

最新教学艺术全书

教科研艺术 (六)

郭雅 主编

吉林摄影出版社

图书在版编目(CIP)数据

最新教学艺术全书/郭雅主编. —长春: 吉林摄影出版社, 2004

ISBN 7-80606-720-6

I. 最… II. 郭… III. 执法工作—中国—汇编
IV. D922.851

中国版本图书馆 CIP 数据核字(2004)第 053253 号

出版发行: 吉林摄影出版社
(长春市人民大街 124 号 130021)

责任编辑: 李乡壮

经销: 全国各地新华书店

印刷: 北京施园印刷厂

版次: 2004 年 3 月第 1 版

书号: ISBN 7-80606-720-5/ D · 201

定价: 399.00 元

目 录

◎教育科研方法系列之八：教育统计 与测量方法.....	1
◎教育科研方法系列之九：科学研究的 思维过程与教育科学研究的步骤.....	2 8
◎教育科研方法系列之十一：教育科研 计划的制定.....	7 8
◎教育科研方法系列之十二：教育科研的 实施和成果表述.....	9 2
◎教育科研论文撰写系列之一：教育 科研论文的基本概念.....	1 1 8
◎教育科研论文撰写系列之二：教育 科研论文的撰写.....	1 5 8
◎与教师谈参与教育科研.....	1 9 6

◎教育科研方法系列之八 教育统计与测量方法

统计和测量都是运用数学方法的研究手段，二者紧密相关，相辅相成。测量是统计的前提，有了测量提供的数据，统计才得以进行；而统计又是测量的基础，从事测量的人员若没有统计的基础知识，也就无法整理、分析测量的结果。因此，将统计与测量合在一起阐述。

目前，教育科研的精确化、科学化、规范化的要求愈来愈高，加上电子计算机技术的日益普及，为定性与定量相结合的分析研究提供了可能。这样，测量与统计便更成为教育科研中不可缺少的手段与环节。因此，必须熟练地掌握统计与测量的运用技术。

一．教育统计在教育科研中的运用

（一）教育统计的基本内容

教育统计，就是应用数理统计学的一般原理和方法，对教育科研和教育实践中所获得的数据，进行整理、计算、分析与解释。其主要内容包括以下两个方面：

1. 描述统计

描述统计，就是将数据资料加以整理、简缩，使之有序化，制作成次数分布表或分布图；或根据数据

的分布特征，如集中趋势、离中趋势、相关强度等，计算出平均数、标准差、相关系数等概括性的统计量数，以便人们从杂乱的原始数据中获得有意义的信息，进行比较，作出结论。

2. 推断统计

推断统计，是从样本统计量来推断它来自总体的特性，并标明可能发生的误差的统计方法。在现实的教育研究中，限于人力物力，总是从总体中抽取有代表性的样本，然后从样本统计量对总体的特征进行推断，即进行相应的“显著性检验”等统计分析工作。在推断统计的基础上，研究者将对所研究的问题做出自己的解释、预测或估价。

(二) 数据的特征量及其计算

除了编制原始数据的次数分布表或绘制其次数分布图以了解数据分布的概貌之外，作为描述统计的主要方法，仍是用数据的特征量来表现其分布状况。

1. 集中量数

描述集中趋势的统计量，叫做“集中量数”，简称“集中量”。常用的集中量数有三种：算术平均数，中（位）数和众数。这里就某实验组和对照组某次考试的原始数据为例作些说明。

(1) 算术平均数，简称平均数、均数或均值。

其符号为“ $\bar{x}$ ”(读作 $\bar{x}$ 杠),它起着衡量一定数据的集中趋势和大致水平的作用,是最常用的集中量,其计算公式是

从算术平均数可以看出,实验组与对照组的平均水平是否一样。

(2) 中数(符号为 M_n),是依一定顺序(如由大到小)排列的一组数据居中间位置的一个点的数值,所以又叫中位数。如果数据个数 N 为奇数时,中位数的位置在 $(N+1)/2$ 处,若 N 为偶数,就以居中的两个数据的平均数作中位数。

