

# 假设检验两类错误的动态可视化教学研究

刘素兵, 陈春梅

火箭军工程大学基础部, 陕西 西安

收稿日期: 2025年8月12日; 录用日期: 2025年10月7日; 发布日期: 2025年10月16日

## 摘要

假设检验中的两类错误是概率论与数理统计教学的难点。针对学生普遍存在的概念混淆、计算障碍及关系理解困难等问题, 文章基于正态分布构建动态可视化模型, 系统解析了样本量固定时两类错误概率此消彼长的动态平衡关系, 并量化展示了第一类错误概率、第二类错误概率与样本量的内在联系。通过理论机制分析、数值模拟与几何解释三位一体的研究方法, 有效化解了知识抽象带来的认知障碍, 助力学生深入理解核心内容。

## 关键词

假设检验, 正态总体, 两类错误, 动态可视化

# Research on Dynamic Visualization Teaching of Hypothesis Testing for Two Types of Errors

Subing Liu, Chunmei Chen

Basic Course Department, Rocket Force University of Engineering, Xi'an Shaanxi

Received: August 12, 2025; accepted: October 7, 2025; published: October 16, 2025

## Abstract

The two types of errors in hypothesis testing represent a challenging topic in the teaching of probability theory and mathematical statistics. To address students' widespread difficulties, such as conceptual confusion, computational barriers, and struggles in understanding relationships, this paper constructs a dynamic visualization model based on the normal distribution. This model systematically analyzes the dynamic balance relationship where the probabilities of the two types of errors vary inversely when the sample size is fixed. It further quantifies the intrinsic connections among

the probability of Type I error ( $\alpha$ ), the probability of Type II error ( $\beta$ ), and the sample size ( $n$ ). Through a tripartite research approach integrating theoretical mechanism analysis, numerical simulation, and geometric interpretation, the study effectively overcomes cognitive obstacles arising from the abstract nature of the topic, thereby helping students gain a deeper understanding of the core concepts.

## Keywords

Hypothesis Testing, Normal Population, Two Types of Errors, Dynamic Visualization

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

参数的假设检验是数理统计推断的重要内容之一, 在许多领域具有广泛的应用, 特别是在科学实验和质量检验中[1]-[4]。但是, 学生在学习假设检验时, 特别是关于假设检验中的两类错误时, 普遍存在以下问题, 这些核心问题既是认知难点, 也是教学中的关键挑战, 具体体现为以下几个方面: 一是概念混淆, 对两类错误理解不到位; 二是关系理解困难, 无法深入理解在一定条件下, 两者此消彼长的关系; 三是计算障碍, 无法从概率角度量化。笔者针对上述问题, 对假设检验中的两类错误进行系统分析和梳理, 以帮助学生深入理解知识, 解决学习中的困惑。

## 2. 假设检验基本原理

假设检验是一种统计推断方法, 用于确定收集到的数据是否充分支持某个特定的假设。假设检验中最常见的是参数检验, 即根据收集到的数据对总体分布含有的未知参数或未知参数的函数进行判断, 给出事先对它们做出的假设正确与否的决定[5]。

为了推断总体的某些性质, 需要先提出关于总体的假设, 然后从总体中随机抽取样本, 根据样本的检验和分析, 对提出的假设作出是接受还是拒绝的判断。

参数的假设检验具体为[1]:

设  $(x_1, x_2, \dots, x_n)$  为来自某一个参数分布族  $\{F(x, \theta), \theta \in \Theta\}$  的样本,  $\Theta$  为参数空间  $\Omega$ ,  $\Theta_0, \Theta_1 \subset \Theta$ , 且  $\Theta_0 \neq \emptyset$ ,  $\Theta_0 \cap \Theta_1 = \emptyset$ , 提出的一对假设为:

$$H_0: \theta \in \Theta_0, H_1: \theta \in \Theta_1$$

$H_0$  称为原假设,  $H_1$  称为备择假设。

对检验问题  $(H_0, H_1)$ , 构造样本空间  $\Omega$  的集合分割:  $W$  和  $\bar{W} = \Omega - W$ ,  $W$  称为检验的拒绝域,  $\bar{W}$  称为检验的接受域; 确定检验规则, 若检验统计量的值落入拒绝域, 则拒绝原假设  $H_0$ ; 否则, 接受原假设  $H_0$ 。因此, 检验就是将定义空间划分为两个不相互的部分  $W$  和  $\bar{W}$ , 通过一个检验的统计量完成。当拒绝域确定了, 检验法则也就确定了。假设检验所依据的是小概率事件原理。

