

# 教育统计的世界

统计原理与SPSS应用

张 黎 著



π

新 华 出 版 社

# 教育统计的世界

统计原理与SPSS应用

张 黎 著

新 华 出 版 社

## 图书在版编目 (CIP) 数据

教育统计的世界：统计原理与 SPSS 应用 / 张黎著.

北京：新华出版社，2016.8

ISBN 978-7-5166-2777-8

I. ①教… II. ①张… III. ①教育统计—统计分析—软件包 IV. ①G40-051

中国版本图书馆 CIP 数据核字 (2016) 第 201884 号

## 教育统计的世界：统计原理与 SPSS 应用

作 者：张 黎

出 版 人：张百新

封面设计：中联学林

责任编辑：张 谦

出版发行：新华出版社

地 址：北京石景山区京原路 8 号 邮 编：100040

网 址：<http://www.xinhupub.com> <http://press.xinhuanet.com>

经 销：新华书店

购书热线：010-63077122

中国新闻书店购书热线：010-63072012

照 排：中联学林

印 刷：北京天正元印务有限公司

成品尺寸：170mm×240mm

印 张：21.5

字 数：386 千字

版 次：2017 年 1 月第一版

印 次：2017 年 1 月第一次印刷

书 号：ISBN 978-7-5166-2777-8

定 价：58.00 元

图书如有印装问题，请与印刷厂联系调换：010-89587322

# 前 言

如果要问教育学或心理学等相关专业的学生,在你所学的专业课程中最难的是哪门?一直到毕业后也没学明白的是哪门?为此我还专门针对心理学本科毕业生做了一个小调查,44名毕业生中有15名对这两个问题的回答都是统计学。估计该回答让很多教统计学的老师很尴尬,同时也会觉得理所当然,统计学就是很难嘛!然而教育统计学不仅仅是一门专业课,不仅仅在写毕业论文或做科学研究的时候会用到,教育统计学还会影响到学生理解数据、理解科学、思考问题和认识世界的方式!因此,真正的学会、学懂教育统计学,爱上教育统计学,能够分析和理解数据背后的世界是十分必要的!

本书作者从事心理学本科生的统计学教学多年,之后在为小学教育学、学前教育学专业的本科生、教育学研究生、教育硕士讲授教育统计学时发现当前已出版的教育统计类书籍较少,而且其中的举例多为心理学研究类案例,教育统计学的理论知识与SPSS应用之间也缺少衔接,学生在学习《教育统计学》、《教育统计与测评技术》、《SPSS软件应用》等课程时存在一定的困难。因此作者力争将教育统计学的基础知识以通俗易懂、颇具趣味的形式表达出来,帮助学生轻松地理解统计学原理,爱上教育统计学。总的来说,本书具有以下特点:

每章内容以学生生活或专业学习中的实例引入,设置出问题情境,激起学生的好奇心,想要去学习和探索解决上述问题的统计学方法。

大部分例题的背景均是精选自教育学、学前教育学、心理学中经

典科学研究的实例,学生在结合自己的专业背景轻松学习统计学原理的同时,能够了解到各种经典的研究结论是如何通过统计分析得到的,从而使统计学与其他教育类专业课程的学习有机融合起来。

在本书中尽量将晦涩的统计知识图形化,尽量用图形表达抽象的统计学原理和概念。尤其是在推论统计部分,这部分是本书中的难点(很多同学学习统计学时就在此部分迷失了,无法进入统计学的世界了)。例如,平均数的显著性检验,如果用三幅图形来表示整个检验过程,既可省略复杂的文字描述,而且学生也能轻松理解假设检验的过程,此外学生对图形的记忆比对文字的记忆更牢固,学习的内容能够长久保持下来,从而达到让学生轻松学习和应用统计知识的教学目的。

在本书中将教育统计学和 SPSS 软件应用有机结合起来,使学生在掌握教育统计学原理的基础上,能够轻松使用 SPSS 软件实现对数据的分析和处理,使对教育统计学的学习不再仅停留在让学生深恶痛绝的纸笔计算上,而是能够应用 SPSS 软件进行复杂的数据分析和处理。

