

# 03 统计学中的危险方程

Speaker: 许文立

[wlxu@cityu.edu.mo](mailto:wlxu@cityu.edu.mo)

August-November, 2025

Faculty of Finance, City University of Macau



# 目录

---

引言：危险方程的启示

01

---

教育数据中的误导现象

02

---

棣莫弗方程的统计学意义

03

---

假设检验与P值

04

---

实际应用案例

05

---

总结与展望

06

# Part. 01

引言：危险方程的启示

---

CHAPTER

# 危险方程的 引言

## 1 霍华德·韦纳（Howard Wainer）的观点

霍华德·韦纳指出，某些方程在我们不了解时会释放其危险性，尤其是莫瓦尔方程。他强调，这些方程隐藏着深刻的统计学原理，若被忽视，可能导致错误的决策和误导性的结论。

## 2 莫瓦尔方程（Moivre's equation）的重要性

$$SE = \frac{\sigma}{\sqrt{n}}$$

莫瓦尔方程是统计学中的基础方程，它揭示了样本均值的标准误差SE与样本量n之间的关系。这一关系对于理解数据的不确定性至关重要，尤其是在小样本情况下。

## 3 引出后续内容

本讲将深入探讨莫瓦尔方程及其在统计学中的应用，通过实际案例分析，揭示其在数据分析中的重要性和潜在的危险性，帮助大家更好地理解和应用统计学方法。



## 莫瓦尔方程的数学形式

莫瓦尔方程的数学形式为

$$SE = \frac{\sigma}{\sqrt{n}}$$

- SE是样本均值的标准误差， $\sigma$ 是总体标准差， $n$ 是样本容量。
- 该方程表明，样本均值的标准误差与样本量的平方根成反比。



## 方程的统计学意义

莫瓦尔方程在统计学中具有基础性作用。它帮助我们理解样本均值的变异性，尤其是在样本量较小时，样本均值的不确定性较大，可能导致极端结果。通过该方程，我们可以更好地评估统计估计的精确性。



# Part. 02

教育数据中的误导现象

---

CHAPTER

# 巴西ENEM成绩分析

## ENEM成绩数据展示

巴西ENEM考试是全国性的教育评估工具，通过分析ENEM成绩数据，我们发现表现最优的学校往往具有学生数量较少的特点。

### 前10名学校的特点

通过数据表格展示前10名学校的详细信息，发现这些学校的学生数量普遍较少。这种现象可能引发误导性结论，例如认为小规模学校模式是提高成绩的关键。

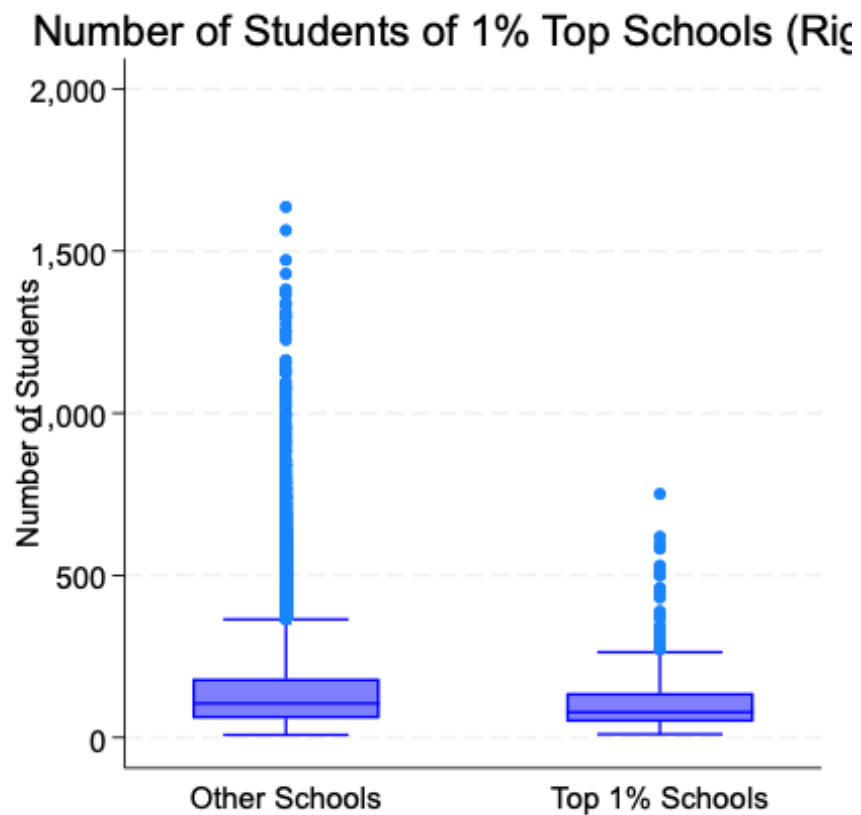
year	school~d	number~s	avg_sc~e
2007	33062633	68	82.97
2007	33065403	172	82.04
2005	33062633	59	81.89
2005	33065403	177	81.66
2007	29342880	43	80.32
2007	33152314	14	79.82
2007	33065250	80	79.67
2007	22025740	144	79.52
2007	31311723	222	79.41
2007	33087679	210	79.38