## 3. 假设检验两类错误的认知难点与关系剖析

### 3.1. 两类错误的定义与概率计算

在统计检验中, 作出决策结论的依据是“小概率事件在一次试验中几乎不可能发生”的实际推断原

理, 但这并不等于不发生[6]; 又因依据样本而作出拒绝或接受原假设的推断, 而抽样具有随机性和不确定性, 因此, 对于给出的拒绝域  $W$ , 我们做出的判断不可能绝对正确, 它可能会犯两类错误。

### 3.1.1. 第一类错误

若  $H_0$  为真, 但由于随机性使得样本观测值落在拒绝域  $W$  中, 我们拒绝了原假设  $H_0$ , 从而犯了“弃真”错误, 这种错误称为第一类错误, 其发生的概率称为犯第一类错误的概率, 或称为弃真概率, 通常记为  $\alpha$ , 即

$$\alpha(\theta) = P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\} = P_\theta(X \in W), \quad \theta \in \Theta_0$$

其中,  $X = (x_1, x_2, \dots, x_n)$  表示样本。

### 3.1.2. 第二类错误

若  $H_0$  不真(原假设本来不成立, 即  $H_1$  为真), 但由于随机性使样本观测值落在接受域  $\bar{W}$  中, 我们接受了原假设  $H_0$ , 从而犯了“取伪”错误, 其发生的概率称为犯第二类错误的概率, 或称为取伪概率, 通常记为  $\beta$ , 即

$$\beta(\theta) = P\{\text{接受 } H_0 | H_1 \text{ 为真}\} = P_\theta(X \in \bar{W}), \quad \theta \in \Theta_1$$

综上, 明确起见, 两类错误以表格形式给出[7] [8], 如表 1 所示。

Table 1. Two types of errors in hypothesis testing

表 1. 检验的两类错误

| 判断       | 情况       |          |
|----------|----------|----------|
|          | $H_0$ 成立 | $H_1$ 成立 |
| 拒绝 $H_0$ | 犯第一类错误   | 判断正确     |
| 接受 $H_0$ | 判断正确     | 犯第二类错误   |

### 3.1.3. 两类错误的计算

以单正态总体均值的右尾  $Z$  检验为例, 对两类错误的关系进行解释。

例 1 设样本  $X_1, X_2, \dots, X_n$  来自正态总体  $X \sim N(\mu, \sigma^2)$ , 其中  $\sigma^2$  已知, 检验假设  $H_0: \mu \leq \mu_0$ ,  $H_1: \mu > \mu_0$ , 显著性水平为  $\alpha$ , 样本量为  $n$ , 检验统计量为  $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ , 且  $Z \sim N(0, 1)$ , 拒绝域为  $W = \{z \geq z_\alpha\}$ , 计算犯第一类错误的概率  $\alpha$  和犯第二类错误的概率  $\beta$ 。

解: 由上述定义, 知道犯第一类错误的概率  $\alpha$  为显著性水平。下面计算犯第二类错误的概率  $\beta$ 。

第二类错误发生在原假设  $H_0$  为假(即备择假设  $H_1: \mu > \mu_0$  为真)时, 但我们却接受了原假设。接受原假设的区域  $\bar{W} = \{z < z_\alpha\}$ , 故对于给定的真实均值  $\mu$  (且  $\mu > \mu_0$ ), 有

$$\begin{aligned} \beta(\mu) &= P\{\text{接受 } H_0 | H_1 \text{ 为真}\} \\ &= P\{Z < z_\alpha | \mu > \mu_0\} \\ &= P\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < z_\alpha | \mu > \mu_0\right\} \end{aligned} \tag{1}$$

将式(1)中  $\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$  的分子拆为两部分, 变形为:

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{(\bar{X} - \mu) + (\mu - \mu_0)}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$$

$$\text{于是, } \beta(\mu) = P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma/\sqrt{n}} < z_\alpha \mid \mu > \mu_0\right\} = P\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}} \mid \mu > \mu_0\right\}$$

而  $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$ , 有

$$\beta(\mu) = \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right) \quad (2)$$

其中,  $\Phi(\cdot)$  为标准正态分布的分布函数, 临界值  $z_\alpha$  满足  $\Phi(z_\alpha) = 1 - \alpha$ 。

