标准分制度及其应用

扈涛 刘刚 编著

科 普 出 版 社

标准分数制度及其应用

扈涛　刘刚　编著

科学普及出版社
北　京

（京）新登字 026 号

科学普及出版社出版
北京海淀区白石桥路 32 号　邮政编码：100081
新华书店北京发行所发行
各地新华书店经售
开封新新印刷厂印刷
※
开本 787×1092 毫米 1/32　印张 1.25　字数 16.25 千字
1991 年 2 月第 1 版　　1991 年 2 月第 1 次印刷
ISBN7－110－03710－X G・1182
印数：1——10000 册　　定价：2.70 元

序

党的十一届三中全会以来，我国以经济建设为中心，在各条战线进行了一系列的改革。普通高校招生改革事关人才培养的大局，起步早，成效显著，受到社会各界瞩目。1985年国家教委首先在普通高等学校招生全国统一考试中推行了标准化考试，它使高校选拔优秀新生更加准确，更加公平，也带动了考试手段的现代化，使考试信度不断提高。高考标准化指考试全过程的标准化。随着命题，考试实施和阅卷评分的逐步标准化，分数解释与分数合成的标准化必然地提到了议事日程。分数解释与合成的标准化是高考标准化的重要内容和主要指标。以往人们习惯用以高考卷面得到的原始分数或者以简单累加的总分去衡量考生的优劣，以当年的原始分与往年比较来评价考生水平的变化，这实际上不科学的，我们河南省从1994年高考开始实行标准分数制度，向高考标准化第二阶段目标迈进。这是一件大事，它将大大推进高考标准化进程，进一步提高高考质量。

为了使人们全面系统地了解标准分数制度，我们组织编写了《标准分数制度及其应用》一书。作者在重视其 系统性、科学性的同时，努力做到深入浅出，通俗易懂。该书将使广大招生考试工作者、考生、高校、中学和社会各界较全面地了解什么是标准化制度，标准分制度的基本原理，以及如何建立和正确使用标准分进行招生录取和考试评价工作等等。该书的出版将有助于在高考中推行标准分制度，促进教育考试的改

革。

标准分制度是考试分数制度的一项重大改革。它将改变教育界和社会各方面对考试分数的传统观念，它的优越性将会迅速被人们所认识。希望广大教育工作者尽快了解它，正确使用它，以促进我们整个教育考试的改革和发展。

侯福禄

1993年12月

目　录

第一部分　基础知识

无论是标准化考试还是其它类型考试都是以考试分数来反映考生知识和能力状况的信息。如何对考试分数进行有效地处理，科学地分析，从而得到有价值的结论是教育工作者的基本任务，国家教委考试中心决定从 1993 年开始在高考中逐步推行标准分数制度，这是高考改革的重大举措，是一种新生事物。为了帮助读者更好地认识和理解标准分数制度，正确地使用这些分数，首先对于了解和使用标准分数制度所应具备的教育统计学基础知识作简要的介绍。

一、次数分布

每次考试后，我们常常需要对考生分数有一个直观的整体了解。统常的做法是把各分数段的人数统计出来。例如某校数学考试后，把全部高三学生的考试情况以表格形式列出：

表 1—1　　某校高三学生数学成绩次数分布表

分数段	0—10	11—20	21—30	31—40	41—50	51—60	61—70	71—80	81—90	91—100
人数	0	1	3	7	13	17	24	18	13	4

这种反映数据散布在各个分数段的情况叫次数分布。用表格形式表示的次数分布叫次数分布表，根据次数分布表画出的图叫次数分布图。我们从次数分布表或图上可以看出一组分数的分布特点以及集中情况和彼此差异的情况。不同类型的数据有不同的次数分布图，其中有一种最常见的次数分布图，如图 1—1，统计学中把它叫做正态分布图。正态分布的

