

教育统计与测量

傅亚强



杭州师范学院教育学院

2013.7

引言

一、统计课的意义

我这门统计课就是一门可以直接转化为技能的课程。你们应该在今后的学习工作中有意识地应用这门课上所讲到的各种方法。

- ◇ 多元评价的需要；
- ◇ 发展性评价的需要；
- ◇ 教师评价的需要；
- ◇ 因果探寻的需要；

二、统计是干什么的？

1、统计有点难学！仅凭讲解或者练习还很难真正理解其中的原理，只有在实际应用当中才能做到这一点。

2、统计的功能：数据的分析。

(1) 潜能班，即分层教学：是否有效提高问题学生的成绩。

(2) **CAI** 对教学效果、学生的能力培养有没有促进作用？是否会导致信息超载？是否会限制师生的发散思维？

(3) 《基础教育课程改革纲要》中强调对学生的“发展性评价”：促进发展必须建立在对已有经历的总结上。

(4) 研究性学习：小发明、小作品、小游戏、创造性思维、自学能力、探究能力。

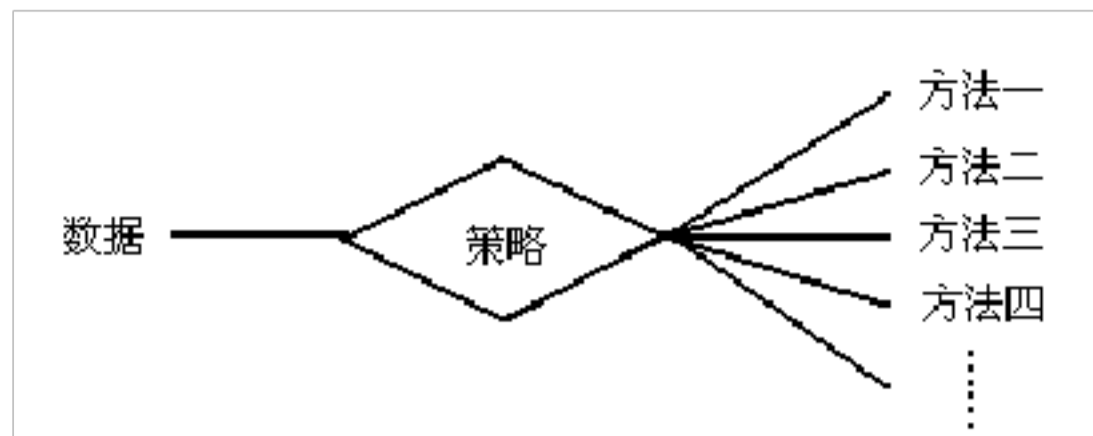
3、从以上例子中可以看出我们**搜集的数据具有一个共同的特点：随机性**，即同一个人观察同一对象时，不同时间观察结果不同；不同人同一时间观察同一对象时也会得到不同结果。