(3) 众数(符号为“ M_0 ”),指一组数据中出现次数最多的那个数值。在众数甚至没有的情况下,一般可看众数段,即哪个分数段的次数多,就以该段中点值作众数。

以上三个集中量中,平均数是无偏的客观量数,又最便于代数运算法则处理,从样本数值推断总体集中量时,平均数比中数、众数可靠,其缺点是易受两极端数值的影响。

2. 差异量数

差异量数是描述次数分布中“离中趋势”这一特征的统计量,简称“差异量”。一组数据,若离中趋势小,则集中量的代表性就大;反之,若离中趋势大,

则集中量的代表性就小。但是，仅考虑集中量数是不够的。要了解两组学生成绩分布的全貌，还必须研究两个组的差异量数。最常用的差异量有全距、平均差和标准差。

(1) 全距(符号为“R”)，指一组数据中由最大量数到最小量数的距离。R小说明离散程度小，比较整齐。

(2) 平均差，指一组数据内的每个数与均数差的绝对值的算术平均数，通常用 A D 表示。平均差的计算公式为：

(3) 标准差，指一组数据中每一个数值与它们的平均数之差的平方的算术平均数的平方根，其符号为“S”(样本标准差)、“ σ ”(总体标准差)。其计算公式为：

S 越大表明离散程度越大，数据不均匀，集中量的代表性小。

3. 相对位置数

平均值与标准差用来考察与分析同质的统计资料是有价值的，但对于不同质的考试，如不同学科、或同一学科不同考试意义就不大。这样就要计算相对位置量数。相对位置量数有百分等级与标准分数两种。这里就常用标准分数作些介绍。

标准分数，又称 Z 分数，它是一种以平均数为参照点，以标准差为单位的，表示一个分数在团体分数中所处位置的量数，其计算方法为：由原始分数与平均分数的离差除以标准差所得的量数，其符号为“Z”，计算公式是：

例：有某生三次数学考试的成绩分别为 70、57、45，三次考试的班平均为 70、55、42，标准差分别为 8、4、5。如何看待该生的三次考试成绩的地位？如果仅从原始分数看，肯定认为第一次最好，其实不然，要计算出各次的标准分数，才能说明问题。

根据公式得出：

$$Z_1 = (70 - 70) / 8 = 0 \quad Z_2 = (57 - 55) / 4 = 0.5 \quad Z_3 = (45 - 42) / 5 = 0.6$$

这说明，原始分数为 70，其位置正在平均线上，而原始分数为 57 的，其位置在平均线上 0.5 处，而原始分数为 45 的，其位置在平均线上 0.6 处。

4. 相关系数

在教育研究中，常涉及到两个事物(变量)的相互关系问题，例如，学习成绩与非智力因素的关系，数学成绩与物理成绩的关系，男女生学习成绩的关系，等等。其关系表现为以下三种变化；第一，正相关：一个变量增加或减少时，另一个变量也相应增加或减

少；第二，负相关：一个变量增加或减少时，另一个变量却减少或增加；第三，无相关：说明两个变量是独立的，即由一个变量值，无法预测另一个变量值。统计学中，就用“相关系数”来从数量上描述两个变量之间的相关程度，用符号“ r ”来表示。