概率密度函数为：

$$P(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

式中：σ 为标准差；

μ 为平均数；

x 为随机变量的某一值。

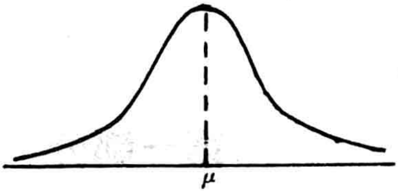


图 1—1　正态分布图

在自然界，在人类社会中，教育中大量现象的概率分布都按正态形式分布。实践证明，人的能力的高低，学生成绩的好坏，人们的社会态度等符合正态分布，也就是说，考试分数的分布从整体上看符合正态分布。因此，它是关于考试的教育测量学研究的前提之一，正态分布从其分布曲线的形状看，呈钟形线，中间高两头低。它是以过平均数点的垂线为对称轴，左右对称。正态分布在平均数及其附近的数据次数最多，远离平均数的数据的次数最少，也就是说，正态曲线的中央点（即平均数点）最高，然后逐渐向两侧下降，曲线的形状是先向内弯，然后向外弯，拐点位于正负 1 个标准差处，曲线两端向靠近基线（横轴）无限延伸，但终不能与基线相交。

正态曲线下的面积为 1（即把全体数据个数看作整体 1）。如果我们想求正态分布中任一数据 x_K 以下的数据个数占全部数据个数的概率如图 1—2，只需要根据正态分布的主要参数平均数和标准差直接查正态分布表即可得某数据以下的面积（即概率）。正态分布表是统计学家根据密度函数的积分而编制的，使用起来十分方便。

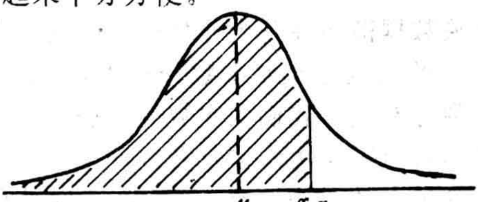


图 1—2　正态分布 X_K 以下的面积

在统计学中我们把一组数据出现次数最多的数据值叫做众数。如一组数据：1、3、4、5. 5、5、7、9、10，其众数是 5。显然，在正态分布中平均数等于众数。然而在实际中次数分布不全是正态分布，那么这时次数分布中平均数和众数不相等，使次数分布曲线发生偏斜。我们把平均数大于众数的连续次数分布叫做正偏态，而把平均数小于众数的分布叫做负偏态。其形状如图 1—3。

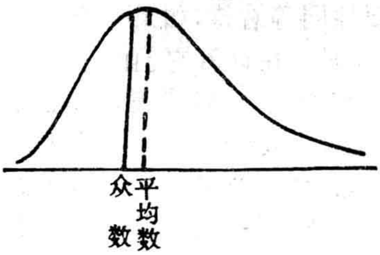
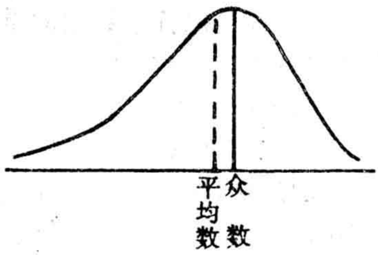


图 1—3　正偏态



负偏态

就整体而言，人的能力，学生的成绩符合正态分布，由于每次考试的难度不同，则对于一次考试，或一个地区的考试分数分布常常呈现为偏态分布。

二、平均数

考试分数经过初步整理后，从次数分布表上可以看出分布在各分数段的次数有多有少，但大部分数据趋向于某一点。例如表 1—1 某校高三数学成绩，40 分以下的是少数，41 分到 50 分的多一些，51 分到 60 分的较多，61 分到 70 分的最多，71 到 80 分的较多，81 分到 90 分的稍微多一些，91 分到 100 分的是少数。这种向某点集中的趋向称为集中趋势，表示集中趋势的量数则叫做集中量数。

集中量数是一组数据的代表值，它有两种功用。第一，可以用来描述和代表研究对象的一般水平；第二；用它与同质的另一研究对象作比较。例如，就一个省来说，它是全省某科高

考分数的代表，可以用它来代表这个省某科高考的程度，并可用它和其他省的这一科高考成绩作比较，集中量数包括有：算术平均数，中数，众数。在实际应用最广泛的是算术平均数。算术平均数简称平均数，计算公式很易理解，就是把所有的数据相加除以数据的个数。即：

$$\overline{X}=\Sigma X/N \quad (1\text{—}1)$$

（总体平均数用 μ 表示）

平均数是把各个数据的重要性同等看待，如果各个数据的重要性不同，分量大小不一样，那么在计算它们的平均数时，就应考虑到各个数据的分量、重要性，应以代表其重要性大小的数量作权数，来区别各个数据的重要性，使各个数据对平均数的影响与重要性相符合。这种把各数据的重要性考虑在内的平均数叫做加权平均数。它是各个数据分别乘以权数的和除以各个权数之和，即 $\overline{X}_W=\Sigma WX/\Sigma W$（$W$ 为权数，ΣWX 为加权总分，ΣW 为权数之和）。例如，计算学生各科学习的平均成绩，就需用加权平均数来计算，从而进行学生与学生之间各科平均成绩的优劣比较。"加权"的概念非常重要，在统计一个学校或地区统考成绩时经常遇到，不同学科成绩合成总分时也要用到。权数的确定，应尽可能依据客观的标准。