### 3.2. $\alpha$ 与 $\beta$ 关系的理论机制分析与动态可视化验证

#### 3.2.1. 理论机制分析

由式(2)知, 第二类错误概率  $\beta$  受显著性水平  $\alpha$ 、真实均值  $\mu$ 、样本量  $n$  的共同影响。当样本量  $n$  固定时, 进一步分析  $\alpha$  和  $\beta$  的关系:

由式(1)知, 当其它不变时,  $\alpha$  越小,  $z_\alpha$  越大( $z_\alpha$  为标准正态分布的上  $\alpha$  分位数), 于是  $z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}$  增大, 由于分布函数  $\Phi(\cdot)$  单调递增, 从而  $\beta$  就越大。反之, 若  $\alpha$  越大,  $\beta$  就越小。

为具体说明  $\alpha$  与  $\beta$  在固定  $n$  下的此消彼长关系, 给出数值示例。取  $\mu_0 = 100$ ,  $\sigma = 15$ ,  $n = 25$ ,  $\mu = 105$ , 当  $\alpha$  分别取 0.01、0.05、0.1 时, 依据式(2)计算对应  $\beta$  的值, 结果如表 2 所示。

**Table 2.** Two types of errors in hypothesis testing

**表 2.**  $\alpha$  分别取 0.01、0.05、0.1 时得到的  $\beta$  值

| $\alpha$ 值 | $z_\alpha$ | $\beta$ 计算                            | $\beta$ 值 |
|------------|------------|---------------------------------------|-----------|
| 0.01       | 2.33       | $\Phi\left(2.33 - \frac{5}{3}\right)$ | 0.75      |
| 0.05       | 1.64       | $\Phi\left(2.33 - \frac{5}{3}\right)$ | 0.49      |
| 0.10       | 1.28       | $\Phi\left(2.33 - \frac{5}{3}\right)$ | 0.35      |

由表 2 可清晰地看到, 在样本量  $n$  及其他参数固定的条件下, 随着显著性水平  $\alpha$  的增大, 犯第二类错误的概率  $\beta$  显著减小。

上述数值示例说明, 在样本量  $n$  固定且其他条件不变的情况下, 期望同时将  $\alpha$  和  $\beta$  都控制得很小是不可能的。

#### 3.2.2. 动态可视化验证

如果增大样本量会如何影响  $\alpha$  和  $\beta$  呢? 为了更清晰地展示显著性水平  $\alpha$ 、第二类错误概率  $\beta$  和样本量  $n$  三者之间的关系, 利用 MATLAB 绘制了三维动态可视化图进行直观呈现。为此, 编写了交互式 MATLAB 程序, 创建了一个图形用户界面。该界面包含两个滑块: 一个用于控制显著性水平  $\alpha$ , 另一个用于控制样本量  $n$ 。学生可以通过动态调节这  $\alpha$  和  $n$  参数, 实时观察参数变化如何影响第二类错误概率  $\beta$ , 从而深入理解三者间的动态关联。如固定  $\alpha$ , 滑动  $n$  滑块, 观察样本量增大如何显著降低  $\beta$ ; 如固

定  $n$ , 调节  $\alpha$  滑块, 理解显著性水平变小, 如何导致  $\beta$  升高。图 1 展示了该可视化工具的一个截面, 动态呈现了  $\alpha$ 、 $n$ 、 $\beta$  三者构成的关系图, 这种直观、动态的关联呈现, 有效克服了传统静态讲解的局限, 使学生能够主动探索并深刻理解参数间的复杂相互作用。