为了激发学生的学习兴趣,了解学习的目标,在每章开始部分设置了“本章你将学到什么”版块;为了鼓励学生独立思考,每章中设置了“需要思考”版块,帮助学生对所学内容进行更为深入的消化和思考,从而提升学生思维的复杂性;为了帮助学生整理所学到的知识,在每章中设置了“需要记忆”版块,帮助学生掌握学习和应用的重点,从而提高学生学习的效率。

为了提高学生对学习教育统计学的兴趣,书中还增加了统计轶事、应用实例、统计与网络、统计与生活等内容,课后习题的设置也尽量本着题量丰富、内容有趣的原则,从而激起学生课上学习和课后自主学习的欲望。

本书的写作得到了青岛大学研究生教材建设项目和研究生重点课程建设项目的资助,丁月增老师帮助绘制了本书中大部分的图形,在因素分析部分参照了我的导师张奇教授关于主成分因子分析的诸多想法和心得,在此一并表示衷心的感谢!

本书中所用到的 SPSS 数据库,读者可以利用书中例题的数据自行录入,也可发邮件到 zhangli201207@163.com 向作者索要。

近几年一直为本书的写作准备大量的材料,但是也难免在写作过程中出现错误,还请各位读者给予批评指正!

张 黎

于青岛大学师范学院心理系

# 目 录

## CONTENTS

### 绪 论

第一节	统计学概述 .....	1
第二节	统计中经常使用的术语 .....	4
第三节	数据的特点和类型 .....	9

### 第一部分 描述统计

第一章	集中量数 .....	15
第一节	平均数 .....	15
第二节	众数和中数 .....	19
第二章	差异量数 .....	24
第一节	全距和平均差 .....	24
第二节	方差和标准差 .....	25
第三节	差异系数 .....	28
第三章	地位量数 .....	30
第一节	百分等级 .....	30
第二节	标准分数 .....	32
第四章	描述统计的 SPSS 操作与结果分析 .....	36
第一节	SPSS 软件简介 .....	36

第二节	数据文件的建立 .....	37
第三节	集中量数、差异量数和地位量数的计算与结果分析 .....	42
<b>第五章</b>	<b>相关系数</b> .....	<b>52</b>
第一节	相关关系 .....	52
第二节	相关计算方法 .....	60

## 第二部分 推论统计

<b>第六章</b>	<b>概 率</b> .....	<b>85</b>
第一节	概率和分布 .....	85
第二节	基本随机变量的分布 .....	90
第三节	抽样分布 .....	98
<b>第七章</b>	<b>假设检验</b> .....	<b>115</b>
第一节	假设检验的一般原理 .....	115
第二节	样本平均数的显著性检验 .....	123
<b>第八章</b>	<b>平均数差异的显著性检验</b> .....	<b>133</b>
第一节	独立样本平均数差异的显著性检验 .....	134
第二节	相关样本平均数差异的显著性检验 .....	144
<b>第九章</b>	<b>方差分析</b> .....	<b>152</b>
第一节	单因素方差分析 .....	153
第二节	重复测量的单因素方差分析 .....	164
第三节	完全随机两因素方差分析 .....	176
<b>第十章</b>	<b>非参数检验</b> .....	<b>199</b>
第一节	单样本的非参数检验 .....	200
第二节	两独立样本的非参数检验 .....	208
第三节	两相关样本的非参数检验 .....	219
第四节	多独立样本的非参数检验 .....	225
第五节	多相关样本的非参数检验 .....	230
<b>第十一章</b>	<b>线性回归</b> .....	<b>237</b>
第一节	回归概述 .....	237
第二节	简单线性回归方程的建立和检验 .....	239
第三节	多元线性回归分析 .....	249