4、对随机现象的探索，涉及一个最基本的统计学思想，这一思想是这门课程应该学到的最重要内容之一：

总体中抽样→属性量化→统计整理→推论总体特征（差异、发展趋势）拿到数据如何选择正确的统计方法

依据上述思路，我们可以纷繁复杂的数据中找到规律。

5、统计学中我们会学到常用的统计方法，其它的只能靠大家自学。但了解一个一个的方法还不是主要的学习目的，重要的领会何时何地使用何种方法：



6、结论建立在可能性基础之上。

四、测量是干什么的？

1、从考试说起

如果一个大学毕业生应掌握的英语单词数是 4000 个，如果想知道他的词汇水平如何——
测量目的；

最常用的方法是出几道题考考他，正所谓“是骡子是马，拉出来溜一溜！”——以测验为
工具；

但是不可能 4000 个单词都考一遍，我们只能从中抽取一部分单词来考，如 100 个。——
行为样本；

接下来就是一个形式问题，出拼写题、听写题、选择题、阅读理解题都可以选取，——题
目形式选择。

2、教育测量的其它形式还有很多，如情景模拟、面试等。

3、测量在评价、选拔、培训中有着十分重要的作用。

五、统计与测量有什么关系？

(1) 教育统计可以比作基础性工具，许多学科都要用到它，教育测量也不例外，其中要用到
大量的统计学原理和方法；

(2) 统计所需要数据又需要各种测量手段来获得。

介绍《社会科学统计软件包 (SPSS)》。

第一章 绪论

第一节 统计学与教育统计学

一、什么是统计学：

对某一类对象的若干属性进行考察（数量化），进行（1）数据搜集；（2）数据整理；（3）数据分析，达到推断这一类对象属性的目的。

二、教育统计学

分析在教育实践、教育实验、教育调查中获得的数据

关键在“应用”二字上，以后我们会讲到许多方法、公式，记住这些东西是必要的但不是最重要的，首要的有两条：（1）掌握在什么条件下用什么方法和公式；（2）要学会对统计结果进行解释。这里的“解释”有两个含义：（1）相对大多数人的相对结果；（2）获得这一统计结果的具体情景，不能任意扩大结果应用范围。

三、教育统计的内容

1、实验设计的一般模式

例 1：分散学习方式与集中学习方式对学习效果的影响

- ◇ 总体中抽取样本
- ◇ 自变量
- ◇ 因变量
- ◇ 统计：比较差异
- ◇ 因果推论

2、描述统计

（1）数据分组（最考察一个人的研究智慧的，在研究中谁与谁比是要整体规划的）

（2）编统计表

（3）绘统计图 是一篇文章的窗口 要画的规范、清晰、明了，让人家一看就懂

（4）特征量计算

- ◇ 集中量—平均数
- ◇ 差异量—标准差
- ◇ 分布形态

◇ 相关系数

描述一组数所的全貌，目的在于将大量零散杂乱无序的数据进行整理、归纳、简化、概括，使事物的全貌明确清晰地表现出来。

3、推断统计（学习重点）

通过局部数据（样本）所提供的信息，推论总体的情况。

比如要分析一下本校学生的心理健康情况，总不可能让全校学生都来测一遍，所以要挑一小部分学生一样本，测量他们的心理健康水平，这样就有一个根据这一部分人的数据推测全校学生一总体的心理健康情况的问题。

包括总体参数特征值估计、假设检验等内容。

如：由样本平均数推断总体平均数，由样本标准差推断总体标准差。

	一中	二中
焦虑水平测验（50）	78	86

注意：推论的正确率。这两组数据是随机的，是样本。样本平均数据数不能说明两者之间的差异。不能说二中的比一中 焦虑。

● t 检验

● 方差分析

4、以上三方面的关系

(1) 实验设计是基础，描述统计和推论统计都要在良好的实验设计基础及其所提供的真实的数据之上；

(2) 描述统计是推论统计的基础；

(3) 实验设计必须以统计原理为依据，必须符合统计原理和考虑统计方法的可行性。

第二节 几个基本概念

一、随机变量

1、随机现象

在相同条件下，进行实验或观察，其结果可能不止一个，事先无法确定，这类现象称为随机现象。

先举例：抛硬币实验

(1) 一次试验有多种结果，其所有可能的结果是已知的；

(2) 试验之前不能预料哪一种结果会出现；

(3) 在相同的条件下可以重复试验。

2、随机现象的每一种结果叫基本随机事件

摸奖：一口袋中有 6 只球，其中 4 只白球，2 只红球，从袋中取先后取两次，每次取一只球（放回抽样），可能的基本事件有 36 种。

3、在统计学上把随机现象称为：变量、随机变量，用大写字母表示，如身高用 X 表示，语文考试成绩用 Y 表示。

◇ 变量：数量或质量上可变的事物属性

◇ 自变量：由实验者操纵的、对实验对象的反应产生影响的变量。

◇ 因变量：实验对象的某种反应。如父母亲的智力对孩子的影响，父母的智力就是自变量，孩子的影响就是因变量

◇ 例：大学生心理健康状况调查，复习时间安排对学习成绩的影响

4、随机事件称之为变量的取值，用 $X_1, X_2, X_3, \dots, X_n$ 表示。随机变量

5、在对随机变量的研究中，可以许多的规律：

(1) 随机现象的结果虽然不可预知，但其中有一个规律：即每个结果都以 50% 的等概率出现。

(2) 智力随年龄的变化规律，散点图。

(3) 语文成绩与英语成绩，散点图。

(4) 智力的概率分布

二、总体和样本

1、总体：具有某种特性的研究对象的个体总和。通过地域、年龄、职业、性别等维度定义一个总体。有限总体和无限总体。如《沿海开放地区小学生品德的性别差异》。本研究范围内，杭州什么区 以什么学校为例。在什么范围内而谈

