

◆ 潘玉进 编著

教育与心理统计

——*SPSS* 应用



ZHEJIANG UNIVERSITY PRESS
浙江大學出版社

教育与心理统计

——SPSS 应用

潘玉进 编著

浙江大学出版社

图书在版编目(CIP)数据

教育与心理统计——SPSS应用 / 潘玉进编著. —杭州:
浙江大学出版社, 2005.12

ISBN 7-308-04575-7

I.教... II.潘... III.①统计分析—统计程序—
应用—教育统计—多元分析②统计分析—统计程序—应
用—心理统计—多元分析 IV.①G40-051②B841.2

中国版本图书馆 CIP 数据核字 (2005) 第 148809 号

责任编辑 李桂云

出版发行 浙江大学出版社

(杭州市天目山路 148 号 邮政编码 310028)

(E-mail: zupress@mail.hz.zj.cn)

(网址: <http://www.zjupress.com>)

排 版 浙江大学出版社电脑排版中心

印 刷 浙江省良渚印刷厂

开 本 787mm×1092mm 1/16

印 张 17.75

字 数 453 千

版 印 次 2006 年 2 月第 1 版 2006 年 5 月第 2 次印刷

书 号 ISBN 7-308-04575-7/G·1004

定 价 23.00 元

目 录

第一篇 教育与心理统计的基本原理

第一章 绪 论	2
第一节 概述	2
一、什么是教育与心理统计	2
二、教育与心理统计基本的内容	3
三、教育与心理统计的作用	3
第二节 统计学中的几个基本概念	4
一、随机现象与随机事件	4
二、总体与样本	4
三、参数与统计量	5
四、误差	5
第三节 数据的种类	6
一、点计数据与度量数据	7
二、离散数据与连续数据	7
三、称名数据、顺序数据、等距数据与比率数据	7
第二章 数据的特征量	9
第一节 集中量	9
一、算术平均数	9
二、中位数	9
三、众数	10
第二节 差异量	10
一、常用的差异量	10
二、方差与标准差的特点	11
第三节 相关量	12
一、相关的概念	12
二、相关的种类	12

三、相关系数	14
第三章 概率及概率分布	15
第一节 概率的一般概念	15
一、概率的定义	15
二、概率的性质	16
第二节 二项分布	16
一、二项试验	16
二、二项分布函数	17
三、二项分布图及其平均数与标准差	17
四、二项分布的应用	18
第三节 正态分布	18
一、正态曲线	18
二、标准正态分布	19
三、正态分布曲线的特点	19
第四章 统计推断的基本原理	20
第一节 统计推断概述	20
一、统计推断的意义	20
二、随机抽样	20
三、统计推断的内容	21
第二节 抽样分布	21
一、抽样分布的概念	21
二、关于样本平均数抽样分布的两个定理	22
三、样本平均数与总体平均数离差统计量的形态	22
第三节 总体平均数的估计	23
一、点估计和区间估计的概念	23
二、区间估计的原理和方法	24
第四节 假设检验的基本原理	25
一、假设检验的基本原理	25
二、假设检验的几个基本概念	26

第二篇 SPSS for Windows 基本操作

第五章 SPSS for Windows 概述	30
第一节 SPSS for Windows 的特点	30
第二节 SPSS for Windows 对环境的要求	31
一、SPSS for Windows 对硬件环境的要求	31
二、SPSS for Windows 对软件环境的要求	31
第三节 SPSS for Windows 的启动与退出	31
一、启动 SPSS for Windows	31
二、退出 SPSS for Windows	32
第六章 数据文件的建立与编辑	33

