01$ ，前者表示事物发生的可能性等于或小于0.05，后者表示事物发生的可能性等于或小于0.01。我们常常把这两个概率值作为衡量事物差别是否有显著意义的界限（详见第五章）。

二、总体和样本

总体是同质的个体所构成的集体。总体所包含的个体数可以为无穷大，而且往往是设想的或抽象的。例如我们要了解业体校学生的身体发育情况，所有业体校的学生都具有业体校这个同质的特征，是一个总体。又如，我们要了解儿童少年脊柱侧弯情况，所有脊柱侧弯的儿童少年都具有脊柱侧弯这个特征，也是一个总体。构成总体的每个个体称为变量。在实际调查研究中，往往不需要把所有的研究对象一个不漏的都调查一遍，而从总体中抽出一部分调查就行了。从总体中抽出的这一部分称为样本。抽取的过程称为抽样。抽样调查的方法在科研中最为常用，但是抽样调查的目的决不满足于样本所得的结果，而是通过样本的调查，期望知道总

体的情况。

这样，对抽样调查便提出了一个重要的问题——样本的可靠性问题。解决这个问题，必须注意样本的同质性，样本量的大小要合理，抽样要随机化。属于这几个方面的问题，我们将在以后有关部分做介绍。

三、随机化

为了使样本对总体有较好的代表性，抽样时不能存在偏见或主观愿望，而应遵守随机化原则。所谓随机化，就是在抽样时要使总体中每个单位有同等的被抽取的机会。随机抽样方法很多，有抽签法和随机数字表法等。

四、计数资料和计量资料

统计资料大体上可分为二类，一类是计数资料，另一类是计量资料。它们在统计分析方法上是不同的。计数资料是按研究单位的类别，通过清点数目所得的数据，如体检中发现某病的病人数，体育动作成功和失败的次数，成绩考核的优秀、良好……等数据。常用相对数来表示。

计量资料是对每个研究单位进行测量所得的数据，如身高、体重、胸围、脉搏、血压、血细胞数等数据。常用平均数表示。

第二章 统计资料的 收集和整理

统计工作包括资料的收集、整理和分析三个基本步骤。在这一章里介绍前两个步骤，统计资料的分析放在以后几章里再讲。

第一节 资料的收集

收集资料是统计工作最重要的一步。如果收集资料的计划不周密，原始记录不正确，往往造成整理分析上的困难，甚至得出错误结论，而这些缺点是难以用统计分析来补救的。因此，在收集资料之前必须根据研究目的深思熟虑，不仅要考虑到研究的内容和项目，而且还要考虑如何选择研究对象以及用什么方法进行统计分析，这就是研究工作的设计问题。

体育卫生统计资料的来源有两个方面：①经常性资料，来自医疗卫生机构的原始记录，如医院或卫生室的伤病登记等；②一时性资料，来自专题调查，这是体育卫生资料的主要来源方式。

收集资料需要制定调查表，它的制定是根据调查的目的和所要分析的问题而提出的。调查表包括备考项目和分析项目两个部分。备考项目是为了复核用的，如姓名、住址等；分析项目是按予定要分析的指标而确定的，如调查身体发育

情况需要有身高、体重、坐高、胸围、呼吸差、肺活量、握力、背肌力项目等。

调查表可分为一览表（表1）和单一表或卡片（表2）两种形式。一览表即把许多观察单位写在一张表上，此表比较节约，但不便于整理，易出差错，适用于项目较少、数量较小的调查。单一表或卡片，把每个观察单位写在一张表或卡片上，此法便于整理，适用于大量调查工作，较为常用。

表1 身体发育调查表（一览表举例）

单位地址及名称											编号
检查号	姓名	性别	出生年月日	足岁	身高	体重	胸围	呼吸差	肺活量	坐高	备注

表2 身体发育调查卡片（举例）

姓名	性别	出生 年月日	足岁	编号
住址		职业		
身高	体重	胸围	呼吸差	
肺活量	坐高	握力	背肌力	
检查日期		检查者		

表中的项目要求精简，必要的项目不能没有，不应有无关项目；项目的提法要明确，不应含糊不清，以保证答案的统一性。必要时制定填表说明。

第二节 资料的整理

通过调查所得到的资料比较分散，要了解事物总的特征和发展情况，必须对这些分散的原始资料进行科学的分组归纳，使资料系统化，便于进一步统计分析以反映被研究事物的规律性，这个过程称统计资料的整理。整理工作可分为资料检查、分组、拟订整理表和归纳四个步骤。

一、资料检查

资料检查工作最好是边调查边进行，这样可以使某些错误随时得到纠正。

检查时要查看原始资料有否错漏，并予以纠正。例如发

育调查需要按性别、年龄进行统计，但调查表或卡片上没有指明性别、年龄，这个资料有漏项，是不完整的，应令调查者查明补上。如果发现资料中有自相矛盾的项目，如男性调查表里填了女性指标，或发育指标与实际相差悬殊，如8岁女孩身高记录为180公分等，这样的调查表或卡片则应挑出，叫调查者查明纠正，不能纠正的应淘汰作废。