2、个体：构成总体的每一个基本元素，不一定是人，如：产品质量抽查，题库质量分析。

3、样本：从总体中抽取的作为观察对象的一部分个体。

4、样本容量：样本中所包含的个体数目， N 表示。 $N \leq 30$ ，小样本； $N > 30$ ，大样本。如：问卷调查 300 个

三、统计量和参数

◇ 统计量相对样本而言 \bar{X}, σ_x, r ，是描述样本的各种特征的数量

◇ 参数相对总体而言 μ, σ, ρ ，是总体的各种特征的数量。一般情况下，我们通过分析统计量来对总体参数作出推论。

第二章 数据的初步整理

第一节 数据的来源、种类及其分类

统计是为了说明总体的属性、特征，或是两个或两个以上总体在某些属性上的差异。这些属性或其差异的说明最好只用简单的数字或简单的几句话。比如：升学率、及格率，配戴眼镜率、平均成绩、成绩分布情况、两个班或两个学校的成绩是否有显著差异、学生的成绩在不同时期的波动情况，上升、下降、起伏不定，而这些简单的结论是在对大量原始数据的整理基础上获得的。

因此，统计工作的第一步是搜集数据，即对客观事物进行量化的过程。这一过程要坚持实事求是，不能随意更改数据。在统计分析过程中亦是如此。

获得数据的能力也是综合素质的重要组成部分，在研究目标的指引下，平时要收集数据。教育实验

一、教育统计资料来源

1、什么是数据

数据是随机变量的观察值，它是用来描述对客观事物观察测量结果的数值。

举例：抛硬币实验的结果，考试成绩结果。

2、经常性资料：日常工作记录和统计报表，如师生档案，工资表，成绩表，综合素质考核表，业绩表。

3、专题性资料

(1) 教育调查：在没有预定因子、不施行控制的条件下，对现成的教育方面有关客事实所进行的观察和分析。

- ◇ 现情调查：对当前正在发生的事情或存在事物进行调查，如教师学历的调查。
- ◇ 回顾调查：调查过去发生过和事情或现在已不存在的事物，如青少年犯罪原因调查。
- ◇ 追踪调查：在今后的一段时间内调查对象的发展过程。实施教改后学生学习态度和学习行为的变化。

(2) 教育与心理测量：测验内容、程序、环境、评分记分、分数解释都有严格的规定。如智力测验、人格测验。

- ◇ 大学生焦虑发展趋势及其诱因变化，焦虑的诱因有学习、未来的工作、谈朋友。
- ◇ 小学生焦虑发展趋势及其诱因变化，焦虑的诱因有学习、上辅导班、父母吵架、老师批评、同伴争吵。