第一节 数据文件的建立	33
一、定义变量	33
二、输入数据	36
三、保存数据文件	37
四、读入其他类型的数据文件	38
第二节 数据文件的编辑	38
一、数据的搜索	38
二、变量的插入与删除	39
三、记录的插入和删除	39
四、根据已存在的变量建立新变量	39
五、变量排序	41
第七章 数据文件的整理	45
第一节 数据的排序与行列转置	45
一、数据的排序	45
二、数据的行列转置	45
第二节 数据文件的分组、筛选与合并	46
一、数据文件的分组	46
二、数据文件的筛选	47
三、数据文件的合并	48

第三篇 SPSS for Windows 统计分析过程

第八章 描述性统计分析	52
第一节 Frequencies 过程	52
一、概述	52
二、SPSS 操作过程	53
三、实例	55
第二节 Descriptives 过程	57
一、概述	57
二、SPSS 操作过程	58
三、实例	58
第三节 Crosstabs 过程	59
一、概述	59
二、SPSS 操作过程	59
三、实例	62
第九章 总体平均数的假设检验	64
第一节 单一样本 t 检验	64
一、概述	64
二、SPSS 操作过程	64
三、实例	65
第二节 总体平均数差异的显著性检验	67
一、概述	67

二、统计检验量	68
第三节 相关样本 t 检验	69
一、概述	69
二、SPSS 操作过程	69
三、实例	70
第四节 独立样本 t 检验	72
一、概述	72
二、SPSS 操作过程	72
三、实例	74
第十章 方差分析	78
第一节 方差分析的基本原理	78
一、概述	78
二、方差分析的基本原理	78
三、方差分析的假设检验	80
四、方差分析中的几个术语	80
五、方差分析过程	82
第二节 One-Way ANOVA 过程	83
一、概述	83
二、SPSS 操作过程	83
三、实例	86
第三节 Univariate 过程	90
一、概述	90
二、SPSS 操作过程	91
三、实例	96
第四节 Multivariate 过程	107
一、概述	107
二、SPSS 操作过程	108
三、实例	109
第五节 Repeated Measures 过程	117
一、概述	117
二、SPSS 操作过程	118
三、实例	120
第六节 协方差分析	148
一、概述	148
二、SPSS 操作过程	149
三、实例	149
第十一章 非参数检验	167
第一节 概述	167
一、参数检验与非参数检验	167
二、非参数检验的种类	167
第二节 χ^2 检验	168
一、概述	168
二、拟合优度 χ^2 检验的 SPSS 操作过程	169

三、独立性或同质性 χ^2 检验的 SPSS 操作过程	170
四、实例	171
第三节 单一样本 K-S 检验	174
一、概述	174
二、SPSS 操作过程	175
三、实例	175
第四节 两个独立样本的非参数检验	177
一、概述	177
二、SPSS 操作过程	178
三、实例	179
第五节 多个独立样本的非参数检验	183
一、概述	183
二、SPSS 操作过程	184
三、实例	184
第六节 两个相关样本的非参数检验	187
一、概述	187
二、SPSS 操作过程	188
三、实例	188
第七节 多个相关样本的非参数检验	191
一、概述	191
二、SPSS 操作过程	192
三、实例	193
第十二章 相关分析	197
第一节 相关分析的概念与过程	197
一、概述	197
二、相关分析的 SPSS 过程	197
第二节 积距相关与等级相关	198
一、概述	198
二、SPSS 操作过程	199
三、实例	201
第三节 质与量的相关和品质相关	204
一、概述	204
二、SPSS 操作过程	206
三、实例	206
第四节 偏相关分析	216
一、概述	216
二、SPSS 操作过程	217
三、实例	217
第十三章 回归分析	221
第一节 线性回归	222
一、概述	222
二、SPSS 操作过程	226
三、实例	230

