

# 统计思想的构建及其应用

孙天义, 贾梦阳, 李沛宸

信阳师范大学教育科学学院, 河南 信阳

收稿日期: 2025年11月3日; 录用日期: 2025年11月24日; 发布日期: 2025年12月5日

---

## 摘要

统计思想是做研究不可或缺的科学思维方式, 它贯穿于研究设计、数据分析和结果解读的各个环节。通过学习统计学基础知识, 包括数据类型、基本统计量、统计方法等, 研究新手可以奠定构建统计思想的基础。进行实际数据分析操作, 运用统计软件处理数据, 从简单案例逐步深入, 能够将理论知识转化为实践能力。阅读统计学相关文献, 了解不同领域的应用案例, 学习统计方法的选择与运用技巧以及结果解读方式, 可以拓宽视野, 提升统计应用水平。文章对研究者, 尤其是研究新手如何快速掌握统计工具有一定的启发。

---

## 关键词

统计思想, 构建, 应用

---

# The Construction and Application of Statistical Thought

Tianyi Sun, Mengyang Jia, Peichen Li

School of Educational Science, Xinyang Normal University, Xinyang Henan

Received: November 3, 2025; accepted: November 24, 2025; published: December 5, 2025

---

## Abstract

Statistical thinking is an indispensable scientific way of thinking for conducting research, running through all aspects of research design, data analysis, and result interpretation. By learning the basic knowledge of statistics, including data types, basic statistics, and statistical methods, novice researchers can lay the foundation for building statistical thinking. Conducting practical data analysis operations, using statistical software to process data, and gradually delving into more complex cases can transform theoretical knowledge into practical skills. Reading literature related to statistics,

**understanding application cases in different fields, and learning the selection and application techniques of statistical methods, as well as result interpretation methods, can broaden one's horizons and enhance the level of statistical application. This article provides some inspiration for researchers, especially novice researchers, on how to quickly master statistical tools.**

## Keywords

**Statistical Thinking, Construction, Application**

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在数据驱动的研究时代，统计思想对于研究新手而言犹如指南针，引领着他们在复杂的数据海洋中探索。从社会科学领域中对不同群体行为模式和观点态度的研究，到自然科学里对实验数据的分析以及医学中对临床试验结果的解读，统计学无处不在[1]。据相关统计，在发表于核心学术期刊的论文中，超过 70% 的研究使用了不同程度的统计分析方法来支持其研究结论[2]。对于初涉研究领域的学生来说，构建起坚实的统计思想，是做好研究设计、准确进行数据分析以及深度挖掘数据价值的关键[3][4]。文章旨在为这些新手提供一套全面且实用的指南，助力他们逐步构建起统计思想，在研究道路上迈出坚实的步伐。

## 2. 统计思想及其核心要素

统计思想是一种科学的思维方式，它贯穿于整个研究过程，从数据的收集、整理、分析到解释结果，都离不开统计思想的指导[5]。其核心在于通过对数据的研究来揭示现象背后的规律和特征，并且在不确定性中做出合理的推断和决策。

统计学研究对象具有变异性[6]。以医学研究为例，不同患者对同一种药物的反应存在差异，即使是年龄、性别、病情相似的患者，其治疗效果也不尽相同。这种变异性使得我们不能仅凭个别案例来得出结论，而需要运用统计方法对大量数据进行分析。通过收集众多患者的治疗数据，计算治愈率、不良反应发生率等统计指标，并进行统计检验，我们才能较为准确地评估药物的疗效和安全性。

统计思想强调从样本推断总体[7]。在实际研究中，往往由于时间、成本等因素限制，无法对总体中的每个个体进行研究，只能抽取一部分个体作为样本。例如，在一项关于全国初中生心理健康状况的研究中，不可能对所有初中生进行调查，而是从各个地区、不同类型高校中抽取一定数量的学生作为样本。通过对这些样本数据的分析，运用适当的统计推断方法，如区间估计和假设检验，来推断总体初中生的心理健康状况。这种从样本到总体的推断过程，是统计思想的重要体现。

统计结果具有概率性。这意味着我们所得到的结论不是绝对的，而是在一定概率水平下成立[8]。例如，在进行假设检验时，我们设定一个显著性水平(如 0.05)，如果计算得到的 p 值小于这个显著性水平，我们就拒绝原假设，认为存在显著差异或效应。但这并不意味着原假设绝对不成立，只是在当前样本数据下，有 95% 的把握认为存在差异。也就是说，还有 5% 的可能性是我们犯了错误(即第一类错误，拒绝了实际上正确的原假设)。这种概率性的认识，让我们在解读统计结果时更加谨慎和客观。