(3) 教育实验：在给实验对象施加预定控制因子的影响的条件下，对教育方面的事实进行观察和分析。

A、被试间设计

例 1

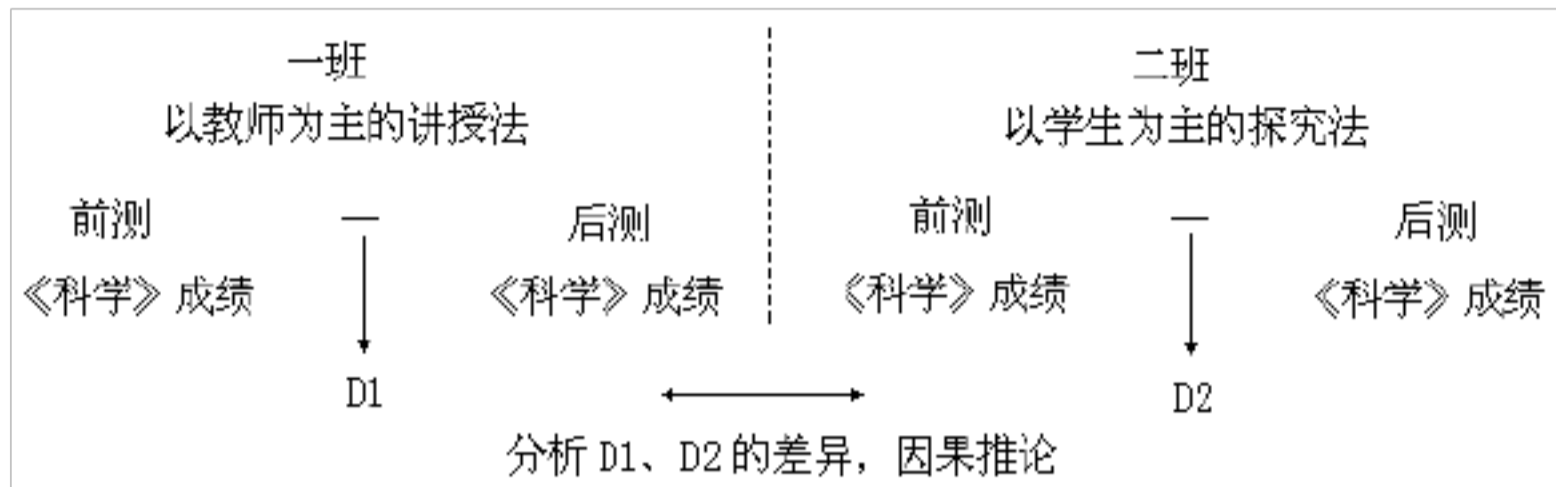


图 被试间设计

B、被试内设计

例 1

一个班学生在某学期施行了语文阅读教学改革，学期初与学期末分别进行了一次阅读理解测验，通过比较前后测成绩的变化，可以判断这次教学改革的效果。



例 2

处理	张三	李四	王五	赵六
纸张文本阅读	1	2	1	2
电子文本阅读	2	1	2	1

二、数据的种类——选择不同统计方法的依据

1、按来源来分

(1) **点计数数据**：如人数，班级数，男女数等。数人数的，班级几人，及格几人，优秀几人，

(2) **度量数据**：借助于一定的测量工具或一定的测量标准而获得的，如 17.2 厘米、58 千克、19℃ 测量工具如尺子，天平。

✧ **比率数据**：有绝对零点和等距单位，如身高、体重，反应时，感觉阈值等，差、倍数

有意义。（从零点定义的角度来理解）

✧ 等距数据：等距，但无绝对零点，如温度，各种能力测验分数等，倍数无意义，只有差有意义。

✧ 顺序数据：既不等距，也无绝对零点，如等级定，喜爱等级，品质等级等。如第一名、第二名 如考试成绩。顺序数据是不能进行任何的加减乘除的

(3) 称名数据：反映事物在组别、种类上的不同。不说明事物之间的数量差异，中按事物的性质不同，将被观测事物物加以划分。如：令男为 1、女为 0；职称：高级 2、中级 1、初级 0。

0、1 没有任何数量上意义，只是标记

2、按取值情况来分：

(1) 间断型随机变量数据——离散数据：单位不能再划分成更小的单位，一般取整数，如计数、顺级数据。电阻箱。

(2) 连续性随机变量数据——连续数据，单位可划分无限细小的单位，如等距数据，比率数据。滑动变阻器。

三、数据的统计分类（分组）

数据的统计分类，是指按照研究对象的本质特征，根据分析研究的目的、任务，以及统计分析时所用统计方法的可能性，将所获得的数据进行分组归类。如：将成绩分成男生成绩和女生成绩，将作文成绩分成优、良中、差；将健康状况分成好、中、差；按实验处理不同分组。

1、关键点：分类的标志，如：心理健康水平，考试成绩。

要求：明确、前后一致、能将全部数据包含在内、依据对象的本质属性、依据研究目的、依据统计方法

2、分类的标志可分为：

✧ 性质类别：按事物的不同性质进行分类，如性别，实验组数据和对照组数据。

✧ 数量类别：按数值大小来分组，如成绩分布。

例：《学生网络问题调查》

问题：

✧ 我认为“黑客”很能干，我很羡慕他们。

✧ 互联网比现实生活更具有真实性。

✧ 我常常为不能控制上网次数而懊恼。

✧ 上网时我会冒充其他人。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/07610400300010104>