第二节 曲线估计	239
一、概述	239
二、SPSS 操作过程	240
三、实例	242
第十四章 因子分析	246
第一节 因子分析的基本原理	246
一、概述	246
二、因子分析的基本原理	246
第二节 因子分析过程	248
一、概述	248
二、SPSS 操作过程	248
三、实例	252
第十五章 信度分析与多选题分析	260
第一节 同质性信度	260
一、概述	260
二、SPSS 操作过程	260
三、实例	262
第二节 分半信度	265
一、概述	265
二、SPSS 操作过程	265
三、实例	265
第三节 再测信度	266
一、概述	266
二、SPSS 操作过程	266
三、实例	267
第四节 多选题分析	267
一、概述	267
二、SPSS 操作过程	268
三、实例	270
参考文献	276

第一篇

教育与心理统计 的基本原理

第一章 绪 论

第一节 概 述

一、什么是教育与心理统计

教育与心理统计是运用数理统计学的原理与方法,研究教育与心理现象的一门学科,其主要任务是研究如何收集、整理和分析教育与心理研究中的数字资料,并以此为依据进行统计推断,揭示教育与心理现象所蕴含的客观规律。

例如,要研究小学生的孤独感与父母教养方式的关系,其研究对象应该是所有的小学生,从地域上看,不仅包括研究者所在的学校、县、市,还应该包括所在的省,乃至全国所有的小学生;从时间上看,不仅包括现在的小学生,更应该包括未来若干年的小学生,否则将由于研究结果不能推广应用而使研究变得毫无意义。要全面观测如此庞大的研究对象,需要投入大量的人力、物力、财力与时间,这对一个小小的课题组来说是很困难的,而且对于未来若干年的小学生而言也是不可能的。在这种情况下,教育与心理统计就大有用武之地了。研究者可以根据研究目的,在研究对象的范围内适量地随机抽取一小部分小学生,比如随机抽取 200 个小学生进行观测,测量这 200 个小学生的孤独感及其父母教养方式,然后根据数理统计学中业已证明的有关原理与方法在一定的可靠性程度上(比如在 95% 的可靠性程度上)推测所有小学生的孤独感与父母教养方式的关系。

在这里,所有的小学生在统计学上称为“总体”,是研究的对象;而从中随机抽取的 200 个小学生则称为来自于这个总体的“样本”,是研究中进行观测的对象。也就是说,可以根据从总体中随机抽取的一定量的样本数据来推测总体的特征,这就是统计推断的基本思想,如图 1.1 所示。因此,统计学的重要意义就在于对不可能获得的总体,能对其各种特征及分布性质作出一定可靠程度的估计和推测。

数理统计学是以概率论为基础,对统计数据数量关系的模式加以解释,对统计原理和方法给予数学的证明,是数学的一个分支。数理统计学的原理与方法在各个领域中的应用称为应用统计学。教育与心理统计就是数理统计学的基本原理与方法在教育与心理研究领域中的具体

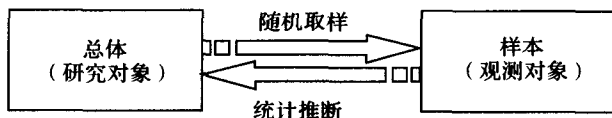


图 1.1 统计推断的基本思想

运用。

二、教育与心理统计基本的内容

教育与心理统计的基本内容包括:描述统计、推断统计和实验设计三部分。

1. 描述统计

对所收集的数据进行整理、概括,显现其分布特征的统计方法,称为描述统计(Descriptive Statistics)。对教育与心理研究中所获得的大量数据进行归纳、整理后,可通过编制统计图表来直观形象地反映其分布特征,也可通过计算各种数字特征量来概括地反映其分布特征。例如,计算集中量(如算术平均数、中位数、众数等)来反映一组数据的集中趋势,计算差异量(如平均差、标准差、差异系数等)来反映一组数据的离散程度,计算相关量(如积距相关系数、等级相关系数等)来反映两组数据之间变化的一致性程度。

2. 推断统计

根据样本的数字特征量,运用数理统计学的原理与方法,在一定的可靠性程度上,对总体特征进行估计和推断,称为推断统计(Inferential Statistics)。推断统计包括参数估计(分为点估计和区间估计)和假设检验(如 t 检验、方差分析、 χ^2 检验等)两部分。