### 3. 统计思想在研究中的作用

#### 3.1. 助力研究设计

统计思想在研究设计中发挥着系统性指导作用，主要体现在以下几个方面，一是研究设计的优化与验证。统计思想通过实验设计原则(如随机化、对照组设置)确保研究可行性，同时帮助合理分配资源以提升效率。例如，医学研究中通过样本容量计算和随机抽样减少偏差，增强结论的可靠性。二是数据规律的科学挖掘。统计方法(如描述性统计、回归分析)为数据简化、特征描述及变量关系分析提供工具，使研究结果更具解释力。例如，在社会学研究中，推断性统计通过参数估计和假设检验实现从样本到总体的科学推论。三是结论可靠性的量化评估。统计学工具(如置信区间、 $p$  值)可量化结果的可信度，并支持因果推断。如在临床试验中，效应量评估和 Meta 分析可以进一步强化结论的普适性。四是跨学科应用与成果转化。统计思想在流行病学(如 RR 值计算)、公共卫生(趋势预测模型)等领域扩展研究边界，推动循证决策和知识积累。其方法论价值贯穿从数据收集到结论应用的全流程。

#### 3.2. 指导数据分析

完成数据收集后，统计思想用于指导选择数据分析的方法。比如描述性统计是通过计算均值、中位数、众数等统计量来描述数据的集中趋势。均值是所有数据的算术平均值，能反映数据的平均水平，但容易受到极端值的影响。例如，在一组员工工资数据中，如果有少数高管的工资极高，那么均值会被拉高，不能很好地代表大多数员工的工资水平。此时，中位数(将数据按照大小顺序排列后位于中间位置的数值)可能更能反映一般员工的工资情况。方差、标准差等统计量用于描述数据的离散程度。方差是每个数据与均值之差的平方的平均值，标准差是方差的平方根。较大的方差表示数据的离散程度较大，即数据较为分散。在分析学生考试成绩时，如果一个班级成绩的标准差较小，说明学生成绩较为均匀，差距不大。而标准差较大，则说明学生成绩差异较大。

而推断统计是通过样本数据对总体参数特征进行推断。例如，相关分析可以帮助我们了解两个或多个变量之间的关联程度。但需要注意的是，相关性并不等同于因果关系，即使两个变量高度相关，也不能直接得出一个变量导致另一个变量变化的结论；回归分析则可以建立变量之间的数学模型，用于预测和解释。在经济学研究中，常使用回归分析来研究消费( $Y$ )与收入( $X$ )之间的关系。通过收集消费者的收入和消费数据，建立回归方程，如  $Y = \beta_0 + \beta_1 X + \varepsilon$ (其中  $\beta_0$  和  $\beta_1$  是回归系数， $\varepsilon$  是误差项)，可以根据收入的变化来预测消费的变化，并且通过回归系数  $\beta_1$  来解释收入对消费的影响程度。

#### 3.3. 辅助结果解读

统计思想对于正确解读研究结果至关重要。在进行假设检验时，我们需要根据统计结果判断是否接受或拒绝原假设。例如，在比较两种教学方法对学生成绩影响的研究中，原假设可能是两种教学方法没有差异。通过对两组学生成绩数据进行  $t$  检验，如果得到的  $p$  值小于设定的显著性水平(如 0.05)，我们就拒绝原假设，认为两种教学方法存在显著差异，即一种教学方法比另一种更能提高学生成绩。但如果  $p$  值大于显著性水平，我们不能拒绝原假设，此时需要谨慎解释结果，不能简单地认为两种教学方法没有差异，而可能是由于样本量不足等原因导致未能检测出差异。

置信区间也为结果解读提供了重要信息。例如，在估计总体均值时，我们得到一个 95% 置信区间，这意味着如果重复进行抽样和计算置信区间的过程，理论上有 95% 的置信区间会包含真实的总体均值。例如，通过样本数据计算出某地区居民月平均收入的 95% 置信区间为(5000 元, 6000 元)，我们可以说有 95% 的把握认为该地区居民的真实月平均收入在这个区间内。这让我们对估计结果的可靠性有了更直观的认识，而不仅仅依赖于点估计值。