在高考中我们习惯使用各科原始总分来比较考生的优劣，也是应用加权的概念，只不过使用的是加权总分（即 ΣWX），而不用加权平均数。例如理科总分，是把语文、数学两科定为 120 分，生物定为 70 分，其余四科分别为 100 分，把七科考试分数相加得到理科总分。实质上是把语文、数学的权数定为 1.2，生物的权数定为 0.7，其余四科权数分别为 1，七科分别均以 100 记分。

即：$X_{理科}=1.2X_{语文}+1.2X_{数学}+1X_{政治}+1X_{物理}+1X_{化学}+1X_{外语}+0.7X_{生物}$

在考试结束后，我们常常用平均数来代表单科成绩的水平，用加权总分或加权平均数代表各科的综合水平。以便在班级之间，学校之间或不同地区之间的比较和考生之间的综合水平的比较。

三、标准差和方差

对于一组数据，仅用集中量数描述其集中趋势还不够，还要同时了解它的离散程度（即数据之间彼此差异的程度）才能反映这组数据的全貌。教师们有这样的经验：如果一个班级中两个组学生数学成绩是：

甲组：60、65、70、75、80、85、90

乙组：50、60、70、75、80、90、100

这两组的数学成绩虽然平均数都是 75 分，但甲组的分数比较集中，数据间的差异较小，75 分的代表性比较大，而乙组的分数比较分散，数据间的差异较大，75 分的代表性比较小。即两组分数各自围绕平均数变化的程度不同。这说明两组学习情况有差别。目前多数学校在评估全校或全年级考试成绩时，已意识到只用平均数还不够，还应有反映分数之间变异情况的量数。

我们把反映一组数据彼此离散程度的统计量叫做差异量数。一般说，描述一组数据的全貌应有两类指标，一是集中量数；二是相应的差异量数。集中量数在量尺上是一个点，表示各量数所在的位置，差异量数在量尺上是一段距离，表示一个量数与另一量数或中心点之间相差的距离。显然，一组数据的差异量数愈大，集中量数的代表性愈小；差异量数愈小，则集中量数的代表性愈大。

在实际工作中，既会遇到集中量数相同，差异量数不同的情况，也会遇到差异量数相同，而集中量数不同或集中量数与差异量数均不相同的情况，这些情况的次数分布曲线，如图

1—4

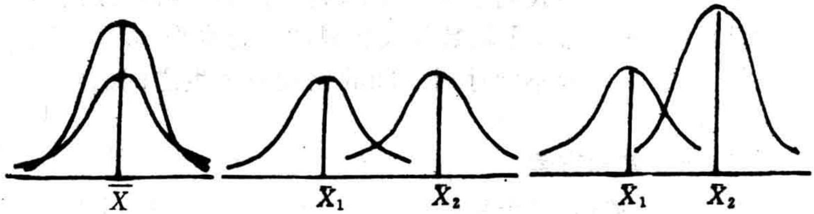


图 1—4　　几种次数分布

常用的差异量数有：全距、平均差、四分差、标准差等，其中标准差是应用非常广泛的重要的差异量数，它是与平均数相对应的差异量数，通常以 S 或 SD 表示标准差。标准差的平方 S^2，简称方差，也是非常重要的差异量数。它们的计算公式为：

$$S=\sqrt{\frac{\Sigma(X-\overline{X})^2}{N}} \tag{1—2}$$

$$S^2=\frac{\Sigma(X-\overline{X})^2}{N} \tag{1—3}$$

例：10 名学生的外语考试分数为 87、93、86、91、78、82、68、75、76、84，求这 10 个分数的平均数、标准差和方差。

$$\overline{X}=\frac{87+93+\cdots\cdots+84}{10}=82$$

$$S=\sqrt{\frac{(87-82)^2+(93-82)^2+\cdots\cdots+(84-82)^2}{10}}$$

$$=7.376$$

$$S^2=7.376^2=54.4$$

在反映一组数据差异程度或者比较两组测量单位相同的数据差异程度，标准差或方差是最好的统计量。但是，如果测量单位不同，或者单位相同而平均数相差很大时，则往往不能使用标准差或方差比较其差异程度。譬如，某年级身高平均 170 厘米，标准差 14 厘米，体重平均 62 公斤，标准差 8.1 公