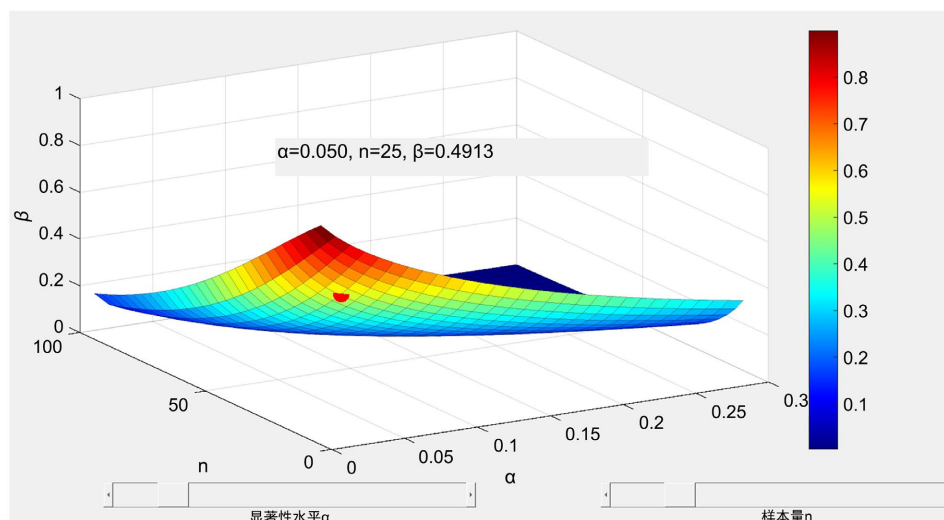


Figure 1. Teaching interface for visualizing  $\alpha$ - $n$ - $\beta$  dynamics

图 1.  $\alpha$ - $n$ - $\beta$  关系动态可视化界面示意图

### 3.2.3. 关系总结

综上, 对于  $\alpha$  和  $\beta$  的关系, 可以明确以下两点:

(1) 固定样本量  $n$  时,  $\alpha$  和  $\beta$  的关系就像跷跷板, 此消彼长,  $\alpha$  小,  $\beta$  就大;  $\beta$  小,  $\alpha$  就大, 不可能使得  $\alpha$  和  $\beta$  同时充分小[9]。

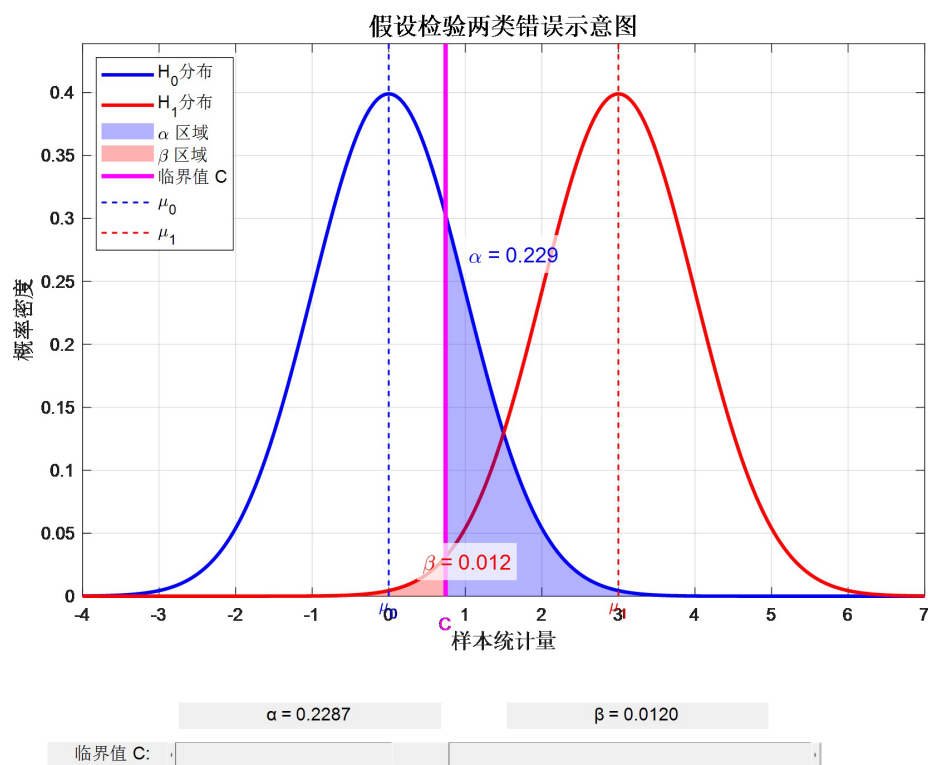
(2) 增大样本量  $n$  的作用: 增加样本量是协调降低  $\alpha$  和  $\beta$  矛盾的关键途径。增大  $n$  能显著降低  $\beta$ , 或者在保持  $\alpha$  不变的前提下更有效地降低  $\beta$ 。然而, 增大  $n$  意味着更高的数据收集成本, 在实际应用中, 需权衡成本与对两类错误控制精度的要求, 有些情况是不现实的。因此, 统计实践中通常预先设定可接受的显著性水平  $\alpha$ , 然后通过功效分析确定能够将  $\beta$  控制在目标值, 或达到期望检验效能, 所需的最小样本量  $n$ 。

