PEARSON
Prentice
Hall

高等院校双语教材
社会学系列

社会研究中的基础统计学
（第十版）

Elementary Statistics in Social Research
(Tenth Edition)

[美] 杰克·莱文（Jack Levin）
詹姆斯·艾伦·福克斯（James Alan Fox） 著
王卫东 改编

中国人民大学出版社

PEARSON
Education

Cast of Characters (in order of appearance)

Symbol	*Meaning/Description*
N	Number of cases
f	Frequency
$\%$	Percentage
cf	Cumulative frequency
$c\%$	Cumulative percentage
m	Midpoint
PR	Percentile rank
Mo	Mode
Mdn	Median
$\overline{X}$	Mean
R	Range
MD	Mean deviation
s^2	Variance
s	Standard deviation
μ	Population mean
σ^2	Population variance
σ	Population standard deviation
P	Probability
z	z score
ME	Margin of error
CI	Confidence interval
$\sigma_{\overline{X}}$	Standard error of the sample mean (true)
$s_{\overline{X}}$	Standard error of the sample mean (estimated)
α	Level of significance
df	Degrees of freedom
t	t ratio
P	Sample proportion
π	Population proportion
s_P	Standard error of the sample proportion

Symbol	*Meaning/Description*
$\overline{X}_1 - \overline{X}_2$	Difference between (sample) means
$\sigma_{\overline{X}_1 - \overline{X}_2}$	Standard deviation of the distribution of the difference between means
$s_{\overline{X}_1 - \overline{X}_2}$	Standard error of the difference between means
s_D	Standard deviation of the distribution of the (before/after) difference between means
$s_{\overline{D}}$	Standard error of the (before/after) difference between means
P*	Pooled sample proportion
$s_{P_1 - P_2}$	Standard error of the difference between proportions
SS	Sum of squares
MS	Mean square
F	F ratio
HSD	Honestly significant difference
q	Studentized range
χ^2	Chi-square
r	Pearson's correlation
ρ	Population correlation
$r_{XY.Z}$	Partial correlation
a	Y-intercept
b	Slope
e	Error term
r^2	Coefficient of determination
$1 - r^2$	Coefficient of nondetermination
R^2	Multiple coefficient of determination

导 言

杰克·莱文和詹姆斯·艾伦·福克斯所著的《社会研究中的基础统计学》是美国流行的统计学教材之一。这本教材自推出以来，已经发行到了第十版。经历了这么多年的教育实践和激烈的教材市场竞争仍能保持旺盛的生命力，正是这本书质量的最好说明。

这本书的特点是简单、可读性强、而又覆盖全面，非常适合中国非统计学专业的人文和社会科学专业本科生用作学习统计学的英文教材。中国的学生通过原版教材学习某门课程时最常遇到的问题是课程本身就有较大的难度，而外文行文又晦涩难懂，使得同学们望而却步或半途而废。而这本书的内容可以说是非常简单易懂，对使用者的教学背景基本没有任何要求，完全是“手把手”地讲解一个个统计学原理和案例；同时，这本书中的英文也非常浅显，没有什么生僻的单词和语汇，更没有大段的复杂从句；另外，这本书中的例子也非常生动，完全来自现实生活，而不是统计学课本中常见的“从黑球中摸白球”这类枯燥的例子。说这本书易懂并不是说这本书的内容浅，实际上，这本书最大的优点是能深入浅出地讲解统计学。学习过这本书后，对于大学本科阶段所需掌握的统计学知识就有了一个全面的掌握，可以为进一步的学习打下坚实的基础。