3. 实验设计

实验设计(Experimental Design)是研究者为了揭示自变量与因变量的因果关系所制订的实验计划与方案,其中包括自变量如何操作、因变量如何检测、控制变量如何控制;如何取样、取多少被试、被试如何分组;实验数据如何采集、用何种统计方法进行处理与分析,等等。

三、教育与心理统计的作用

教育与心理统计具有以下三方面作用:

1. 为教育与心理研究提供一种科学方法

教育与心理研究的根本任务就是对教育与心理现象进行观测,揭示蕴藏于其中的客观规律。要提高对各种教育与心理现象的观测与分析研究能力,就必须运用科学的方法。如图1.1所示,教育与心理统计正是提供了这样一种科学方法。

2. 教育与心理统计是开展教育与心理研究的重要工具

对教育与心理现象的研究有定性分析与定量分析。在教育科学与心理科学发展的早期阶段,定量分析很少得到运用,许多研究者仅仅从个人的经验出发,用思辨的方法对教育与心理现象进行研究,其研究结果往往具有较大的主观随意性。随着教育与心理测量的逐渐发展,教育与心理统计得到越来越广泛的运用,成为教育与心理研究中重要的定量分析工具,它可以对研究中所获得的数据进行统计分析,以揭示其中所蕴含的客观规律。

3. 可顺利阅读运用统计方法进行定量分析的研究文献

在现代教育与心理研究文献中,大多采用各种统计方法来表述、说明与解释其研究成果。如果不掌握统计的符号、概念及思维方法,就不能顺利阅读有关文献,无法学习先进的研究成果。

第二节 统计学中的几个基本概念

一、随机现象与随机事件

1. 随机现象

在自然界和人类社会中,有时一旦满足某种条件,就必然会发生某种现象,这称为确定性现象。如化学中的强酸与强碱的中和反应,或是一个盒子中有十个完全相同的白球,搅拌后从中任意摸取一球,必是白球,等等。但也有些现象是非确定性现象,也就是说,在一定的条件下,可能发生也可能不发生的现象,这就是随机现象(Random Phenomenon)。

随机现象具有三个特征:

- (1)一次试验有多种已知的可能结果;
- (2)在试验之前其结果无法预料;
- (3)在相同的条件下可重复试验。

例如,从某班级随机抽取两名学生,有三种已知的可能结果:二男、二女、一男一女;究竟是哪种结果,试验之前无法预料;在相同的条件下(每次抽完后归队)可以重复抽取。又如,抛一枚硬币,有两种已知的可能结果:正面朝上或反面朝上;究竟哪一面朝上,试验之前不能预料;相同的条件下(如刚性平面、举到一定的高度等)可以重复抛掷。以上这些都属随机现象。

2. 随机事件

随机现象的某种可能结果,称为随机事件(Random Event)。随机事件在一次试验中,可能发生,也可能不发生,表现出随机性;但是,在大量的重复试验中,它们又表现出统计规律,即随着试验次数的增加,每一个随机事件发生的次数 f (也称为频数)在试验总次数中所占的比率 f/N (这个比率,称为该随机事件的频率,用 W 表示,即 $W = f/N$) 都会逐渐稳定地逼近某一恒定值。

例如,设硬币的正面朝上为随机事件 A ,反面朝上为随机事件 B 。在抛一次硬币时,随机事件 A 可能发生,也可能不发生,但如果重复抛许多次,随机事件 A 的发生就会呈现出某种规律性,即它出现的频率接近 0.5,如表 1.1 所示。

表 1.1 抛掷硬币的试验情况

试验次数		10000	20000	30000	40000	50000	频率所逼近 的恒定值
随 机 事 件							
正 面 朝 上 (A)	频 数	4875	9922	14941	19934	24946	0.5
	频 率	0.4875	0.4961	0.4980	0.4984	0.4989	
反 面 朝 上 (B)	频 数	5125	10078	15059	20066	25054	0.5
	频 率	0.5125	0.5039	0.5020	0.5016	0.5011	