## 3.3. 基于正态分布的动态可视化分析

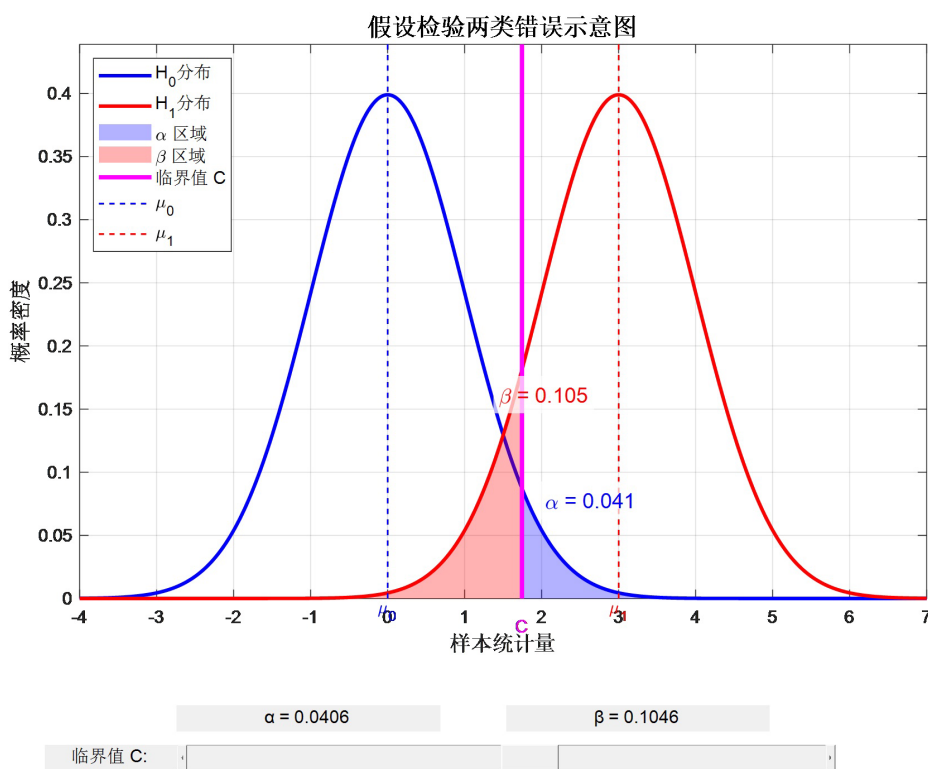
以正态分布的右尾检验为例, 对假设检验问题中两类错误及有关问题, 进行几何上的说明和解释, 图形展示两类错误的概率的变化关系, 使抽象概念具象化。

### 3.3.1. 临界值移动对两类错误的影响(固定样本量 $n$ )

设样本  $X_1, X_2, \dots, X_n$  为来自正态总体  $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  已知, 检验问题:  $H_0: \mu \leq \mu_0$ ,  $H_1: \mu > \mu_0$ , 样本均值  $\bar{X}$  在  $H_0$  下服从  $N\left(\mu_0, \frac{\sigma^2}{n}\right)$ , 在  $H_1$  下服从  $N\left(\mu_1, \frac{\sigma^2}{n}\right)$ , 即  $H_0$  成立时,  $\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$ ; 当  $H_1$  成立时,  $\bar{X} \sim N\left(\mu_1, \frac{\sigma^2}{n}\right)$ 。如图 2(a) 为初始状态  $\bar{X}$  的两个分布, 其中蓝色曲线表示在  $H_0$  下, 以  $\mu_0$  为中心的  $\bar{X}$  的分布曲线, 红色曲线表示在  $H_1$  下, 以  $\mu_1$  为中心的  $\bar{X}$  的分布曲线。 $H_0$  成立时,  $\alpha$  为  $\bar{X} > C_1$  的概率, 是  $H_0$  曲线下  $C_1$  右侧的尾部面积;  $\beta$  为  $\bar{X} \leq C_1$  的概率, 是  $H_1$  曲线下  $C_1$  左侧的尾部面积。



(a)



(b)