数据和代码: <https://wenzhe-huang.github.io/python-causality-handbook-zh/03-Stats-Review-The-Most-Dangerous-Equation.html>

# 巴西ENEM成绩分析

## ENEM成绩数据展示

巴西ENEM考试是全国性的教育评估工具，通过分析ENEM成绩数据，我们发现表现最优的学校往往具有学生数量较少的特点。



数据和代码: <https://wenzhe-huang.github.io/python-causality-handbook-zh/03-Stats-Review-The-Most-Dangerous-Equation.html>



# 巴西ENEM成绩分析

## 结论

一个自然而然的结论是，小规模学校能带来更高的学业表现

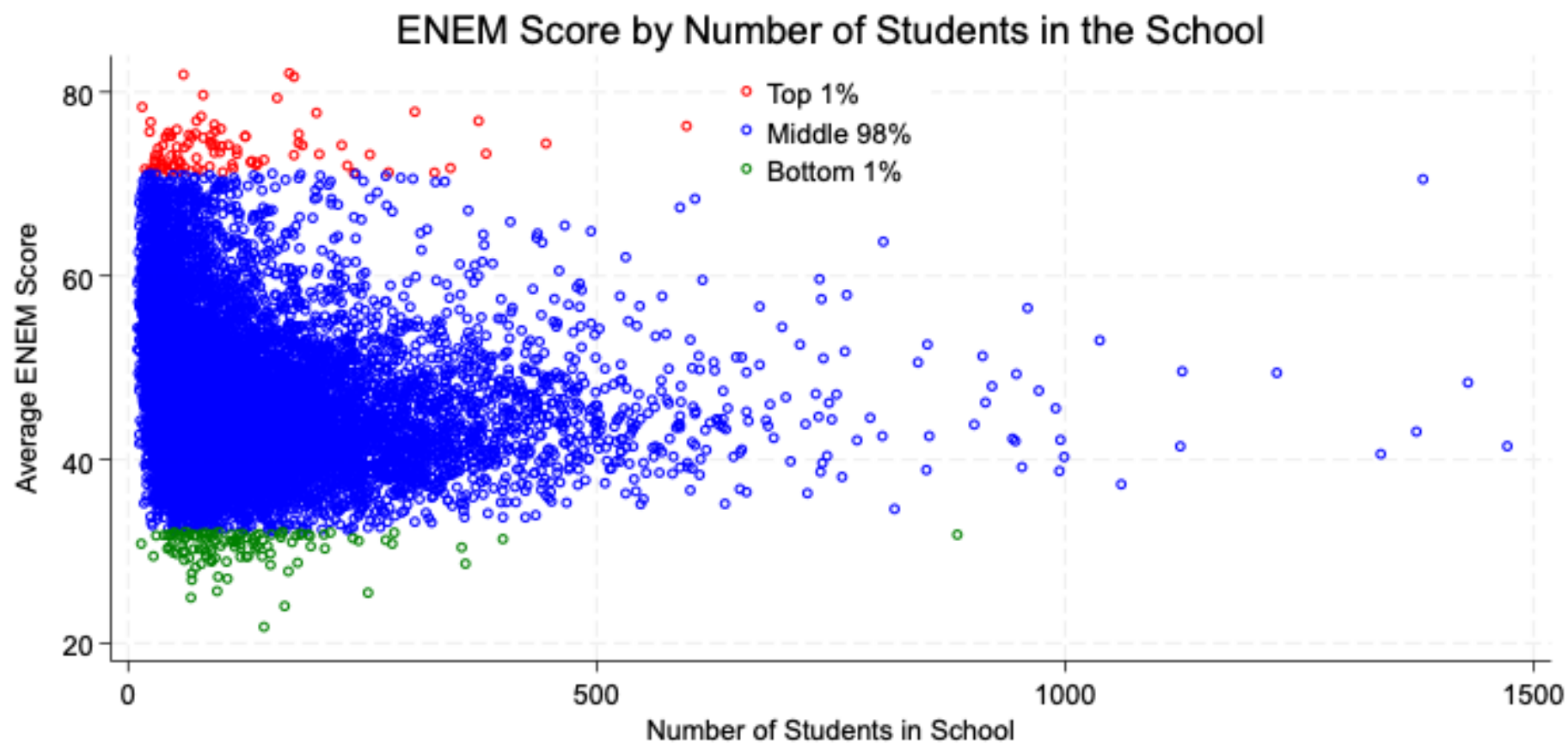
## 直觉

- 因为我们相信每位教师对应的学生越少，教师就能给予每个学生更多关注；
- 20 世纪 90 年代，提倡缩小学校规模的做法盛行一时。众多慈善组织和政府机构资助拆分大型学校，因为小规模学校的学生在考试成绩优异群体中占比过高。

