

浅谈标准分数及标准差

李正龙

怎样评价一个班级教学任务完成的好坏?怎样评价一个学生学业水平的高低?目前广为采用的办法,对班级是看平均分,及格率,有的学校还看高分率(即80分以上的人数比例);对于学生个体,则以学生原始分数的高低或以各科总分之和来衡量。这些办法是不够完善的,有的甚至是不够科学的。随着标准化考试的逐步实施,对考试团体及个人的成绩评估,还必须引入标准差及标准分数这两项内容。掌握这两项评估数据是今后每位教师的基本功。由于计算器为我们提供了运算工具,掌握它们并不困难。本文旨在对标准差和标准分数的意义及计算办法作一简介。

一、标准差简介

标准差是与平均数相对应的差异量数,它反映了考试团体中考生成绩分布的离散程度,即波动程度。标准差较大,说明分数分布的离散程度较大,考生有两极分化现象;标准差较小,说明分数分布的离散程度较小,考生分数差距不大。假如每个考生的成绩都相同,则标准差为零,但在实际上是不可能的。

在统计学上,标准差分为总体标准差和样本标准差,单位是分。我们经常采用的是样本标准差,以下提到的标准差均指样本标准差。

标准差为多少分比较好呢?这涉及到试卷的总分,难度及考试的总人数等因素,不能一概而论。一般来说,相比较而言,适中较好。标准差过大,说明考生水平参差不齐,不能体现面向全体学生,让多数学生学有所得,或说明试题区分度较大;标准差过小,说明考生水平接近,尖子很难脱颖而出,差生也不明显,或者说明试题区分度较

组别	分数分布情况	平均分	标准差
甲组	69, 70, 70, 70, 71	70	0.71
乙组	66, 70, 70, 70, 74	70	2.83
丙组	54, 60, 70, 73, 93	70	14.95
丁组	38, 57, 69, 88, 98	70	23.99

小。为了说明这个问题，让我们看以下实例：

例 1，有四组学生，每组 5 人，他们的某科考试成绩如上表，假定试题难度适中，试评价各组学生学业水平。

分析：虽然这四组学生平均成绩都相等，但甲组最整齐，乙组次之，丁组最分散，因此这四组学生的学业程度是不一样的，其中丙组的分布较为理想。因此标准差适中较好。这里应当说明的是：当几个考试团体进行比较时，考生人数必须相差不大，考生学前水平也应当基本接近，这样按标准差来评估教学时才会得出让人信服的结论。

二、怎样计算标准差

在统计学中，标准差按以下公式计算：

$$S = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

S 代表标准差， x 代表各考生原始分数， \bar{x} 代表团体平均分， Σ 表示求和，即将每个考生的得分与平均分相减后平方再逐一相加， n 表示考生人数。

在不使用计算器或电子计算机时，标准差的计算是很繁琐的，容易发生错误。但我们借助于计算器时，就十分容易了。不过计算器必须是函数计算器，才具有统计功能。现用国产 ST-550 型计算器加以说明，若用其它函数计算器方法也大同小异，参看使用说明书即可。

例 2，计算例 1 中四组学生的平均分和标准差。

操 作	显 示	说 明
开机, \boxed{iNV} \boxed{MODE}	0. (并出现 SD 字样)	进入统计状态
\boxed{iNV} \boxed{AC}	0.	消除内残
$\boxed{69}$ \boxed{x} $\boxed{70 \times 3}$ \boxed{x} $\boxed{71}$ \boxed{x}	71	输入数据
\boxed{x}	70	甲组平均分
$\boxed{\sigma_{n-1}}$	0.707106781	甲组标准差
\boxed{iNV} \boxed{AC}	0	消除内残
$\boxed{66}$ \boxed{x} $\boxed{70 \times 3}$ \boxed{x} $\boxed{74}$ \boxed{x}	74	输入数据
\boxed{x}	70	乙组平均分
$\boxed{\sigma_{n-1}}$	2.828427125	乙组标准差
\boxed{iNV} \boxed{AC}	0.	消除内残
$\boxed{54}$ \boxed{x} $\boxed{66}$ \boxed{x} $\boxed{70}$ \boxed{x} $\boxed{73}$ \boxed{x} $\boxed{93}$ \boxed{x}	93	输入数据

续上表

操 作	显 示	说 明
\bar{x}	70	丙组平均分
σ_{n-1}	14.94991639	丙组标准差
$\overline{INV} \quad AC$	0	消除内残
38 \times 57 69 \times 88 \times 98 \times	98	输入数据
\bar{x}	70	丁组平均分
σ_{n-1}	23.99958107	丁组标准差
$\overline{INV} \quad MODE$	0.(SD字样消失)	退出统计状态