斤。这两个标准差测量单位不同，不能直接比较，绝对不能说身高的离散程度大于体重的离散程度。又如：某年级数学成绩$\overline{X_1}=85.3$分，$S_1=7.9$分，物理成绩$\overline{X_2}=70.1$分，$S_2=7$分。这时，$S_1>S_2$，但不能说数学成绩的离散程度大于物理成绩的离散程度。在这种情况下要比较两组数据的离散程度，则需用相对标准差，也称变异系数，其公式为：

$$CV=\frac{S}{\overline{X}}\times 100 \qquad (1—4)$$

上例中，$CV_{数学}=\frac{7.9}{85.3}\times 100=9.26$

$$CV_{物理}=\frac{7}{70.1}\times 100=9.99$$

因此，数学成绩比物理成绩的离散程度小。

在使用标准差时，常常还需要从几个已知的、具有相同测量单位的标准差合成总的标准差。其公式为：

$$S_T=\sqrt{\frac{(N_1S_1^2+N_1S_2^2+\cdots\cdots+N_KS_K^2)+(N_1d_1^2+N_2d_2^2+\cdots\cdots+N_Kd_K^2)}{N_T}}$$

$$=\sqrt{\frac{1}{N_T}(\Sigma N_iS_i^2+\Sigma N_id_i^2)}=\sqrt{\frac{1}{N_T}\Sigma N\backslash -i(S_i^2+d_i^2)} \qquad (1-5)$$

式中：S_T—总标准差；N_T—总人数；$(N_T=\sum N_i)$；S_i—第i组的标准差；d_i—第i组平均数与总平均数之差。

例：某年级三个班，一次考试后，甲班：$\overline{X_1}=80$分，$S_1=10$分，$N_1=30$人；乙班：$\overline{X_2}=75$分，$S_2=8$分；$N_2=40$人；丙班：$\overline{X_3}=70$分，$S_3=9$分，$N_3=45$人，求全年级的总标准差。

解：先求总平均数

$$\overline{X_T}=\frac{N_1\overline{X_1}+N_2\overline{X_2}+N_3\overline{X_3}}{N_1+N_2+N_3}$$

$$=\frac{30\times 80+40\times 75+45\times 70}{30+40+45}=\frac{8550}{115}=74.35$$

再求出d：

$d_1=\overline{X_1}-\overline{X_T}=80-74.35=5.65$

$d_2=\overline{X_2}-\overline{X_T}=75-74.35=0.65$

$d_3=\overline{X_3}-\overline{X_T}=70-74.35=-4.35$

代入公式(1—5)：

$$S_T=\sqrt{\frac{30\times(10^2+5.56^2)+40\times(8^2+0.65^2)+45\times[9^2+(-4.35)^2]}{30+40+45}}$$

$=9.74$(分)

故全年的总标准差为 9.74 分。

以上我们介绍了怎样从集中、离散两方面对一组数据进行全面描述。在教育工作中，人们不仅关心团体数据的情况，还关心个人分数在团体分数中所处的位置。如在一个班学生中，一个学生的成绩为 80 分，家长要追问班里有几个 80 分以上的？全班平均分是多少？也就是要问他在班上的成绩位置如何？比他好的占百分之几？名列第几？因为水涨船高，单凭一个分数 80 分是看不出学生水平的。在这种情况下我们则需用相对位置量来刻化。相对位置量数是用以表明一个量数在其总体量数中所处的位置。在考试中，如果有两个学生同是 80 分，分别在两个班，若两个班的成绩分布不同，或人数多少不同，则两个学生在其班上的位置也不同。同样，在高考或其他升学考试中，如果有一考生在两科(比如语文、数学)考试中均是 80 分，若两科的成绩分布不同，其分数不等值，则两个分数在其学科上的位置也不相同。为了相比较，需要用一个相同的量尺，从而量出相对的位置，即相对位置量数。它主要有百分等级和标准分数。现分别介绍。