## 成绩较差的1%学校

人们却忽视了去关注成绩最差的 1%学校。当我们这样做时，瞧！这些学校的学生人数同样寥寥无几！

# 巴西ENEM成绩分析





## 前1%学校的分布

通过图表展示前1%顶尖学校的分布情况，发现这些学校的学生数量普遍较少。这种现象可能引发错误决策，例如盲目推广小规模学校模式。

## 小样本量的影响

小样本量可能导致极端成绩的出现，这些成绩可能只是随机波动的结果，而非真正的教育模式优势。

## 避免错误决策

在分析教育数据时，需要警惕小样本量带来的误导性结论，避免因样本量不足而做出错误的决策。

# 前1%学校的特征

# 成绩最差学校的对比

Moving on to the next point

## 成绩最差学校的特征

展示成绩最差的1%学校的特征，发现这些学校的学生数量同样较少。通过图表对比顶尖学校和最差学校的分布，揭示这种现象的随机性本质。



## 随机性本质的启示

这种对比揭示了小样本量学校的随机性本质，无论是成绩最好还是最差，都可能是样本量不足导致的极端结果，而非真正的教育质量差异。



# 巴西ENEM成绩分析

## 结论

上述现象正是莫瓦方程（Moivre's equation）所预期的现象：随着学生人数的增加，平均成绩变得越来越精确