在实际研究中，可能会出现一些看似矛盾或复杂的统计结果，此时需要运用统计思想进行深入分析。例如，在一项关于某疾病危险因素的研究中，单因素分析发现某个因素与疾病发生存在显著关联，但在多因素分析中，该因素的显著性消失了。这可能是因为该因素与其他因素存在共线性，在多因素模型中，其他因素对疾病发生的影响掩盖了该因素的作用。通过进一步分析变量之间的关系，运用统计方法(如方差膨胀因子检验来检测共线性)，可以更准确地理解结果，避免得出错误结论。

## 4. 构建统计思想的方法

### 4.1. 学习统计学基础知识

#### (一) 从基础概念学起

统计学中有许多基础概念，是构建统计思想的基石。数据类型的理解至关重要，数据可分为定量数据和定性数据。定量数据又可细分为离散型数据和连续型数据。离散型数据取值是有限个或可列个，如班级学生人数、家庭拥有汽车数量等；连续型数据在一定区间内可以取任意值，如身高、体重、时间等。定性数据则分为分类数据和顺序数据。分类数据各类别之间没有顺序之分，如性别(男、女)、民族等；顺序数据类别之间存在顺序关系，如产品质量等级(优、良、中、差)、疾病严重程度(轻度、中度、重度)等。不同的数据类型适用不同的统计分析方法，正确识别数据类型是进行有效分析的前提。

#### (二) 掌握基本的统计方法

描述性统计方法是数据分析的第一步。除了计算各种集中趋势和离散程度的统计量外，还包括数据的可视化展示。通过绘制直方图，可以直观地了解数据的分布形状，判断数据是否近似服从正态分布等常见分布。例如，在分析一批产品质量数据时，绘制直方图发现数据呈现中间高、两边低的钟形分布，初步判断可能近似正态分布。箱线图则可以同时展示数据的中位数、四分位数、最大值、最小值以及异常值，方便比较不同数据集的分布特征。在比较不同班级学生考试成绩时，绘制箱线图可以清晰地看出各班成绩的集中趋势、离散程度以及是否存在异常值。

推断统计方法是从样本推断总体的重要工具。假设检验的基本步骤包括提出原假设和备择假设、选择合适的检验统计量、确定显著性水平、计算检验统计量的值并与临界值比较或计算  $p$  值进行决策。例如，在检验某品牌灯泡的平均使用寿命是否达到宣传的 1000 小时时，原假设为平均使用寿命等于 1000 小时，备择假设为平均使用寿命不等于 1000 小时。根据样本数据计算出  $t$  检验统计量的值，与给定显著性水平下的  $t$  分布临界值比较，或者计算  $p$  值，如果  $p$  值小于显著性水平，则拒绝原假设，认为该品牌灯泡平均使用寿命与宣传不符。

置信区间的计算也是推断统计的重要内容。以总体均值的置信区间为例，当总体方差已知且样本量较大时，使用  $Z$  分布来计算置信区间；当总体方差未知时，使用  $t$  分布来计算。例如，在估计某地区居民家庭月平均支出时，从该地区随机抽取 100 户家庭作为样本，计算出样本均值和样本标准差，由于总体方差未知，根据  $t$  分布计算出该地区居民家庭月平均支出的 95% 置信区间，为决策者提供了关于总体参数的一个估计范围。

#### (三) 借助教材和课程系统学习

许多统计学教材可供选择。对于初学者而言，《现代心理与教育统计学》(张厚粲、徐建平著)是一本广泛使用的入门教材，它以通俗易懂的语言介绍统计学的基本概念、方法和应用，帮助读者理解和掌握统计知识。《商务与经济统计》(安德森等著)则侧重于统计学在商务和经济领域的应用，详细讲解了如何运用统计方法解决实际商业问题，如市场调研、质量控制、销售预测等。在学习教材时，要注重理论知识的理解，同时结合书中的例题和练习题进行实践，加深对知识点的掌握。

在线课程也是学习统计学的有效途径。中国大学 MOOC 平台上有许多高校开设的统计学课程，如北京大学的“统计学”课程，系统地介绍了统计学的基本原理和方法，通过讲解、案例分析、作业等环节，帮助学生逐步掌握统计学知识。Coursera 上的“Statistics with R”课程，不仅教授统计学理论，还结合 R 语言进行实际数据分析操作，让学生在实践中学习统计方法的应用。通过观看课程视频、参与在线讨论、完成作业和测验等方式，可以更系统地学习统计学知识，并且与其他学习者交流心得，共同进步。

