

统计学基本概念

13.1 统计学实例

统计学(Statistics)：收集与分析数据的科学与艺术。

统计学与概率论的宗旨都是把不确定现象量化。

概率论是数学，其基本特征是从法则到结果(from rules to results)，

统计学是一门科学，其基本特征是从结果到法则(from results to rules)。

统计学能够发挥作用的领域不胜枚举，举例如下：

- 科学：实证的科学研究离不开搜集和分析数据；
 - 技术：技术的创新和改进离不开作试验和对试验数据的分析；
 - 工农业生产：改进质量或提高产量离不开作试验和对试验数据的分析；
 - 经济金融：对经济金融形势的分析与展望需要建立模型，离不开对大量数据的分析；投资、保险、股票等；
 - 政府或公司的管理和决策——进行量化的管理和决策；
- 天气、医药、人口、社会卫生、收入保障体系的制定等等。

实例一 统计学帮助政府解决难题

印度独立不久，首都发生暴乱，大量难民聚集到被称为红色堡垒的被保护区域。

政府有责任给难民提供食品等必需品，将这个任务委托给了承包商。

政府担心承包商夸大商品的需求量，从中牟取暴利，但无从入手解决

有人建议让统计学家完成这个任务。但问题显得很困难。

全体避难者一天所需的米、豆和盐的总量为 R, P, S 。

普通人的消费调查，每人每天平均所需米、豆、盐的量分别为 r, p, s

R/r, P/p, S/s

实例二 小儿麻痹症疫苗的有效性

问题： 小儿麻痹疫苗问世后，1954 年进行了一项研究以评价它在预防幼儿麻痹及死亡方面的有效性。两组幼儿参加了这项研究。试验组按规定接受三次疫苗，对照组则不接受疫苗。对照组作为证实疫苗有效性，进行对比是必须的。比较的最重要的判据是两组中发生麻痹以及死亡的幼儿数。由于小儿麻痹症发病率极低，两组都需要大量的幼儿以保证有足够的病例发生，从而为比较提供可靠的基础。Meier 的文章称该项研究是“有史以来最大规模的公共卫生试验”。两组人数都略多于 200000 名小孩。

实例二 小儿麻痹症疫苗的有效性

随机决定每个小孩是否接受疫苗。

表 13.1 试验组和对照组小儿麻痹发病率

组别	幼儿人数	发病人数	发病率（每十万人）
试验组	200745	33	16
对照组	201229	115	57

测试中存在随机因素，如何在概率意义下量化差异？两组幼儿的发病率是否有本质的差异？

回答这些问题是统计推断的重要内容。

统计学有自己独特的思维方式与方法

● 统计的目的是回答实际领域中提出的各种问题，对科学结论提供定量分析（而不是单纯的定性分析）的依据；为发现新的理论模型提供线索；预测未来，为决策提供支持等。因此统计学以问题为导向，而不以理论为导向。统计学本质上是一门应用性、方法性的学科。

● 统计的对象——数据是部分的、具有不确定性的，因此所有的统计结论都可能是错的！通过统计学不可能得到绝对正确的结论，只能设法尽量降低因犯错误所造成的损失。

● 对具体问题所产生数据的统计分析结果不必具有数学和科学意义一般性，必须与实际问题的背景相结合加以解释。简单地说“**Yes**” or “**No**”，会让你误入歧途！

● 没有一种统计方法是“最好的”，反之，对同一组数据有可能用不同的方法去分析，甚至得到相互矛盾的结论！任何一组数据都有一定的背景。用什么方法分析可以得到“好的”或“更好的”结果？这需要大量统计分析的实践经验。在某种程度上可以说：统计既是科学，又是艺术。