## 误导性结论的产生

由于样本量较小，这些学校的优异成绩可能只是随机波动的结果，而非真正的教育模式优势。这种误导性结论可能影响教育政策的制定。

## 数据的随机性本质

这种现象揭示了数据的随机性本质，小样本量可能导致极端结果，而这些结果并不一定具有普遍性和可推广性。

## 启示

莫瓦方程揭示了一个关于现实中信息与数据记录的基本事实：数据永远是不精确的。那么问题就变成：它到底有多不精确？

# Part. 03

莫瓦方程的统计学意义

---

CHAPTER



# 样本量与不确定性

## 样本量与不确定性的关系

莫瓦方程揭示了样本量与不确定性之间的关系。随着样本量的增加，样本均值的不确定性逐渐减小。小样本量可能导致极端结果，而大样本量则能更准确地反映总体特征。



*Moving on to the next point*





# 标准误差的计算

## 标准误差的定义

标准误差是样本均值的标准差，用于衡量样本均值的变异性。公式

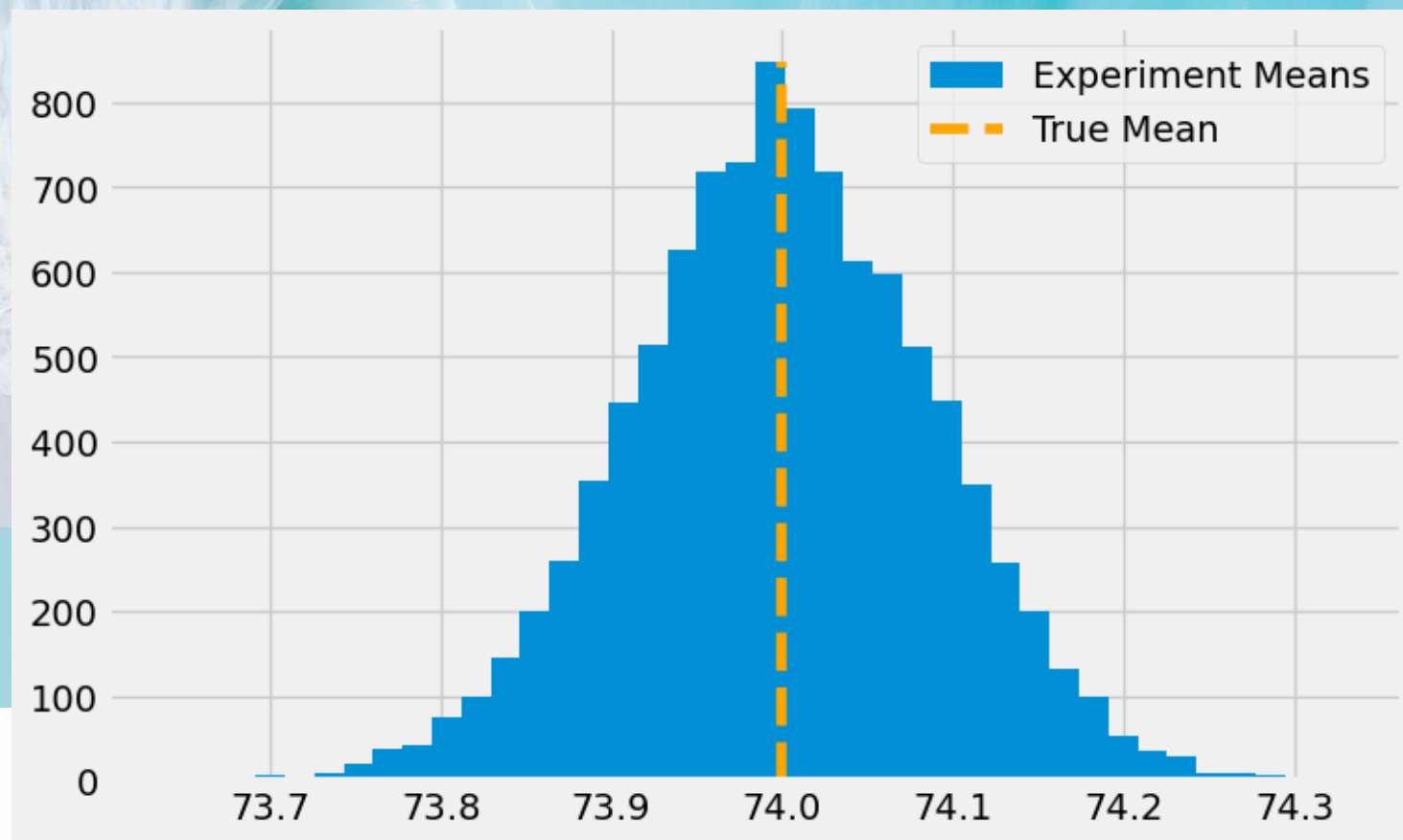
$$\hat{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