## 4.2. 进行实际数据分析操作

### (一) 寻找合适的数据资源

在实际操作中，首先需要获取合适的数据。政府公开数据平台是丰富的数据来源之一，如国家统计局官网，提供了大量宏观经济数据、人口数据、行业数据等。例如，研究人员可以从该网站获取历年的国内生产总值(GDP)数据、各行业的就业人数数据等，用于经济领域的相关研究。

学术数据库中的数据具有较高的研究价值。Web of Science、中国知网等数据库收录了众多学术期刊论文，其中许多论文附带了研究数据。在医学研究领域，通过检索相关疾病的研究论文，可以获取到患者的临床数据，如症状表现、治疗方法、治疗效果等，用于进一步的分析和研究。

### (二) 运用统计软件处理数据

统计软件是进行数据分析的有力工具。SPSS 就是一款专业的统计分析软件，操作相对简单，适合初学者。它提供了丰富的统计分析模块，包括描述性统计、假设检验、方差分析、回归分析等。在进行问卷调查数据分析时，使用 SPSS 可以轻松地对问卷数据进行录入、编码和分析。例如，对于一份关于消费者满意度的调查问卷数据，通过 SPSS 的可靠性分析可以检验问卷的信度，通过因子分析可以提取影响消费者满意度的主要因素，通过相关性分析和回归分析可以研究不同因素与消费者满意度之间的关系。

R 语言是一种强大的开源统计编程语言，在数据分析和统计建模方面具有广泛的应用。它拥有丰富的软件包，涵盖了各种统计方法和数据分析技术。例如，使用 ggplot2 软件包可以进行精美的数据可视化，绘制出各种复杂的统计图表；使用 dplyr 软件包可以高效地进行数据处理和转换；使用 lm ()函数可以进行线性回归分析，使用 glm ()函数可以进行广义线性回归分析等。

### (三) 从简单案例逐步深入

在开始进行数据分析实践时，可以从简单的案例入手。例如，分析自己班级同学的考试成绩数据，计算成绩的均值、中位数、标准差，绘制成绩的直方图，观察成绩的分布情况，比较不同科目成绩之间的相关性等。通过这个简单案例，熟悉数据处理和分析的基本流程，掌握描述性统计方法的应用。

随着对统计方法的熟悉，可以逐渐增加案例的难度。比如，开展一项关于校园内学生消费行为的调查，收集学生的月生活费、消费项目、消费金额等数据。运用统计软件对数据进行分析，使用聚类分析方法将学生按照消费行为模式进行分类，分析不同消费群体的特征；通过相关性分析研究月生活费与消费金额之间的关系，以及不同消费项目之间的关联。在这个案例中，涉及到了更复杂的数据收集、整理和分析过程，通过实践可以进一步掌握推断统计方法和多元统计分析方法的应用，提高解决实际问题的能力。

## 4.3. 阅读统计学相关文献

### (一) 了解不同领域的应用案例

阅读统计学相关文献，可以让研究新手了解统计学在各个领域的广泛应用。在医学领域，统计学被用于临床试验设计、疾病诊断准确性评估、药物疗效分析等。例如，阅读关于某种新药物疗效的临床试验文献，我们可以了解到如何设计试验方案，包括样本量的确定、对照组的设置、随机化分组的实施等。

在一项关于新型降压药物的临床试验中，研究者根据前期预试验数据和统计学公式，确定每组需要纳入 150 名高血压患者，以确保有足够的检验效能来检测药物与安慰剂之间的差异。通过随机化分组，将患者分为试验组(接受新型降压药物治疗)和对照组(接受安慰剂治疗)，以消除可能的混杂因素影响。在分析数据时，运用 t 检验比较两组患者治疗前后血压变化的差异，通过生存分析评估药物对心血管事件发生风险的长期影响等。这些实际案例展示了统计思想如何贯穿于医学研究的全过程，从研究设计到结果分析，为研究结论的可靠性提供保障。

在社会科学领域，统计学在市场调研、社会调查等方面有着广泛应用。例如，阅读关于消费者对某品牌电子产品满意度调查的文献，我们可以看到如何运用抽样调查方法选取具有代表性的样本，通过设计合理的问卷收集数据，并运用描述性统计分析消费者的基本特征(如年龄、性别、收入等分布情况)，使用因子分析提取影响消费者满意度的关键因素(如产品质量、功能、外观、售后服务等)，再通过回归分析研究这些因素与满意度之间的定量关系。在一篇关于某城市居民环保意识的社会调查文献中，研究者采用分层抽样方法，按照城市不同区域和居民收入水平进行分层，抽取了 1000 名居民进行问卷调查。通过数据分析发现，年龄、教育程度与环保意识呈正相关，高收入群体在环保行动上更为积极等结论，为城市环保政策制定提供了依据。这些案例让我们了解到统计学在社会科学研究中如何帮助研究者从大量复杂的数据中提炼出有价值的信息，为解决社会问题提供支持。