相关系数取值范围限于： $-1 \leq r \leq +1$

相关系数表示的意义

相关系数 (r) $0.000.00 - \pm 0.3 \pm 0.30 - \pm 0.50$

相关程度无相关微正负相关实正负相关

相关系数 (r) $\pm 0.50 - \pm 0.80 \pm 0.80 - \pm 1.00$

相关程度显著正负相关高度正负相关

例：某语文实验班随机抽 10 名学生，参加市里的语文数学竞赛，其成绩如下表，求这两门成绩的相关系数。

10 名学生语文与数学相关系数计算表

语文 X 数学 $Y_x = X - \bar{X}$ $2 y = Y - \bar{Y}$ $2xy$

58684.722.0910803.428.46

416712.3151.290.800.649.84

49534.318.4913.20174.2456.76

57553.713.6911.20125.4441.44

62808.775.6913.8190.44120.06

46657.353.291.210448.76

48585.328.098.267.2443.46

59675.732.490.80.644.56

53770.30.0910.8116.643.24

60626.744.894.217.6428.14

Σ 533622 440.10 697.69159.4

相关系数的计算公式是由英国统计学家皮尔逊提出的“积差相关”公式：公式中， $r=X$ 与 Y 两数列之间的相关系数；

$x=X-\bar{x}$ ，即 X 数列中各量数与其平均数之差；

$y=Y-\bar{y}$ ，即 Y 数列中各量数与其平均数之差；

$S_x=X$ 数列的标准差；

$S_y=Y$ 数列的标准差；

Σxy =各对离差积的总和；

N =成对量数的次数，即总对数。

计算步骤：

1. 先计算出 ΣX 、 ΣY 、 X^2 、 Σx^2 、 Y^2 、 Σy^2 、 Σxy 分别填入表。

2. 求 X 和 Y 的标准差

3. 求语文数学两门成绩的相关系数：

4. $r=0.29$ 在0与+0.3之间，属于微正相关，说明这10名学生的语文成绩与数学成绩就样本本身而言，是有一定相关的。

如果用的是等级评分法，就不宜用积差相关公式，而要用英国统计学家斯皮尔曼提出的等级相关公式：

式中， r_R —斯皮尔曼等级相关系数； D —各对 X 与 Y 的差数， N —总对数。

（三）统计检验及其应用

统计检验，就是对样本的特征量能否反映总体特征的问题，或两种不同样本数量标志的参数的差异性问题，作出定量分析与推断。

1. 统计检验的原理与方法

（1）统计检验的基本原理

统计检验的理论依据是概率论中的“小概率事件实际上的不可能性”原理。所谓“小概率事件”，即假定某个事件在实验中出现的概率很小，则在一次实验中，该事件实际上是不会出现的，例如，把小于 0.05 或 0.01 的概率，视为“小概率”。

（2）统计检验的一般方法

一般来说，统计检验先对总体的分布规律作出某种假说，然后，根据样本提供的信息，对假说作出肯定或否定的决策。具体步骤为：

①提出假设。如“假设两个总体平均数没有差别”，其数学符号为：“ $H_0: \mu_1 = \mu_2$ ”，这种对总体所

作的“无差别”的假设，称为“零假设”或称虚无假设，用符号“ H_0 ”表示。与此同时实际上存在第二种假设，“两个总体平均数有差别”，其符号为：“ $H_0: \mu_1 \neq \mu_2$ ”，称为备择假设。显然，“零假设”与“备择假设”是两个对立的假设，肯定此，必否定彼。

②根据不同条件和样本提供的信息即数据，从零假设出发，代入相应的公式，计算出零假设的概率。

③作出统计决断，根据“小概率事件实际上不可能性”原理，研究 H_0 成立的概率。如果 H_0 的概率 $P < 0.05$ ，表示零假设不是一个小概率事件，则 H_0 成立，便否定被择假设 H_1 从而定“ $\mu_1 = \mu_2$ ”。如果 H_0 的概率 $p \leq 0.05$ ，表明是个小概率事件 H_0 不成立，就肯定备择假设 H_1 的成立，从而确定“ $\mu_1 \neq \mu_2$ ”。

(2) 统计检验的具体操作

① Z 检验

Z 检验是一般用于大样本(即样本容量大于 30) 平均值差异性检验的方法。它是用标准正态分布的理论来推断差异发生的概率，从而比较两个平均数的差异是否显著。

例 1987 年上海市初中三年级语文教学调查中，对男女生语文测试成绩作如下统计，试检验男女生语

文成绩是否存在显著差异，

性别人数总分阅读写作

平均分标准差平均分标准差平均分标准差

男 167 113.74 20.95 71.64 15.11 42.12 9.73

女 159 118.52 19.09 74.87 14.01 44.43 8.83

差值 $|Z|$ 2.15 2.00 2.27

抽取的两个样本均大于 30，属两个独立大样本平均数差异的显著性检验，用 Z 检验。

检验步骤：

①提出零假设 $z: H_0: \mu_1 = \mu_2$ 即假定男女写作、阅读及读写总分均无显著差异，现在的差异是抽样误差所致。

②计算统计量，代人 Z 值公式

③计算出的 Z 值与下表进行对照，作出判断：

ZP 差异显著性

$<1.96 > 0.05$ 差异不显著

$\geq 1.96 \leq 0.05$ 差异显著

$\geq 2.58 \leq 0.01$ 差异极显著

因为 $|Z_{写}| = 2.27$ ，显然， $|Z_{写}| > 1.96$ ，表明概率 $P \leq 0.05$ ，男女生写作成绩差异显著。

因为 $|Z_{读}| = 2.00$ ，显然， $|Z_{读}| > 1.96$ ，表明概率 $P \leq 0.05$ ，男女生阅读成绩差异显著。

因为 $|Z_{\text{总}}|=2.15$ ，显然， $|Z_{\text{总}}|>1.96$ ，表明概率 $P\leq 0.05$ ，男女生语文成绩差异显著。