四、百分等级

百分等级是将一个数列中的量数，按大小顺序排列，将其次数分为 100 等份，从而表明某一量数在此百分等级中所处的位置。如果某人的百分等级为 50，则他正好是团体的平均水平，如果百分等级为 89，则说明他的成绩在团体中高于

89%的人,低于 11%的人。

百分等级的计算可分为两种情况:根据未分组的原始分数计算和根据已分组的次数分布计算。

1、根据原始分数求百分等级

计算步骤为:

(1)先把考试而得的原始分数,按其大小顺序排列,并记上各原始分数的次数;

(2)把各原始分数的次数自上而下累加;

(3)根据以下公式,确定百分等级。

$$P_R = 100 - \frac{100R - 50}{N} \tag{1-6}$$

式中:P_R—百分等级;N—总次数;R—某学生的原始分数在团体中处于由大到小排列的名次。

例:某班的数学考试成绩如下表,求该班 80 分的学生百分等级。

表 1—2

分数 x	次数 f	累加次数 cf	利用公式求百分等级
95	4	4	$N=48$
90	5	9	80 分的名次为大于该分数向下累加次
85	10	19	数 19 加 1,即 $R=19+1=20$
80	8	27	$P_R = 100 - \frac{100\times 20-50}{48} = 59.38$
75	6	33	
70	5	38	
65	4	42	
60	3	45	
55	2	47	
50	1	48	
合计	48		

2、根据次数分布求百分等级

计算百分等级的公式：

$$P_R=\frac{100}{N}\left[\frac{f(X-L_b)}{i}+F_b\right] \qquad (1-7)$$

式中：X—给定的原始分数；

f—该分数所在组的次数；

i—组距；

L—该分数所在组的真实下限(即下限减去 0.5)；

F_b—小于 L 的向上累加次数。

例：求表 1—3 的 112 个原始分数中的 81 分的百分等级，并解释其意义。

表 1—3 某高校高三年级外语成绩次数分布表

组别	次数 f	累加次数 cf
90—94	3	112
85—89	11	109
80—84	16	98
75—79	21	82
70—74	24	61
65—69	15	37
60—64	12	22
55—59	4	10
50—54	5	6
45—49	1	1

解：由表 1—3 可知，$N=112$，$i=5$，81 分在 80—84 这一组内，由此得到：

$f=16$，$L_b=80-0.5=79.5$，$F_b=82$

代入公式：

$$P_R=\frac{100}{112}\left[\frac{16(81-79.5)}{5}+82\right]=77.5$$

故 81 分的百分等级是 77.5，即得 81 分的学生其成绩超过全年级 77.5%的学生。

百分等级在教育工作中的用途是很多的，它可以用来明确考生的相对位置，比较考试成绩的优劣，也可以比较两个群体的成绩.判断成绩的优劣，还在建立标准分制度的分数转换中起着重要作用。

例如：根据某地区高考成绩的次数分布表，求得某生语文分数(86 分)的百分等级为 $p_{86}=91.3$，数学分数(82 分)的百分等级为 $P_{82}=94.7$，即分别超过该地区全体考生的 91.3%和 94.7%，成绩均属优良，都处于较高地位，而且可看出，数学的原始分数虽比语文低，但超前百分数比语文大，所以，该生的数学成绩在全体考生中的地位高于语文成绩在全体考生中的地位。

又如，93 年高考中，某省的录取线为 480 分，根据甲、乙两地区考生总分的次数分布；求得甲地区 480 分百分等级 $P_{480}=74.3$，乙地区 480 分的百分等级 $P_{480}=82.4$。由此得知，甲、乙两地区的上线率分别为 25.7%和 17.6%，所以，甲地区的高考成绩优于乙地区。

五、标准分数

标准分数是反映个人在团体中相对位置的最好统计量。它是将原始分数与其平均数之差除以标准差所得的商数。标准分数又称 Z 分数，通常用 Z 表示。其计算出公式为：

$$Z=\frac{X-\overline{X}}{S} \tag{1—8}$$

从公式可以直观地看出，标准分数是用标准差为单位来衡量某一原始分数与其平均数之差的。它表明此原始分数离平均数的远近，即刻化了原始分数在平均数以上或以下几个标准差的位置上，从而明确此原始分数的相对地位。例如原始