## (二) 学习统计方法的选择与运用技巧

文献中会详细阐述作者在研究中选择特定统计方法的原因和依据，这对于研究新手学习统计方法的选择技巧非常有帮助。例如，在一项比较三种不同教学方法对学生学习成绩影响的教育研究中，作者选择了方差分析方法。原因是研究中有一个分类自变量(教学方法，分为三种类型)和一个定量因变量(学生学习成绩)，且需要比较多个组之间的差异。方差分析能够同时检验多个总体均值是否相等，适合这种多组比较的研究设计。通过阅读此类文献，新手可以学习到根据研究问题的性质、数据类型以及研究设计来选择合适统计方法的思路。

同时，文献也会介绍统计方法在实际应用中的一些技巧和注意事项。在进行回归分析时，如何处理自变量之间的多重共线性问题。一些文献会提到使用方差膨胀因子(VIF)来检测多重共线性，如果 VIF 值大于某个阈值(如 5 或 10)，则表明存在严重的多重共线性。对于存在多重共线性的自变量，可以采用逐步回归法，让软件自动筛选出对因变量有显著影响且不存在严重共线性的自变量；或者使用岭回归等方法，通过对回归系数进行有偏估计来降低多重共线性的影响。研究新手通过阅读这些文献内容，可以了解到在实际运用统计方法时可能遇到的问题及相应的解决办法，避免在自己的研究中犯同样的错误。

## (三) 分析统计结果的解读方式

仔细研读文献中对统计结果的解读部分，能够学习到正确、全面解读统计结果的方法。在许多研究中，作者会首先对统计检验的显著性进行说明，例如在假设检验中，如果 p 值小于设定的显著性水平(如 0.05)，则拒绝原假设，认为存在显著差异或效应。但仅仅关注显著性是不够的，还需要结合效应量来评估结果的实际意义。效应量可以反映变量之间关系的强度或差异的大小。在比较两种治疗方法疗效的医学研究中，除了报告 t 检验的 p 值外，还会报告效应量指标，如 Cohen's d 值。如果 Cohen's d 值为 0.8，表示两种治疗方法之间存在较大的差异，具有实际临床意义；而如果 d 值仅为 0.2，则说明差异较小，即使 p 值显著，其实际应用价值也可能有限。要是比较两种以上治疗方法的疗效，则选择方差分析，除了报告 F 检验的 p 值外，还会报告效应量指标，如  $\eta_p^2$  的值，如果  $\eta_p^2$  的值在 0.14 以上，则认为是大效应量。

文献中还会对统计结果的不确定性进行讨论，如置信区间的解读。通过阅读这些内容，研究新手可以明白置信区间不仅给出了参数估计的范围，还反映了估计的精度和可靠性。在一项关于某地区居民平均收入的研究中，报告的 95% 置信区间为(4500 元， 5500 元)，这意味着有 95% 的把握认为该地区居民的

真实平均收入在这个区间内。如果置信区间较宽，说明估计的精度较低，可能需要进一步增加样本量或改进研究方法来提高估计的准确性。同时，文献中可能会对结果的局限性进行分析，例如由于样本的局限性导致结果可能无法推广到更广泛的人群，或者研究中存在某些未控制的混杂因素可能影响结果的解释等。学习这些内容可以让研究新手在自己解读统计结果时更加严谨和客观，避免过度解读或错误解读结果。

#### 4.4. 学习高级统计方法

随着研究的深入，需要学习更高级的统计方法，以应对复杂的研究问题和数据。多元统计分析方法是重要的进阶内容，包括因子分析、聚类分析、判别分析等。因子分析用于从多个相关变量中提取少数几个潜在的公共因子，简化数据结构。例如，在心理学研究中，通过对大量描述人格特征的量表数据进行因子分析，可以提取出外向性、神经质、开放性等几个主要的人格因子，便于对人格进行分类和研究。聚类分析则是根据变量之间的相似性或距离，将研究对象分为不同的类别。如在市场细分研究中，运用聚类分析可以将消费者按照购买习惯、消费偏好等特征分为不同的细分市场，为企业制定针对性的营销策略提供依据。判别分析用于根据已知类别的样本数据建立判别函数，对新的未知类别的样本进行分类判断。在医学诊断中，通过对患者的症状、体征、检查指标等数据进行判别分析，可以建立疾病诊断模型，辅助医生对患者的病情进行诊断。