二、资料分组

按照研究对象的一定标识将调查总体分成若干组称为分组。标识有两种，一种是质量标识，另一种是数量标识。分组就是按照这两种标识分为质量分组和数量分组两种。

质量分组，就是按事物的性质或类型来分组，如运动员的性别、专项类别、动作成功和失败的情况、技评和达标成绩的情况等。

数量分组是在质量分组的基础上再按数量（变量值）的大小来分组，如年龄、脉搏、血压、身高、体重、胸围、时间、红（白）细胞数等。

数量分组的组数多少，取决于资料的性质和资料的多少。分组不可过多，也不可过少。过多一是计算起来比较麻烦，二是使每组例数减少，看不清资料的规律性；过少常常会隐蔽资料的特点，也会看不清资料的规律性。一般变量值在100—200可分成10组左右，变量值在200—300可分成15组左右，例数增多，组数可适当增加。

数量分组的写法应使组限清楚，以避免混乱。以年龄分组为例，正确写法应为：

0—	0—4
5—	5—9
10—	10—14
15—	15—19
20—	20—24
25—	25—29
30—	30—34
35—	35—39
40—	40—44
⋮	⋮
80岁及以上	80岁及以上

上述两种年龄分组表示方法是相同的，但常用左侧的表示方法。数量分组应注意归组问题，即分组界限要清楚。例如右侧分组中，“0—4”岁组，指刚出生婴儿到未满5周岁，“5—9”岁组，指满5周岁到未满10周岁，余此类推。左侧分组中，“0—”表示出生到未满5周岁，“5—”表示满5周岁到未满10周岁。这种分组法组限可以这样记：下限得足，上限不足。“0—”组中，下限为0，上限为5，“5—”组中，下限为5，上限为10……。

下面分组表示方法是错误的：

0—5	25—30
5—10	30—35
10—15	35—40
15—20	⋮
20—25	80岁及以上

之所以说它是错误的,是因为分组界限不清楚,例如0—5, 5—10, 其中5究竟在0—5组内,还是在5—10组内,显然是不清楚的。

为了便于和其他资料对比,还必须注意到习惯分组方法。年龄分组,成人习惯以5岁或10岁为一组。儿童少年应以1岁一组为宜。身体发育指标有下表具体规定(表3)。

表 8 身体发育指标组距

指 标	年 龄	组 距
体 重	幼年儿童	0.5公斤
	少年	1 公斤
	青年和成年	2 公斤
身 高	儿 童	1 厘米
	青年和成年	2 厘米
胸 围	儿 童	1 厘米
	青年和成年	2 厘米
肺活量	少年	100毫升
	青年和成年	250毫升
握 力	少年	1 公斤
	青年和成年	3 公斤
拉 力	少年	2 公斤
	青年和成年	5 公斤

现以160名男学生身高(表4)为例,说明分组的具体步骤。

表4 某地区某年160名某年龄男学生身高记录(厘米)

162.5	170.9	165.7	165.9	160.8	168.0
166.7	160.3	163.3	161.5	177.6	168.9
159.1	164.1	175.5	164.3	156.9	162.8
167.8	171.7	167.2	169.9	171.6	172.1
157.4	170.6	174.6	169.9	160.2	167.5
<u>154.3</u>	160.8	161.9	163.8	160.3	167.2
161.6	161.7	168.5	176.5	172.7	167.4
166.2	162.2	173.9	173.6	171.4	158.8
170.6	169.2	166.3	171.4	171.6	163.9
167.4	162.7	171.3	164.1	169.2	166.9
165.2	175.8	164.1	163.8	171.2	159.5
162.2	165.9	164.0	156.6	168.7	173.0
170.6	159.6	161.4	<u>180.4</u>	161.3	160.5
171.5	158.9	164.7	169.7	162.1	164.1
168.9	173.3	169.0	162.2	154.9	170.3
169.3	162.9	163.5	173.3	158.4	167.3
160.7	157.8	167.4	174.7	168.7	168.8
161.3	172.6	169.6	167.4	167.1	168.3
161.5	170.8	169.6	169.2	164.8	169.3
165.3	165.7	160.9	165.9	164.7	162.3
170.5	162.9	165.2	168.2	166.4	
155.5	170.3	166.7	163.3	157.4	
157.8	164.3	171.5	165.5	165.0	
158.0	160.7	164.8	162.9	162.1	
164.7	164.8	165.6	166.5	166.3	
164.3	163.0	173.8	166.9	167.7	
168.4	166.0	167.6	170.9	170.6	
166.7	161.7	170.6	156.3	168.4	

1、先求出全距：一套变量值中最大值和最小值之差称为全距，本例为 $180.4 - 154.3$ （取整数） $= 26$ （厘米）。

2、确定组距和组数：组距（ i ）指的是组间宽度，应根据样本例数多少而确定的组数来确定，关于组数多少的问题前面已经讲过，一般分为10—15组为宜。根据这个要求，

本资料可分为13组，组距定为2（厘米）。求法是：

$$\text{组数} = \frac{\text{全距}}{\text{组距}} = \frac{26}{2} = 13$$

3、**确定组限：**定出每一小组的组限。注意第一组要包括最小值，最末一组要包括最大值。第一组开始的数值取等于或小于最小值的整数，本例为154（厘米）。第一组开始界限可由154厘米开始，然后以组距2厘米为距离，每隔2厘米作为一组写出相继各组的组限。各组组限数值小的称为下限，数值大的称为上限，如第一组开始数值—154称为第一组下限，第一组上限是156，此值同时又是第二组的下限，第二组的上限是158，此值同时也是第三组的下限……依此类推，直至末组上限为182包括最大值180.4。

表4资料可分成如下若干组（表5）

表5 160名男学生身高

身高（厘米）	划 记	频 数
154—		
156—		
158—		
160—		
162—		
164—		
166—		
168—		
170—		
172—		
174—		
176—		
178—		
180—		
合 计		