二、总体与样本

1. 总体

总体(Population)是指所研究的具有某种共同特征的个体的总和。总体可以是人,或是人所具有的人格特征、活动方式等,如全市高一学生、全省初二女生、沿海地区初中生的智力发展

水平,等等;也可以是物,如所有的常用字、某次考试的全部试卷、某次测验的所有分数,等等。

组成总体的基本单元称为个体(Individual),如一个学生、一份试卷、一个测验分数等。总体分为有限总体和无限总体。包含有限个体的总体称为有限总体(Finite Population),如某校的学生、某市的小学中高年级学生、参加 2005 年高考的所有考生等;包含无限个体的总体称为无限总体(Infinite Population),如所有的高一学生,理论上包括了现在以及未来若干年的所有高一学生,也包括全校、全县、全省,乃至全国所有的高一学生。

如前文所述,要对总体中的所有个体都进行观测,不仅是不可能的,也是不必要的,因为这要花费大量的人力、物力、财力和时间,一般难以实现。因此,必须从总体中随机抽取一部分个体进行观测,再根据这部分个体所提供的信息在一定的可靠性程度上对总体的特征与分布性质进行推断,如图 1.1 所示。

2. 样本

从总体中随机抽取的作为观测对象的一部分个体称为样本(Sample)。样本所包含的个体数目称为样本容量(Sample Size),一般用 n 表示。如某样本包含 25 个个体,记为 $n = 25$ 。

总体与样本是相对的。如研究某校学生的智力水平,那么全校学生的智力水平就是总体,从中随机抽取的 20 个学生的智力水平就是样本;如要研究全市学生的智力水平,那么该校的学生智力水平就可以作为样本。

三、参数与统计量

描述总体分布特征的量数,称为参数(Parameter)。例如,描述总体集中趋势的有总体平均数(μ),描述总体内个体间离散程度的有总体标准差(σ),描述两个总体之间相关关系的有总体相关系数(ρ),等等。

描述样本分布特征的量数,称为统计量(Statistic)。例如,描述样本集中趋势的有算术平均数(Mean)、中位数(Median)、众数(Mode)等;描述样本离散程度的有四分位数(Quartile)、差异系数(Coefficient of Variation)、标准差(Standard Deviation)等;描述两个样本之间相关关系的有相关系数(Correlation Coefficient)等。

根据样本统计量来估计与推断总体参数,正是统计推断的内容之一。

四、误差

教育与心理统计中所说的误差,泛指观测值与真值之差。主要有以下三种:

1. 系统误差

在数据采集过程中,由于仪器不准确,指导语有暗示性,或对某些标准掌握过宽、过严等原因,导致数据成倾向性地偏大或偏小而引起的误差叫系统误差(Systematic Error)。如每年的高考,虽然各科试卷都有统一的评分标准,但由于各地在掌握标准上有宽严的差异,以致标准掌握得稍宽的地区,考生的得分会成倾向性地偏高,而标准掌握得较严的地区,考生的得分会普遍地偏低,像这样的误差就属于系统误差。

系统误差有一定的倾向性,如恒定性:在多次测试中,如果观测值比真值偏高就都普遍偏高,比真值偏低则都普遍偏低;累增性:用同一方式多次观测时,误差会越来越大;周期性:如果观测系统的某部分发生偏差时,每到观测这部分时都会发生系统误差。

系统误差影响着原始数据的准确性,必须加以消除。但由于系统误差难以从原始数据中分析出来,因此其发现和消除较其他误差更困难。消除系统误差的措施是仔细分析观测的方

法、量表及仪器,查明产生系统误差的原因,然后调整观测系统,重新进行观测;此外还可以改用另一种方法进行观测来确定是否存在系统误差。

2. 随机误差

在数据采集过程中,由于存在一些不易发现或无法控制的偶然因素,致使同一对象经同一方式观测多次,其结果都不一样,这样产生的误差叫随机误差(Random Error)。