这本教材共分 11 章，具体是：

第 1 章：为什么社会研究者要使用统计学。讲述统计学在社会科学研究中的作用。

第 2 章：组织数据。讲述在一般研究中如何对数据进行初步的处理和呈现。

第 3 章：集中趋势的测量。讲述均值、中位数、众数这三个常用的统计指标的原理和计算。

第 4 章：变异性的测量。讲述对变量变异性测量的统计指标和计算。

第 5 章：概率与正态曲线。引入了概率的基本思想。重点介绍了统计学中最重要的正态分布的原理和应用。

第 6 章：样本与总体。讲述了抽样中的统计学原理，以及与抽样相关的重要的分布和统计学概念。

第 7 章：均值间差异的显著性检验。在前面各章的基础上，讲述如何对两个均值的差异性进行检验。

第 8 章：方差分析。引入了统计模型的概念，对方差分析模型从原理到应用进行了全面的讲述。

第 9 章：非参数。讲述了在总体分布未知的情况下如何进行假设检验，重点介绍了单因素卡方检验和双因素卡方检验。

第 10 章：相关。介绍了相关性的思想。重点讲述了皮尔逊相关系数的背景、原理以及计算。

第 11 章：回归分析。引入了回归模型这个因果分析模型，讲述了这个模型的思想和计算，以及和前面各章中介绍过的统计模型的关系。

为了便于同学们的集中学习，本书是对英文原版的改编本。删除了原书中在当前的统计实践中基本不再使用的一些内容。改编对教材的作用没有任何影响，反而会使得对本书的学习更有效率。另外，本教材的第九版的中译本已由中国人民大学出版社出版，两本书结合使用，会对同学们的学习更有帮助。

改编者　王卫东

2008 年 12 月

前 言

《社会研究中的基础统计学》第十版向社会学及相关的刑事司法学、政治学和社会工作专业的学生们提供了一个基本的统计学介绍。这本书不是一本对统计方法的全面参考书，恰恰相反，我们的首目标是它能够被大多数学生所理解，尤其是那些可能并没有很强的数学背景的学生。

像它的前几版一样，第十版拥有一些教学上的特点。其中最突出的是对于全书的所有重点都有详细的统计过程的逐步演示。我们也对社会研究中统计方法的原理和使用给予了清楚而富有逻辑性的解释。同时，在每一章的后面都附有大量的习题，而绝大多数都在书后给出了答案。最后，这本书的每一部分都以一个我们称之为“详细演示”的内容结束，它将带领学生们体验一项假想的对高级中学学生抽烟和喝酒行为的调查研究的整个过程。

本书的最新版本有一些改进与提高。其中最重要的是所有的计算练习和演示都经过了不厌其烦的检查和验算，以使学生们因为计算上的错误而受到的困扰减到最小。由于追求完美是我们的目标之一，我们欢迎老师和学生把他们的评论和指正发送电子邮件到 J.levin@neu.edu。我们增加了书中的例子和习题，以求能更好地代表实际的案例。我们也增加和扩展了对几个重要议题的讨论，具体包括：第 1 章中的测量（离散和连续变量、测量水平的升级和降级），以及第 2 章中的描述技术（组限、条形图和直方图）。对各章后面的小结也做了进一步的改进和扩展，以求它们对学生们更加有用。

在第 1 章的详尽综述之后，这本书被分成四部分。第一部分（第 2 章至第 4 章）向学生介绍了描述和比较数据最常用的方法。第二部分（第 5 章和第 6 章）是一个过渡性的部分，它以对概率的基本概念的讨论开始，然后继续引导学生从了解正态曲线如何作为一个重要的描述工具一直到它在抽样中的基本运用。第三部分（第 7 章至第 9 章）以统计推断为核心，涵盖了几种最常用的显著性检验。第四部分（第 10 章至第 11 章）则包括了如何获取相关系数的过程，并介绍了回归分析。

我们要感谢本书第十版的下列各位评阅人，感谢他们敏锐而有益的建议，他们是：马歇尔大学的 Robert Bickel，佛罗里达国际大学的 David Elite，内华达大学拉斯维加斯分校的 Kenneth F. Fernandez，得克萨斯大学圣·马科斯分校的 Debarun Majumdar，得克萨斯 A&M 大学的 Dorothy J. Pace，内华达大学拉斯维加斯分校的 Satish Cope。我还要感谢东北大学的 Jenna Savage 和 Sarah Cope，她们花了大量的时间来验证我们的演算和答案。Jenna Savage 同时也负责添加了大量优秀的章后习题。我们还要感谢英国皇家学会的文学部主管 Ron A. Fisher 爵士、英国皇家学会的 Rank Yates 博士，以及伦敦的朗曼出版集团，感谢他们准许本书借用他们 1974 年出版的《生物学、农学和医学研究中的统计用表》（Statistical Tables for Biological，Agricultural and Medical Research）（第 6 版）一书中的表 3、表 4、表 5 和表 6。最后，我们还要感谢我们的个人电脑所起的重要作用，没有它们的协助，这次修订是不可能完成的。