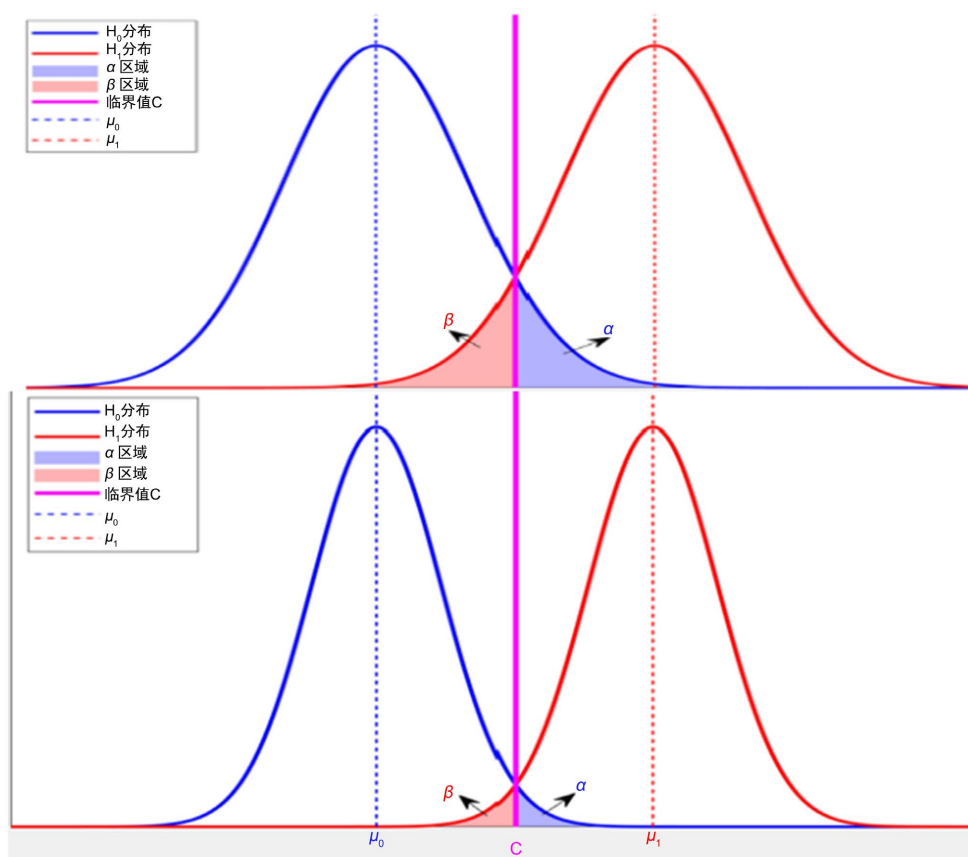
Figure 2. Critical value slider demonstration diagram

图 2. 临界值滑动演示图

运用 matlab 编程, 给出交互模型, 进行临界值滑动演示, 如图 2 所示。此时, 临界值  $C$  的移动直接影响两类错误的概率。临界值  $C$  右移, 则降低  $\alpha$ , 即减小  $H_0$  曲线下  $C$  右侧的尾部面积, 但增大  $\beta$ , 即增大  $H_1$  曲线在  $C$  左侧的尾部面积。反之,  $C$  左移降低  $\beta$  而提高了  $\alpha$ 。因此,  $\alpha$  和  $\beta$  呈现此消彼长的关系, 这是源于固定分布形态下, 改变  $C$  只是移动了划分接受域和拒绝域的边界, 重新分配了曲线重叠区域下两类错误的面积。也就是说, 调节临界值会直接改变定义两类错误概率的区域范围, 从而引起  $\alpha$  与  $\beta$  此消彼长的变化趋势; 需要强调的是, 这种变化并非严格的数学对称关系, 即两者下降或上升的幅度未必相同, 而是在给定样本量, 即固定数据信息精度这一约束条件下, 对弃真风险  $\alpha$  与取伪风险  $\beta$  进行权衡取舍后所接受的动态平衡状态。

### 3.3.2. 样本量增大后, 两类错误同时降低的几何解释

当增大样本量  $n$  后, 会减小  $\bar{X}$  的方差  $\frac{\sigma^2}{n}$ , 也就是意味着分布曲线会变得更加的瘦高和陡峭, 并集中于各自均值  $\mu_0$  和  $\mu_1$  附近, 如图 3 所示。



**Figure 3.** Visualization of simultaneous reduction in two types of errors with larger sample sizes  
**图 3.** 样本量增大后, 两类错误同时降低示意图