随机误差一般都由众多微小的偶然因素造成。努力达到观测系统的精确性,可以缩小随机误差,但无法完全避免。无论观测的方法如何正确,仪器如何精密,量表如何科学,随机误差总是存在的,但应控制在允许的范围之内。

随机误差虽不能完全消除,但它却有一定的统计规律性。这表现在以下几方面:

- (1) 随机误差不恒定,有时为正,有时为负;
- (2) 正负误差出现的可能性大致相等;
- (3) 小误差出现的可能性大于大误差出现的可能性,特别大的误差一般不会出现;
- (4) 同一对象被观测多次时,随着观测次数的增加,误差的算术平均数逐渐接近于零。

3. 抽样误差

随机样本的统计量与总体参数之差,称为抽样误差(Sampling Error)。如样本平均数与总体平均数的差异、样本标准差与总体标准差的差异,等等。抽样误差越小,则样本对总体的代表性越强;抽样误差越大,则样本对总体的代表性越弱。所以,抽样误差是测定样本代表性的指标。

在教育与心理研究中,即使消除了系统误差,并把随机误差控制在允许的范围之内,但随机样本的结构不可能与总体结构完全一致,所以,抽样误差不可避免。

抽样误差受样本容量、总体标准差和抽样方法的制约。在其他条件相同的情况下,样本容量越大,抽样误差越小,样本容量越小,抽样误差越大,样本容量与抽样误差成反比;总体标准差越大,抽样误差也越大,总体标准差越小,抽样误差也越小,总体标准差与抽样误差成正比;在样本容量相同时,分层随机抽样比简单随机抽样的抽样误差小。

样本统计量与总体参数既有差别,又有明显的一致趋势,是抽样误差的显著特点。这一特点表明样本与总体的辩证关系。总体所具有的必然性寓于随机获得的、带偶然性的样本之中,并通过样本数据反映出来。从个别数据看,误差有大有小,好像捉摸不定;但从大量数据看,从整体看,它是有规律性的。抽样误差虽然不能避免,但因它是随机产生的,故可根据概率统计原理估计其大小和范围,并通过抽样程序加以控制,将其缩小到最低限度和控制允许的范围之内。

第三节 数据的种类

随机现象有多种可能结果,这种能表示随机现象各种可能结果的变量称为随机变量(Random Variable),统计处理的变量都是些随机变量。例如,学生的性别、智商、学科考试成绩,教师的年龄、教龄、教学水平,等等。可用 x, y 等符号分别表示不同的随机变量。例如,用 x 表示某班语文测验分数这一随机变量;用 y 表示数学分数这一随机变量。而该班每个学生的语文分数可以分别用 $X_1, X_2, \dots, X_n$ 表示,数学分数可以分别用 $Y_1, Y_2, \dots, Y_n$ 表示。

数据(Data)是随机变量的观测值,它是用来描述对客观事物观察测量结果的数值。当对

某个随机变量进行观测时,事先不能预料会取到什么值;一旦某个值被取定,就称这个值为随机变量的一个观测值,即数据(也是一个随机事件)。

数据的种类不同,统计处理的方法也不同。根据不同的分类标准,数据可分为不同的类型。

一、点计数据与度量数据

根据数据的来源,可将数据分为点计数据和度量数据。点计数据(Count Data)是指通过计算个数所得到的数据。如学校数、班级数、学生数、教师数、教室数、图书册数等。点计数据一般取整数形式。度量数据(Measurement Data)是指用一定的测量工具对事物进行测量,或者按照一定的测量标准给对象赋值而得到的数据。如身高、体重、学科成绩、智商、反应时间、记忆保持量等。