非参数统计方法也是进阶学习的重要内容，适用于不满足参数统计方法假设条件(如正态分布、方差齐性等)的数据。常见的非参数检验方法包括 Wilcoxon 秩和检验、Kruskal-Wallis 检验、Spearman 等级相关分析等。在分析偏态分布的数据(如居民收入数据)时，使用 Wilcoxon 秩和检验比较两组数据的差异，比参数检验的 t 检验更合适。对于有序分类数据(如产品满意度等级)，Spearman 等级相关分析可以用来研究变量之间的关联程度，而不依赖于数据的分布形式。

贝叶斯统计方法是一种基于贝叶斯定理的统计推断方法，与传统的频率学派统计方法不同，它将先验信息与样本数据结合起来，得到后验分布，用于进行统计推断[9]。在机器学习、风险评估、医学诊断等领域有广泛应用。例如，在疾病诊断中，医生可以根据患者的症状和病史等先验信息，结合检查结果的样本数据，运用贝叶斯方法计算患者患病的后验概率，提高诊断的准确性。研究新手可以通过学习贝叶斯统计的基本原理、马尔可夫链蒙特卡洛(MCMC)方法等，逐步掌握这一高级统计方法。

此外，“机器学习”聚焦于从数据中自动学习规律以实现预测或决策，它与贝叶斯统计是现代数据科学领域紧密关联却又各有核心逻辑的两大分支。二者的深度融合(如贝叶斯机器学习)已成为解决复杂问题(如小样本学习、模型可解释性)的关键思路。机器学习的预测思想与传统统计学的推断思想并非对立关系，而是从目标、方法、数据适应度等维度对后者形成了关键补充，二者共同构成了现代数据分析的核心框架。要理解这种补充关系，需先明确两类思想的核心差异，再从具体应用场景中拆解补充的具体体现。

#### 4.5. 统计的常见误区与批判性思考

在数据分析、决策制定和科学的研究中，统计是核心工具之一，但由于对统计概念的误解、数据处理的疏漏或逻辑推导的偏差，人们常陷入各类统计误区。理解这些误区并建立批判性思考能力，是避免“数据说谎”、做出合理结论的关键。

一是忽略“数据分布”——用“通用模型”套“特殊数据”。许多统计方法(如均值、方差、线性回归)依赖“数据符合特定分布”(如正态分布)的前提，但实际中若数据分布异常，强行使用这些方法会导致结论偏差。例如，分析“居民财富分布”时，因财富数据呈“右偏分布”(少数人拥有大量财富，多数

人财富较少), 若用“均值”代表“平均财富”, 会远高于多数人的实际财富, 应改用“中位数”或“众数”。

二是统计显著性的误解——“ $p$  值崇拜”与“实际意义脱节”。在假设检验中, “ $p < 0.05$ ”(统计显著性)常被视为“结论成立”的唯一标准, 但过度依赖  $p$  值会陷入误区。 $p$  值  $\neq$  实际意义, 统计显著仅说明“观察到的结果不太可能是随机产生的”, 但不代表结果有“实际应用价值”。例如, 某减肥药实验显示, 服药组比对照组体重平均减少  $0.5 \text{ kg}$ ,  $p$  值 = 0.03(统计显著), 但  $0.5 \text{ kg}$  的减重效果在实际中几乎无意义——忽略了“效应量”(如均值差异、相关系数大小)这一关键指标。

三是结论的“过度推广”——超出数据的“适用范围”。将基于特定条件、特定群体的统计结论, 无限制推广到其他场景或群体, 导致“结论越界”。例如, 某研究基于“20~30岁城市女性”样本, 得出“每天喝 1 杯牛奶可提高骨密度”的结论, 却被推广为“所有人都应每天喝牛奶补钙”——忽略了“年龄、性别、地域、饮食习惯”等变量的差异(如乳糖不耐受人群喝牛奶可能不适)。

统计是“用数据说话”的工具, 但工具的有效性依赖使用者的“批判性思维”。常见的统计误区本质上是“对数据的懒惰解读”或“对结论的刻意美化”, 而批判性思考的核心是: 不轻易相信“数据表面的结论”, 而是追问“数据如何产生、指标如何选择、逻辑如何推导、结论如何适用”。最终, 统计的价值不在于“证明预设结论”, 而在于“通过数据接近真相”——这需要我们以“审慎质疑”为习惯, 以“多维度验证”为方法, 避免被“虚假数据”或“片面结论”误导。