### 第三部分 测验数据分析

第十二章 项目分析 .....	263
第一节 项目的难度分析 .....	263
第二节 项目的区分度分析 .....	266
第十三章 主成分因子分析 .....	271
第一节 因素分析的基本原理 .....	271
第二节 主成分因子分析的操作与应用实例 .....	278
第十四章 测验信度和效度的评定 .....	296
第一节 测验信度的评定 .....	296
第二节 测验效度的评定 .....	302
附 录	
附表 1 标准正态分布表 .....	307
附表 2 $t$ 值表 .....	312
附表 3 $\chi^2$ 分布表 .....	314
附表 4(1) $F$ 分布表(双侧检验, $\alpha = 0.05$ ) .....	316
附表 4(2) $F$ 分布表(双侧检验, $\alpha = 0.01$ ) .....	318
附表 5 $F$ 分布表(单侧检验,上一格为 $\alpha = 0.05$ ,下一格为 $\alpha = 0.01$ ) .....	320
附表 6 哈特莱方差齐性检验临界值表 .....	323
附表 7 $q$ 分布的临界值表( $\alpha = 0.05$ ) .....	324
附表 8 单样本 $K-S$ 检验临界值表 .....	325
附表 9 曼—惠特尼 $U$ 检验临界值表(双侧检验) .....	327
附表 10 符号检验表(单侧检验) .....	329
附表 11 符号等级检验表 .....	330
附表 12 弗里德曼双向等级方差分析表 .....	331

# 绪 论

本章你将学到什么

1. 什么是统计学
2. 为什么要学习统计学
3. 需要提前掌握的术语
4. 如何对数据类型进行区分

## 第一节 统计学概述

### 一、什么是统计学

统计学,多么高大上的名字!在人们的印象里好像是只有戴着如瓶底厚的眼镜、嘴里成天叨念着公式、不断在纸上演算的学者才需要学习的科目,似乎和教育或心理类专业学生的生活和学习不太沾边,所以有些刚入学的大学生一听到要学教育或心理统计学都很诧异,纷纷抱怨:“我高中数学不好啊!”“我是文科生,一定学不好了!”其实统计学学的好坏与这些都没关系。而且统计学离我们的生活和专业学习并不遥远,已经成了其中不可或缺的一部分。

我们在日常生活中常常思考下面一些问题:

我的身高、体重、相貌在人群中处于什么水平?

追求心仪的对象,我成功的可能性有多大?

对于减肥的人来说,在一周中,哪天吃得最多?(答案是周六。)

新闻联播上天天在说CPI,这究竟是什么东西啊?和猪肉价格有什么关系啊?

财富和幸福之间是什么关系?

努力学习和体重增加之间是否存在相关?

男老师是不是比女老师更愿意给学生高分?

这筐苹果连吃三个都坏了,其余的是不是也坏了?

根据模拟考试的成绩,我高考能考多少分?

这两种减肥方法相比,哪一种更有效?

在专业课学习过程中,老师常常提出这样的问题:

小班孩子的主要游戏类型是什么?

国家对高等教育经费投入的增长率是多少？

某个 15 岁青少年在韦氏智力测验中得了 100 分，这个青少年的智力在同龄人中处于什么水平？

男生和女生谁更容易产生阅读障碍和学习困难？

特殊儿童应该在特殊教育学校接受特殊教育还是在普通学校接受融合教育？

权威型、专制型、放任型三种父母教养方式，哪种教养方式下孩子的自尊水平最高？

高自尊和低自尊的儿童相比，哪组儿童的心理健康水平更高？

先举例子再讲原理和先讲原理再举例子两种新知识讲授方式，哪种方式更有利于学生对知识的学习？

对知识的集中练习和分散练习，哪种练习方式更有利于知识的保持？

当我们在思考和想要回答类似问题时，我们就是在尝试解决统计问题！虽然我们可以根据个人经验回答其中一些问题，但要科学回答上述问题，就需要收集数据，并对数据进行统计分析，统计学这时候就要戎装出场了。事实上，统计学在我们生活中无处不在，人的一生都是在排序、被排序、比较、被比较，甚至是平均、“被平均”中度过的。我们每次的分析、决策过程都有数据分析过程伴随。统计学家鲁尼恩就将人看成非常老道并善于做出决策的统计工具。英国的预测家威尔斯提出：在将来的某一天，统计思维会像读和写一样成为一个人所必须具备的能力。