二、离散数据与连续数据

根据数据的取值情况,可将数据分为离散数据和连续数据。取值个数有限的数据,称为离散数据(Discrete Data)。这种数据的单位是独立的,两个单位之间不能再划分成更小的单位,一般用整数表示。例如,学生人数、某门课程不及格的人数、等级制的学科成绩等。取值个数无限的数据,称为连续数据(Continuous Data)。它们可能的取值范围能连续充满某一个区间,数据的单位之间可以再划分成无限多个细小的单位,一般可以用小数表示。例如,学生的身高、体重、智商、百分制的学科成绩等。

三、称名数据、顺序数据、等距数据与比率数据

根据数据的测量水平,可将数据分为称名数据、顺序数据、等距数据与比率数据。

1. 称名数据

称名数据(Nominal Data)只表示某一事物与其他事物在属性上的不同。它既不能排序,也没有相等单位,更没有绝对零点。如:运动员的号码、学生的学号、性别、衣服的颜色、不同的教学方法等等。在这里,只给不同属性、不同类别的事物以不同的名称,这些名称的排列顺序无关紧要,在处理这类数据时,不能进行加减乘除运算。

2. 顺序数据

顺序数据(Ordinal Data)是指可以就事物的某一属性的多少或大小依次排序,但没有相等单位和绝对零点的数据资料。如:比赛中的第一、二、三名,学科成绩的优、良、中、合格、待合格,等等。对于这类数据,不仅可以从所赋予的名称来区别各事物之间的不同,而且可以排出它们在某一特质方面的多少或大小。顺序数据与称名数据的主要区别在于它有“方向次序”,但在处理这类数据时,也不能进行加减乘除运算。

3. 等距数据

等距数据(Interval Data)是指具有相等单位,但无绝对零点的数据资料。例如气温,假设周一的气温为 10°C ,周二的气温为 15°C ,周三的气温为 20°C ,首先,周一、周二、周三的气温各不相同;其次,周三的气温高于周二,周二的气温高于周一;第三,周三与周二的气温之差,等于周二与周一的气温之差。但等距数据没有绝对零点,因此它只是一个相对的数值,而不能作比率性描述。在对这类数据进行统计处理时,只能进行加减运算,而不能进行乘除运算。

4. 比率数据

比率数据(Ratio Data)是指具有相等单位和绝对零点的数据资料。这类数据可以作比率性描述,在统计处理时,既可进行加减运算,也可进行乘除运算。如:身高、体重等。假如一个成年人的身高是 176cm,一个儿童的身高是 88cm,那么,该成年人的身高比该儿童高 88cm,同时也是该儿童身高的两倍。

不同测量水平的数据之间可以进行转换,但只有比率数据或等距数据可以转换为顺序数据,而且转换后数据所包含的信息量将会损耗。例如,百分制的学科成绩(等距数据)可以按照一定的规则转换为等级制的学科成绩(顺序数据),如:60 分以下为“待合格”,60~69 分为“合格”,70~79 分为“中”,80~89 分为“良”,90 分及以上为“优”。但应该看到,转换后的数据只保留了彼此间大小的顺序关系,而相差的量已被抛弃。

各种不同类型的数据之间存在一定的关系。首先,点计数据必然属于离散数据,而连续数据由于无法计数,必然属于度量数据。例如,学生的人数是点计数据,同时也是离散数据;而学生的身高是连续数据,同时也是无法计数的,必须用工具去度量,属于度量数据。但离散数据可能是点计数据,也可能是度量数据;度量数据可能是连续数据,也可能是离散数据。例如,等级制的学科成绩既是离散数据,同时也是度量数据。其次,比率数据和等距数据属于连续数据,顺序数据和称名数据属于离散数据。例如,学生的体重(比率数据)、百分制的学科成绩(等距数据)均属于连续数据,而等级制的学科成绩(顺序数据)和学生的性别(称名数据)均属于离散数据。