④结论：当 $P\leq 0.05$ 时，拒斥 H_0 ，肯定 H_1 ，1987年调查说明上海市初三语文成绩男女生存在显著差异，女生高于男生。

②t 检验

t 检验是用于小样本(样本容量小于30)时的平均值差异程度检验方法。它是用t分布理论来推断差异发生的概率，从而比较两个平均数的差异是否显著。

例 某校初一年级抽出一组20人，对数学自学辅导教材进行试验，期末全年级测试平均成绩为70分，而这20人的平均分为 $\bar{x}=77.7$ ，标准差为15，试检验实验效果。

本例随机抽样样本容量为20人，属小样本，因此适用t检验。所谓检验实验效果，就是以样本(20人)的平均数 $\bar{x}$ 与某已知总体平均数 μ_0 之间的差异程度，来检验样本所取自(所代表)的总体的平均数 μ 与 μ_0 ，是否有差异。

检验步骤：

①提出零假设： $H_0: \mu=\mu_0$ ，即假定样本所代表的总体平均数与已知平均数无显著差异，如有差异仅是抽样误差所致。本题 $\mu_0=70$ 分

②计算检验统计量 t 值。用如下公式

式中， $\bar{x}$ =样本平均数 77.7； μ_0 =已知总体平均数 70；

s =样本标准差 15； n =样本容量 20，代入公式得

③作出判断。与正态分布曲线不同， t

分布的曲线形式随自由度大小而不同。“自由度”记作“ df ”。作总体平均数的假设检验时，统计量 t 的自由度 $df = n - 1$ 。据此，本题的 $df = 20 - 1 = 19$ 。查 t 值表，得出理论 t 值为：

$$t(19)_{0.05} = 2.093$$

再与计算所得 t 值比较可得： $t > t(19)_{0.05}$

依据《 t 值与差异显著性关系》表，推断 H_0 发生的概率，作出结论。

t 值与差异显著性关系

t P 差异显著性

$< t(df)_{0.05} > 0.05$ 差异不显著

$\geq t(df)_{0.05} \leq 0.05$ 差异显著

$\geq t(df)_{0.01} \geq 0.01$ 差异十分显著

因为 $t = 2.31 > t(df)_{0.05}$ ，从上表可知，概率 $P \leq 0.05$ 时， μ 和 μ_0 之间的差异显著。因此可结论为：拒斥 H_0 ： $\mu = \mu_0$ ，而肯定 H_1 ： $\mu \neq \mu_0$ ，又因 $\mu > \mu_0$ ，故结论表明新教材实验有成效。

如果是依两组样本平均值 1 和 2 的差异程度，检验它的代表的两总体平均数， μ_1 和 μ_2 是否有差异，其检验统计量 t 的计算公式为：

如果是按同一组样本不同情况的测试所得的平均值 1 和 2

来检验 μ_1 和 μ_2 的差异程度，其计算公式为：

式中， D 为两次测试中每对分数之差即 $D = X_2 - X_1$ 。

③ χ^2 检验

Z 检验与 t 检验，通常用于计量资料的分析，而在教育研究中还常有计数资料，如按品质分类，然后按类评等计数，如优良中差，甲乙丙丁，或同意、反对、弃权等。这种计算资料检验就要利用 χ^2 检验的方法。 χ 是希腊字母，读 chi

，通常把 χ^2 读作“卡方”。 χ^2 检验是通过对所得的计数资料与依据某种假设而确定的理论次数二者之间的差异来进行检验的。 χ^2 值是检验实测次数与理论次数之间差异程度的指标。两者相差越大， χ^2 值就越大；两者越接近，则 χ^2 值就越小：如果两者完全相同，那么 χ^2 值就等于零， χ^2 值永远是非负值。

例 某校在本校高一重点班与非重点班分别抽取 100 名学生，对他们的英语口语能力作出检测、评价，

结果如下表，试检验两班差异程度是否显著。

(R)

行(L)列 评价结果合计

优良中差

重点 20 (13.6) 40 (40.9) 30 (31.8) 10 (13.6)

$N_a=100$

非重点 10 (16.4) 50 (49.1) 40 (38.2) 20 (16.4)

$N_b=120$

合计 $N_1=30 N_2=90 N_3=70 N_4=30 N=220$

检验步骤：

①提出零假设： H_0 ：该两班英语口语能力无差异

②确定自由度，根据列联表自由度公式： $df=(R-1) \times (L-1)$ ，求出自由度为： $df=(2-1) \times (4-1)$

③计算 R 行与 L 行的理论次数，计算公式为：

将计算出来的理论次数填入上表相应的实际次数旁边的括号内。

④根据公式， χ^2 值。

缩节《 χ^2 值表》

Df 0.050.010.005

13.8146.6357.88

25.9959.21010.60