## 5. 统计思想应用示例

假设检验是统计推断的主要方法, 其基本思想是概率性质的反证法[10]。为了检验虚无假设, 首先假定虚无假设为真。在虚无假设为真的前提下, 如果导致违反逻辑或违背常识和经验的不合理现象出现, 则表明“虚无假设为真”的假定是不成立, 也就是不能接受虚无假设。若没有导致不合理现象出现, 那就认为“虚无假设为真”的假定是正确的, 也就是接受虚无假设。

例如, 某小学三年级一个班进行智力测验, 班级人数  $n=49$  人, 平均智力为  $m=110$ 。已知小学三年级学生智力测验的总平均(常模)  $u_0=100$ , 标准差  $\sigma_0=16$ 。该班智力水平是否与常模水平有差异?

分析: 该班的平均智力是 110, 在数值上大于总平均 100, 但不能直接就据此认为该班智力水平与常模水平有显著差异。因为 110 大于 100, 109 也大于 100, ……, 100.001 亦是大于 100, 所以不能根据数值大于常模, 就得出与常模有显著差异的结论。

直接证明困难, 可以采用反证法。以  $u_1$  表示该班多次测量结果的总平均(也是该班的真实水平), 则检验的目的是要证实某班智力与已知总体不同。因而研究假设为  $H_1: u_1 \neq u_0$ 。反证法就需要建立的虚无假设为  $H_0: u_1 = u_0$ 。总体是一个正态分布, 置信区间水平  $\alpha = 0.05$ , 取两端各 0.025, 此为小概率事件区域。则抽取样本所得的平均数  $m$  可能比  $u_0$  大, 也可能比  $u_0$  小, 或者  $m = u_0$ , 只要没有落在两端的 0.025 的区域, 则  $m$  与  $u_0$  的差异可以被认为是抽样误差所致, 即  $m$  与  $u_0$  差异不显著。这时就接受需要假设  $H_0: u_1 = u_0$ ; 如果  $m$  落在两端的 0.025 区域, 则小概率事件出现, 而在一次抽样中小概率事件几乎不可能发生。此时就有充分理由拒绝虚无假设  $H_0: u_1 = u_0$ , 即拒绝  $H_0$ , 或者说  $m$  与  $u_0$  差异显著,  $m$  不是来自  $u_0$  这个总体。

具体统计推断过程:

### (一) 建立假设

$$H_0: u_1 = u_0;$$

$$H_1: u_1 \neq u_0$$

## (二) 计算 Z 统计量

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{49}} = \frac{16}{7} = 2.29$$

$$Z = (m - u_0) / SE = (110 - 100) / 2.29 = 4.37$$

## (三) 确定显著性水平

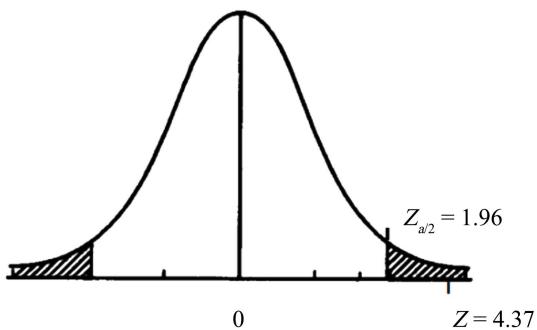
心理及教育常用的显著性水平设为 0.05，即  $\alpha = 0.05$ 。

## (四) 查表或计算 p 值

双侧检验， $\alpha = 0.05$ ，对应的临界值  $Z$  为 1.96，即  $Z$  值小于 -1.96 或大于 1.96 为拒绝  $H_0$  的区域。

## (五) 做出决策

如图 1 所示， $m = 110$  对应  $z = 4.37$ ，这个值大于置信区间的临界值  $z_{\alpha/2} = 1.96$ ，即落在小概率事件( $\alpha = 0.05$ )的范围。而小概率事件在一次抽样中几乎不可能发生，此时发生了，则由此可以做出拒绝原假设  $H_0$  的决策，也就是拒绝“该班智力分数与常模相同”，进而得出该班智力分数与常模有显著差异。需要说明的是，尽管根据  $m$  落在小概率事件区域，我们拒绝虚无假设  $H_0: u_1 = u_0$ ，但这种推断还是存在风险的，毕竟  $m$  还是有 0.05 的概率落在拒绝区域，因此，我们推断的风险概率是 0.05。