那么究竟什么是统计学？

统计学是一种处理数据的方法，通过分析数据，从数据中找出有用的信息并得出结论。统计学又分为数理统计学和应用统计学。数理统计学，主要侧重于统计理论与方法的数理证明。应用统计学主要研究如何将统计原理和方法在教育、心理、金融、医学、管理学、人口统计学等诸多实践领域应用。可见，教育统计学是应用统计学的一个分支。教育统计学就是专门研究如何运用统计学的原理和方法，搜集、整理和分析教育研究中获得的随机性数据资料，并根据这些数据资料传递的信息，进行科学推论找出教育活动规律的一门学科。

## 二、学习统计学的意义

为了动员学生好好学习教育统计学，任课教师在此部分都要大谈特谈一下学习教育统计学的意义。一般来说，我都是先表扬一下我的同学，夸奖他们是我从教以来遇到的最聪明的一届学生，然后请求他们不要把这个消息告诉高年级的学长，因为他们还在以为他们是最聪明的一届学生。然后让学生了解到学习教育统

计学除了有利于学生毕业论文的写作、今后从事科学研究工作外,还能改变学生理解数据、理解科学、思考问题和认识世界的方式!究竟怎么改变的,看完这本书,学完这门课你就可以发现自己的变化!

例如,如果有一项减税政策,将会有 9200 万人享受到减税待遇,人均减税 1083 美元。看到这些数据你会不会支持这项政策?还是需要问对方几个问题?你会问什么问题?一般学过教育统计学的学生都会问下列问题:

这些数据是怎么来的?哪个机构或个人做的研究,数据可信不可信啊?这 9200 万人是谁啊?谁不能享受减税待遇呢?减税的上限是多少?减税的下限是多少?大多数人减税多少?人均减税的中位数是多少?

通过提问,学生发现大多数人才减税 100 美元!现在你还会不会支持这项政策?谁是减税政策的最大受益者?

我曾经给学生看一个牙膏的广告,一位穿着白大褂的外国医生在用极其不标准的普通话为一款添加了新成分的牙膏叫好,声称该牙膏有多重功效。我的一个学生抱怨道:“他说的这些都是无用的信息,要想让我买这个新款牙膏,就要告诉我添加了这种新成分后,这个新款牙膏究竟比老款牙膏或者价钱便宜的普通牙膏好了多少?”

可见,对教育统计学的学习让学生能更好地理解数据的实质和各种统计量的含义,主动去寻找真正支持结论的数据,对各种结论和研究结果更具有批判性精神,能更好地做出理性的决策和行为,而不是简单的被数据所左右。

### 统计应用实例

统计学能够成功解决各个领域中与数据有关的问题。第二次世界大战期间,招募士兵需要对申请者进行血液检查,如果一一进行检查的话,这是一项非常庞大的工程,需要大量的人力、物力、财力和时间。有没有更经济一些的做法呢?

(请聪明的读者在此处停留两分钟稍作思考!)

这时候统计学家出场了,他们建议将血液混合,如果某项疾病平均每 10 人中有一人患病,那么将这 10 个人分为两组,那么有一组的血液会检查出问题,再将出现问题的血液逐一进行检查,则能找出患病的申请者,整个检查下来只需要 6 次,节省了 4 次。最终统计学家还给出了计算最佳的混合样本人数的公式,如果患病率为  $p$ ,那么混合样本人数的最佳值  $n$  应为使  $(1-p)^n - \frac{1}{n}$  的值最大时的  $n$  值。请计算一下如果某项疾病发病率为  $\frac{1}{20}$ ,如何对样本进行混合最为经济。



需要思考的内容

- (1) 统计学与你所学的专业有什么关系?
- (2) 你能不能给出一个统计学成功解决教育领域问题的例子。

## 第二节 统计中经常使用的术语

### 一、变量、自变量和因变量

某研究者要研究父母教养方式对儿童气质类型的影响,在这个研究中“教养方式”和“气质类型”就是两个变量。变量(variable)是一个可以取不同数值的物体属性或事件,在观测不同个体时变量可能具有不同的值。在教育与心理研究中,性别、班级、智力、学习成绩、自我延迟满足时间、教学方法、教育经费等都可作为要研究的变量。通过观测获得的变量值即观测值,也称为具体数据(data)。一般来说,研究要先选取变量,然后通过观察、实验、调查等方法获得变量的观测值或数据,再对数据进行统计分析。