如果考生人数众多,也可用以下办法计算标准差:

例3、某地区语文统考分数段人数如下表,计算标准差。

分数段	0—9	10—19	20—29	30—39	40—49	50—59	60—69	70—79	80—89	90—100
人 数	0	0	37	56	114	278	363	146	83	27

操 作	显 示	说 明
开机, $\overline{INV} \quad MODE$	0.(并出现SD字样)	进入统计状态
$\overline{INV} \quad AC$	0	清除内残
$25 \times 37 \quad \times \quad 35 \times 36 \quad \times \quad 45 \times 114 \quad \times \quad 55 \times 278$	95	取各分数段的中间 值为平均分输入
$\times \quad 65 \times 363 \quad \times \quad 75 \times 146 \quad \times \quad 85 \times 83 \quad \times \quad 95 \times 27 \quad \times$		
n	1104	总 人 数
σ_{n-1}	14.70253923	标 准 差

注意:在输入数据时,必须按“分数 \times 人数”输入,不可按“人数 \times 分数”输入,

例 $25 \times 37 \quad \times$ 不可操作为 $37 \times 25 \quad \times$ 。

下面提供某班某次考试数据,请读者试计算平均分和标准差。

28, 99, 23, 84, 28, 55, 60, 55, 58, 73, 40, 75, 35, 15, 55, 63, 85, 60,
43, 30, 61, 40, 98, 83, 90, 45, 52, 75, 55, 45, 50, 65, 30, 83, 80, 30, 55,
43, 50, 71, 50, 70, 61, 75, 73, 60, 50, 43, 83, 55, 68。

(答案:平均分 $\bar{x} = 57.9$, 标准差 $S = 19.83$)

三、标准分数简介

在评价学生个体水平时，光凭学生的原始分数或各科总分来评价实际上是不科学的。因为各次或各科考试的难度并不一致也很难做到一致，因此各次或各科考试的分数并不等值，就象不同国家的同数额货币并不等值一样。如果引入标准分数，就能较客观地评价一个学生。

标准分数就是选定了共同的参照点，具有共同的分数单位的分数。这个参照点，就是各次考试或各科考试的平均分。成绩高于平均分者，标准分为正分；低于平均分者，标准分为负分；等于平均分者，标准分为零分。在统计学中，计算公式为：

$$Z = \frac{x - \bar{x}}{s}$$

Z 指标准分数， x 指原始分数， \bar{x} 指平均分， s 指标准差。

标准分最高和最低各为几分呢？经测量表明，-3到+3包括了全部数据的99.73%，在特殊情况下（例如标准差很小时），标准分也会在-5到+5之间。

经过这样的分数转换，学生同科各次或各科同次的考试成绩就有了一个科学的参照点，具有了可比性和可加性，请看以下几例实际应用：

例4、比较同一学生不同阶段的考试成绩并作出评估。

时 间	成 绩	全 班 平均分	标准差	标准分	结 论
上学期	85	90	20	-0.25	下学期成绩不但没有退步，而且 进步很快。
下学期	75	65	10	1	

例5、比较同一学生不同学科的考试成绩并作出评估。

学 科	成 绩	全 班 平均分	标准差	标准分	结 论
语 文	66	65	13	0.08	该生在班级中语文成绩优于数学 成绩
数 学	74	75	10	-0.81	

例 6、比较四个学生几科的考试成绩并作出评估。

学 科	全 班		考 试 成 绩				标 准 分 数			
	平均分	标准差	甲	乙	丙	丁	甲	乙	丙	丁
语 文	75	9	83	65	93	72	0.89	-1.11	2.0	-0.3
数 学	50	5	70	70	50	55	4.0	4.0	0	1.0
外 语	32	4	40	50	30	36	2.0	4.5	-0.5	1.0
总 计	—	—	193	185	173	163	6.89	7.39	1.5	1.7

结论：这四名学生的能力强弱顺序是：乙、甲、丁、丙。

下面请读者试按下表所列数据评价甲乙两学生实际能力的高低。

考试科目	全班平均分	标准差	甲生成绩	乙生成绩	甲生标准分	乙生标准分
数 学	48.3	13.9	73	54		
物 理	66.9	18.5	79	94		
化 学	67.2	14.0	75	82		
总 计	—	—	227	230		

(答案：甲生实际能力强于乙生。)

需要说明的是，标准化考试并不是以标准分数来进行测量的考试，但标准分数是标准化考试中的一项评估指标，也可以作为普通考试的评估指标，它对我们的教学评估工作是有积极意义的。

(责任编辑 闵近义)