**Figure 1.** Distribution of intelligence tests

**图 1.** 智力测验的分布

## 6. 总结与展望

统计思想是做研究不可或缺的科学思维方式，它贯穿于研究设计、数据分析和结果解读的各个环节 [11]。通过学习统计学基础知识，包括数据类型、基本统计量、统计方法等，研究新手可以奠定构建统计思想的基础。进行实际数据分析操作，运用统计软件处理数据，从简单案例逐步深入，能够将理论知识转化为实践能力。阅读统计学相关文献，了解不同领域的应用案例，学习统计方法的选择与运用技巧以及结果解读方式，可以拓宽视野，提升统计应用水平。

同时，要警惕统计使用的常见误区，如错误选择统计方法、不理解统计方法的假设条件、数据缺失值和异常值处理不当、过度依赖  $p$  值、混淆相关性与因果关系、忽视结果的适用范围和局限性等，并采取相应的避免方法[12]-[14]。构建统计思想是一个循序渐进的过程，需要从基础到高级，不断学习和实践，结合具体研究领域深入应用，培养批判性思维和创新意识。

随着数据科学的快速发展，统计学在各个领域的应用将更加广泛和深入[15]。构建统计思想，不仅要掌握传统的统计方法，还要关注统计学与其他学科的交叉融合，如与计算机科学、生物学、社会学等学科的结合，形成新的研究方法和技术。例如，统计方法与机器学习的结合，能够更好地处理大数据和复杂数据，提高数据分析的效率和准确性。未来的研究将更加注重数据的质量和伦理问题，研究者在构建

统计思想时，要树立数据质量意识和伦理意识，确保数据的真实性、可靠性和合法性[16]。在数据收集过程中，要遵守研究伦理规范，保护研究对象的隐私和权益；在数据分析和结果报告中，要客观、公正地呈现研究成果，不篡改数据、不隐瞒研究局限性。此外，随着开源软件和在线学习资源的普及，获取统计学习资料和工具的渠道更加便捷。

## 基金项目

本文为2024年河南省研究生教育改革与质量提升工程项目研究生优质课“高级统计”(YJS2024KC33)成果。

## 参考文献

- [1] 司艳丽. 现代统计技术在大数据处理中的应用[J]. 销售与管理, 2024(18): 9-11.
- [2] 杨俐敏, 常湘珍, 张方晨, 等. 医学论文来稿中统计方法的误用及分析[J]. 韶关学院学报, 2014, 35(10): 76-78.
- [3] 全国大学生统计建模大赛组织委员会, 国家统计局统计教育培训中心. 数据新动能的统计测度研究[M]. 北京: 中国统计出版社, 2022.
- [4] 李强. 如何培育学生数据分析素养[N]. 中国教育报. 2025-09-22(005).
- [5] 李金昌. 论统计功能[J]. 学术月刊, 2012(10): 82-90.
- [6] 史周华, 何雁. 中医药统计学与软件应用[M]. 北京: 中国中医药出版社, 2017.
- [7] 周华军. 用样本估计总体 为未来谋划发展[J]. 初中生世界, 2020(43): 44-46.
- [8] 王霞, 沈静, 占晓军. 概率论与数理统计[M]. 重庆: 重庆大学出版社, 2022.
- [9] 约翰 K. 克鲁斯克. 贝叶斯统计方法[M]. 北京: 机械工业出版社, 2015.
- [10] 娄银霞. 关于假设检验的一种合理推导[J]. 电子设计工程, 2012, 20(20): 33-35.
- [11] 马秀麟. 科研统计思维与方法: SPSS 实战[M]. 北京: 人民邮电出版社, 2024.
- [12] “大数据中的统计方法”课题组, 马双鸽. 大数据时代统计学发展的若干问题[J]. 统计研究, 2017, 34(1): 5-11.
- [13] 史芸萍, 刘军廷, 么鸿雁, 等. 统计学方法的合理选择和常见误区[J]. 中国防痨杂志, 2016, 38(5): 346-348.
- [14] 张翠凤. 统计分析应避免的几个误区[J]. 中国统计, 2010(5): 50-51.
- [15] 梁龙龙. 概率论与数理统计在大数据分析中的应用[J]. 科技视界, 2021(10): 34-35.
- [16] 李正风, 王硕. 数字素养、数据权利与数字伦理[J]. 科普研究, 2022, 17(6): 8-14+108.