自变量(independent variable)是被研究者操纵和控制的量。因变量(dependent variable)是研究者检验、测量和描述的变量,即研究者观察和记录的随自变量变化而变化的量。例如,有研究者想要探索教师的反馈类型对学生数学成绩的影响,一种反馈是测验结束后立刻给学生做出反馈,另一种是测验结束后一天给学生做出反馈。在这个研究中自变量就是反馈类型,自变量有两个水平,立刻反馈和延迟一天反馈,因变量是学生的数学成绩。

### 二、总体、抽样和抽样方法

总体(population)是指具有某种共同特征的所有个体或事物的集合。构成总体的每个基本单元称为个体(individual)。一般我们关心的是总体的一些量,如总体的平均数、标准差、相关系数、回归系数等。调查总体的研究即为普查,如我国的人口普查,普查常常需要大量的人力、物力和时间。有句谚语说:“你不必吃完整头牛,才知道肉是老的”,因此可以从总体中抽取一部分个体组成总体的一个样本(sample)。例如,研究中国人的智力,从操作角度来说这是不可能的研究,但在实际研究中,可以从总体中抽取样本,然后将结果从样本推论到总体。

实验中被试的数目,或者一个观测重复的次数,称作样本大小(sample size)或样本容量(capacity of sample),用 $n$ 来表示。样本量越大,样本对总体的代表性越强。一般情况下,样本量超过30被称为大样本,小于等于30的样本被称为小样本。

根据两个样本之间的关系又可分为独立样本和相关样本。独立样本是指两个样本之间不相互影响。例如,男生的身高和女生的身高就是两个独立样本,他们的身高之间不存在相互影响。相关样本是指两个样本数据之间存在一一对应的关系。例如,儿童在3岁时候的智力和4岁时候的智力。

抽样(sampling)是指从总体中抽取出一部分样本进行研究。随机化是抽样研究的基本原则,也是抽样研究的灵魂。随机化可以保证样本对总体的代表性。即在抽样时,每个个体是否被抽取,不是由研究者主观决定的,而是每一个个体按照概率原理,被抽中的可能性是相等的,因此有相当大的可能性使样本保持和总体相同的结构。

可见抽样研究与全面调查相比,具有经济性,能够节省人力、物力。对于某些变化较快的心理和教育现象,抽样调查能缩短研究时间从而提高研究的时效性。虽然全面调查涵盖了总体中的所有个体,但进行全面研究时,由于被试(即被研究者)人数众多,需要大量的研究人员,在研究过程中出现差错的可能性也会增加,很难保证每个被试的数据都是准确的,因而会导致测量结果出现误差。

### 统计应用实例

收视率是怎么得来的?

中国大陆地区的收视率主要是有由央视-索福瑞媒介研究有限公司(CSM)提供的,CSM拥有世界上最大的广播电视受众调查网络,选取了6.11万余户样本家庭,通过监测这些家庭的收视纪录,推测中国内地超过12.7亿电视人口的电视观看情况。

民意测验的结果是怎么来的?

2011年秋,美国的民意调查显示有46%的美国人认可奥巴马作为美国总统的工作表现,同样有46%的美国人不可奥巴马的工作表现。美国的人口大约为3亿左右,要得到上述结果就需要从这3亿人中随机抽出一部分样本作为调查的对象,并认为调查的结果约等于全国3亿人的意见,即再做一次调查,调查结果也与这次的结果基本相近。《赤裸裸的统计学》的作者对抽样做了一个非常形象的比喻:一锅汤,舀了一汤勺出来尝一尝,搅一搅,再舀一汤勺出来,两次的味道应该是差不多,这味道基本就是这锅汤的味道。