分布曲线的收紧显著减小了它们之间的重叠区域。在固定临界值  $C$  的情况下, 它们之间的重叠区域会显著缩小, 这正是能够同时降低  $\alpha$  和  $\beta$  的关键。因此, 在固定临界值  $C$  的情况下, 增大样本量  $n$  可以同时降低犯第一类错误  $\alpha' < \alpha$  和第二类错误  $\beta' < \beta$  的概率。图形直观地展示了, 增大样本量如何通过降低数据离散度协同降低两类错误概率。

## 4. 教学实验设计与效果分析

为客观评估基于正态分布动态可视化的教学方法,开展了两个班的对比教学实验。一个班为对照组,采用传统教学;另一个为实验组,在传统教学的基础上,引入动态可视化教学,两个班前序课程成绩无显著性差异。实验围绕“假设检验中的两类错误”这一知识点展开,为全面评估教学效果,收集并分析学生概念理解试卷、学习兴趣和信心等问卷反馈数据,用以检验方法是否有效。概念理解试卷,包含选择题、判断题和简答题,重点考察学生对两类错误定义、关系及影响因素的理解;对于学习兴趣、信心及对该教学方法的接受度和满意度等,采用 Likert 五级量表问卷,用 SPSS 软件分析数据。

测试成绩进行统计分析,结果如表 3 所示。实验组的平均得分显著高于对照组,表明实验组学生对两类错误的理解等概念性问题上明显优于对照组。

Table 3. Statistical analysis of test scores

表 3. 测试成绩的统计分析

| 组别  | 人数 | 平均分   | T 值   | P 值  |
|-----|----|-------|-------|------|
| 实验组 | 32 | 85.15 | 11.45 | 0.00 |
| 对照组 | 35 | 76.32 |       |      |

同时,学习兴趣问卷调查显示,对假设检验内容的学习兴趣和信心打分,实验组学生平均分为 4.32/5;对照组为 3.56/5,实验组显著高于对照组;经独立样本 t 检验,两组差异具有统计学意义( $t(65) = 5.42, p < 0.001$ )。这表明,基于动态可视化的教学方法有效激发了学生的学习动机,显著增强了其攻克难点的认知信心,这为深化概念的理解提供了积极的情感态度基础。

## 5. 结语

本研究对假设检验中两类错误的认知难点,采用了基于正态分布假设的动态可视化教学手段,通过理论解析、数值模拟与几何解释三者的融合,系统揭示了样本量固定时  $\alpha$  与  $\beta$  此消彼长的动态平衡机制,以及增大样本量在协调两类错误矛盾中的核心调控作用;所开发的交互式 MATLAB 工具,支持学生动态调节  $\alpha$  与  $n$ ,并实时观测  $\beta$  变化规律,这种交互式演示,通过参数实时交互,动态呈现  $\alpha$ 、 $\beta$ ,将抽象概率计算转化为分布曲线重叠区域的可视化表达,从而具象化概念,有效化解了传统教学中概念抽象与关系理解困难等瓶颈问题,从而深化了学生对两者关系的认知。未来研究,可进一步拓展非正态分布检验的广义可视化框架,用可视化桥梁连接理论概率空间与学生认知空间,进一步提升学生的理解和统计推断能力。

## 参考文献

- [1] 赵清波, 刘烁, 吴克坚, 等. 应用假设检验需特别注意的几个问题[J]. 中华医学图书情报杂志, 2017, 26(5): 53-55+71.
- [2] 王嘉澜. 数理统计中关于假设检验的几个要点问题[J]. 高等理科教育, 2005(1): 79-81.
- [3] 盛骤, 谢世千, 潘承毅. 概率论与数理统计[M]. 第 5 版. 北京: 高等教育出版社, 2019.
- [4] 李启寨, 孟珍. 概率论与数理统计[M]. 北京: 高等教育出版社, 2023: 285-286.
- [5] 茆诗松, 程依明, 濮晓龙. 概率论与数理统计教程[M]. 第三版. 北京: 高等教育出版社, 2019.
- [6] 胡政发, 肖海霞. 应用数理统计与随机过程[M]. 北京: 电子工业出版社, 2021: 104-106.
- [7] 天津大学数学系. 概率论与数理统计讲义[M]. 北京: 高等教育出版社, 2012: 200-201.
- [8] 贾俊平, 何晓群, 金勇进. 统计学[M]. 第 8 版. 北京: 中国人民大学出版社, 2023: 136-137.
- [9] 刘凤霞. 假设检验中两类错误的几何解释[J]. 渤海大学学报(自然科学版), 2007(3): 251-253.