杰克·莱文
詹姆斯·艾伦·福克斯

Brief Contents

简明目录

Contents

1 Why the Social Researcher Uses Statistics

A little of the social scientist can be found in all of us. Almost daily, we take educated guesses concerning the future events in our lives in order to plan for new situations or experiences. As these situations occur, we are sometimes able to confirm or support our ideas; other times, however, we are not so lucky and must face the sometimes unpleasant consequences.

Consider some familiar examples: We might invest in the stock market, vote for a political candidate who promises to solve domestic problems, play the horses, take medicine to reduce the discomfort of a cold, throw dice in a gambling casino, try to psych out our instructors regarding a midterm, or accept a blind date on the word of a friend.

Sometimes we win; sometimes we lose. Thus, we might make a sound investment in the stock market, but be sorry about our voting decision; win money at the craps table, but discover we have taken the wrong medicine for our illness; do well on a midterm, but have a miserable blind date; and so on. It is unfortunately true that not all of our everyday predictions will be supported by experience.

The Nature of Social Research

Similar to our everyday approach to the world, social scientists attempt to explain and predict human behavior. They also take "educated guesses" about the nature of social reality, although in a far more precise and structured manner. In the process, social scientists examine characteristics of human behavior called *variables*—characteristics that differ or vary

from one individual to another (for example, age, social class, and attitude) or from one point in time to another (for example, unemployment, crime rate, and population).

Not all human characteristics vary. It is a fact of life, for example, that the gender of the person who gave birth to you is female. Therefore, in any group of individuals, gender of mother is the *constant* "female." A biology text would spend considerable time discussing why only females give birth and the conditions under which birth is possible, but a social scientist would consider the mother's gender a given, one that is not worthy of study because it never varies. It could not be used to explain differences in the mental health of children because all of their mothers are females. In contrast, a mother's age, race, and mental health are variables: In any group of individuals, they will differ from person to person and can be the key to a greater understanding of the development of the child. A researcher therefore might study differences in the mental health of children depending on the age, race, and mental health of their mothers.

In addition to specifying variables, the social researcher must also determine the *unit of observation* for the research. Usually, social scientists collect data on individual persons. For example, a researcher might conduct interviews to determine if the elderly are victimized by crime more often than younger respondents. In this case, an individual respondent is the unit to be observed by the social scientist.

However, researchers sometimes focus their research on *aggregates*—that is, on the way in which measures vary across entire collections of people. For example, a researcher might study the relationship between the average age of the population and the crime rate in various metropolitan areas. In this study, the units of observation are metropolitan areas rather than individuals.

Whether focusing on individuals or aggregates, the ideas that social scientists have concerning the nature of social reality are called *hypotheses.* These hypotheses are frequently expressed in a statement of the relationship between two or more variables: at minimum, an *independent variable* (or presumed cause) and a *dependent variable* (or presumed effect). For example, a researcher might hypothesize that socially isolated children watch more television than children who are well integrated into their peer groups, and he or she might conduct a survey in which both socially isolated and well-integrated children are asked questions regarding the time they spend watching television (social isolation would be the independent variable; TV-viewing behavior would be the dependent variable). Or a researcher might hypothesize that the one-parent family structure generates greater delinquency than the two-parent family structure and might proceed to interview samples of delinquents and nondelinquents to determine whether one or both parents were present in their family backgrounds (family structure would be the independent variable; delinquency would be the dependent variable).

Thus, not unlike their counterparts in the physical sciences, social researchers often conduct research to increase their understanding of the problems and issues in their field. Social research takes many forms and can be used to investigate a wide range of problems. Among the most useful research methods employed by social researchers for testing their hypotheses are the experiment, the survey, content analysis, participant observation, and secondary analysis. For example, a researcher may conduct an experiment to determine if arresting a wife batterer will deter this behavior in the future, a sample survey to investigate political opinions, a content analysis of values in youth magazines, a participant observa-