根据总体的特点和调查目的,可以采用下列几种随机抽样的方法。

简单随机抽样(simple random sampling),如果总体的个数为 $N$ ,将每个个体编上号码,每次从总体中随机抽取一个样本,每个个体被抽到的概率相等。这种抽样方法最能体现随机化原则。看完对简单随机抽样的描述,大家可能立刻想到了彩票,中国福利彩票中的35选7就是采用简单随机抽样的方法选取中奖号码的。

这种抽样方式适用从个数较少的总体中抽取样本。当总体较大或总体的各组成部分之间存在较大差异时,不适合采用简单随机取样。

等距抽样(interval sampling)也称为系统抽样,如果总体的个数为 $N$ ,将每个个体编上号码,然后每隔若干个抽取一个。例如,某高校想了解大学新生对校规校纪的了解情况如何,对200名大学新生进行了编号(1~200),拟从中抽取20人参加校规校纪考试,即每隔10个同学抽取一个。这时只需确定抽取的起点,即从1~10号同学中随机抽取一个同学,之后每个编号加10即可。当总体的编号具有某种周期性变化时不适合采用等距取样。如果编号中的1,11,21……都是班干部,则有可能某次抽取的同学都是班干部,班干部对校规校纪的了解情况不能代表所有大学新生,很可能班干部为了做好本职工作,更加认真地学习了校规校纪。总体的各组成部分之间存在较大差异时,也不适合采用等距抽样,这时可采用分层随机抽样。

分层随机抽样(hierarchical sampling),当总体各部分之间存在较大差异时,根据已有特征,将总体分成几个不同的部分(层),再在每部分(层)中随机取样。划分层次时要保证层内的差异小,层与层之间的差异大。有两种方法可确定各层内的人数。一种是根据各层人数在总体中所占的比例选取被试,即比例分配法。当各层内的标准差已知时,可采用最佳分配法,即在选取被试时,不仅考虑各层人数比例,还考虑各层的标准差,在标准差大的层内多选取被试,在标准差小的层内少选取被试。采用分层随机抽样方法所选取的被试其对总体的代表性要优于简单随机取样,抽样误差也小于简单随机取样。

两阶段随机抽样(two stage random sampling),一般应用在规模非常大的总体中。如对山东省居民的教育消费情况进行随机抽样观察,为了使样本较好地反映总体的情况,又便于调查工作的组织,可以按照该省的行政组织分层,先从省内所有的城市内随机抽取要调查的城市,之后从抽取城市中随机抽取研究对象。两阶段随机抽样简单易行,节省经费,但与简单随机取样相比抽样误差要大一些。

但在实际研究中并不是所有通过随机抽样得到的样本都是最好的样本,抽样也会发生误差,使样本结果与总体的结果不一致。在选取样本时由于机遇所造成的样本结果与总体结果的差异被称为随机抽样误差,随着样本量的增加,随机抽样误差会减小。

不当的抽样方法也会导致抽样误差。例如,有些研究采用自发性回应的方法收集数据,即把调查问卷通过邮件、信件或夹在杂志中发下去,然后收集主动寄回问卷的结果。有一位作家曾经对读者做过一个调查:“如果可以重来一次,你还要孩子吗?”最终收到1万份读者答复信,其中将近70%回答:不要!这是很惊人

的数字。难道说70%的父母都后悔有了孩子吗？当然不是！这是个自发性回应样本。通常对某个议题有强烈感觉的人，尤其是有负面感觉的人，会比较不嫌麻烦地去回应。由于自发回应人群所构成的样本和总体结构差异比较大，所以得到70%这样一个不真实还很吓人的结果。

再比如，很多电视栏目在做一些热点问题的节目时，常常喜欢加入民意调查部分。一想到了解民意，他们就非常喜欢街头采访。其实街头采访的样本构成与总体构成之间存在很大的差异。例如，有电台记者想要调查我国二胎政策放开后，愿意生二胎女性的比例。那么在街头找到的愿意接受采访的女性往往都是时间充裕、喜欢逛街、愿意在电视台露脸，甚至连一胎还没生，没真实考虑过生二胎的女性！那些在家带孩子、无暇逛街的、真正面临生二胎问题的女性则很少在受访群体里，所以得到的结果会与真实的女性二胎生育意愿存在很大的差别。

此外，一些在抽样过程中研究者和被研究者因素也会导致样本结果和与总体的结果不一致。如数据录入错误，尤其是在录入大量数据的时候，很难核查，容易发生错误。比较简单的避免录入错误的方法是采用具有双录功能的软件录入数据，如EpiData软件允许两个人分别录入同一份数据，之后利用软件中的比对功能对录入结果进行比较。如果两次录入不一致，则会发出提醒。被研究者不认真的回应调查问题或者根本不做任何回应也会影响最终调查的结果。一些涉及被调查者利益的问题，很难得到真实的调查结果。例如，上个月我忘记点名了，然后调查选课的学生，上个月的教育统计课你缺勤了几次？一般缺勤过的学生或多次缺勤的学生很少会给出真实的数字。即使有些问题无关被调查者的利益，也很难得到准确的调查结果。如非常简单的问题：你上个月花了多少钱？如果回答者没有记录每天消费的习惯，则很难给出准确的回答。

如果提问时的语言加入暗示性信息，也会影响调查的结果。例如，有心理学研究者随机选了一群被试者观看一个影片，里面有车祸的视频，观看完之后将被试者随机分为两组，问一组的被试者：“两车撞击在一起时，红车的车速是多少？”问另外一组被试者：“两车猛烈撞击在一起时，红车的车速是多少？”在第一种情况下被试者回答的平均速度为50公里每小时，第二种情况下平均速度竟然为65公里每小时！不同的提问方式对速度的估计竟然相差15公里每小时。同样的问题，不同的表述或措辞，得到的结果也是不一样的。举一个比较极端的例子。“老师，我在做统计作业的时候可以玩手机吗？”估计所有老师都不同意。学生换种说法问老师：“老师，我玩手机的时候可以思考统计作业吗？”估计老师会欣然同意的。





### 需要思考的内容

在接受某一研究结论之前,应该进行什么样的批判性思考?

《统计学的世界》一书的作者戴维·穆尔(David S Moore)就提议在相信抽样调查的结果之前要问下面几个问题:

谁做的调查?

总体是什么?

样本是怎么选取的?

样本多大?

受访对象的回应率是多少?

用什么方式联系的受访者?

调查是什么时候做的?

调查问题是怎么问的?

### 三、参数、统计量和变异性

表示总体统计特征的量数,如总体平均数、总体标准差、总体相关系数、总体回归系数被称为总体参数(parameter)。总体是确定的,参数就是常数。总体参数一般用希腊字母来表示,如 $\mu$ 、 $\sigma$ 、 $\rho$ 等,但是更多的时候,由于各种原因无法对总体中所有的个体进行测量,所以往往选择总体中有代表性的个体作为样本进行观测,这时候计算出来表示样本统计特征的量数,如样本的平均数、样本的标准差、样本的相关系数、样本的回归系数,就是样本统计量(statistics)。样本是随机抽样的,统计量是可变的,随着样本的变化而变化。样本统计量一般用英文字母表示,如 $\bar{X}$ 、 $s$ 、 $r$ 等。

变异性(variability)是指在相同条件下,观测不同个体得到的数据结果存在差异的特点。一般来说导致数据变异性的原因有三种:

(1)个体差异,即研究对象之间存在差异。世界上没有两片相同的树叶,也没有完全相同的两个人。每个个体在个性特征、兴趣、爱好、价值观、反应速度等各个方面都是不同的,因此测量结果间必然存在一定的差异。

(2)测量误差,即由于观测人员、观测工具、观测条件等方面的原因导致观测误差的存在。在心理和教育研究中,几乎所有的测量结果都会包含着某种程度的测量误差。例如,对儿童集中注意能力的测量,可采用划消测验。具体做法是给儿童呈现一些数字,要求儿童在3分钟内将其中的数字9划掉,如果9前面是5就不能划掉。有些研究者在开始测验前没有给儿童示范如何对划错的数字进行修