其中 ( $\sigma$ ) 是总体标准差，(N) 是样本容量。



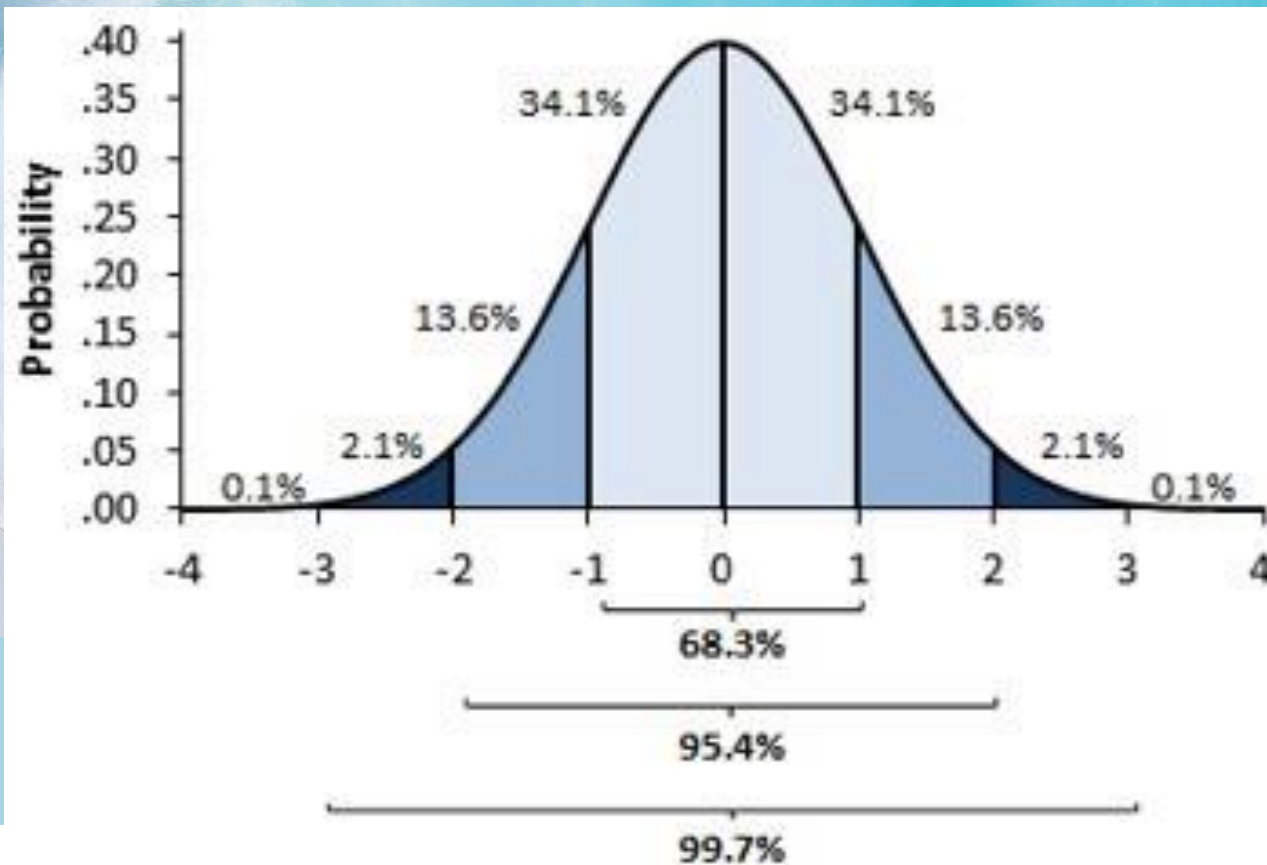


## 置信区间的构建



我们永远无法确定某次实验得到的样本均值就恰好等于真实均值

# 置信区间的构建



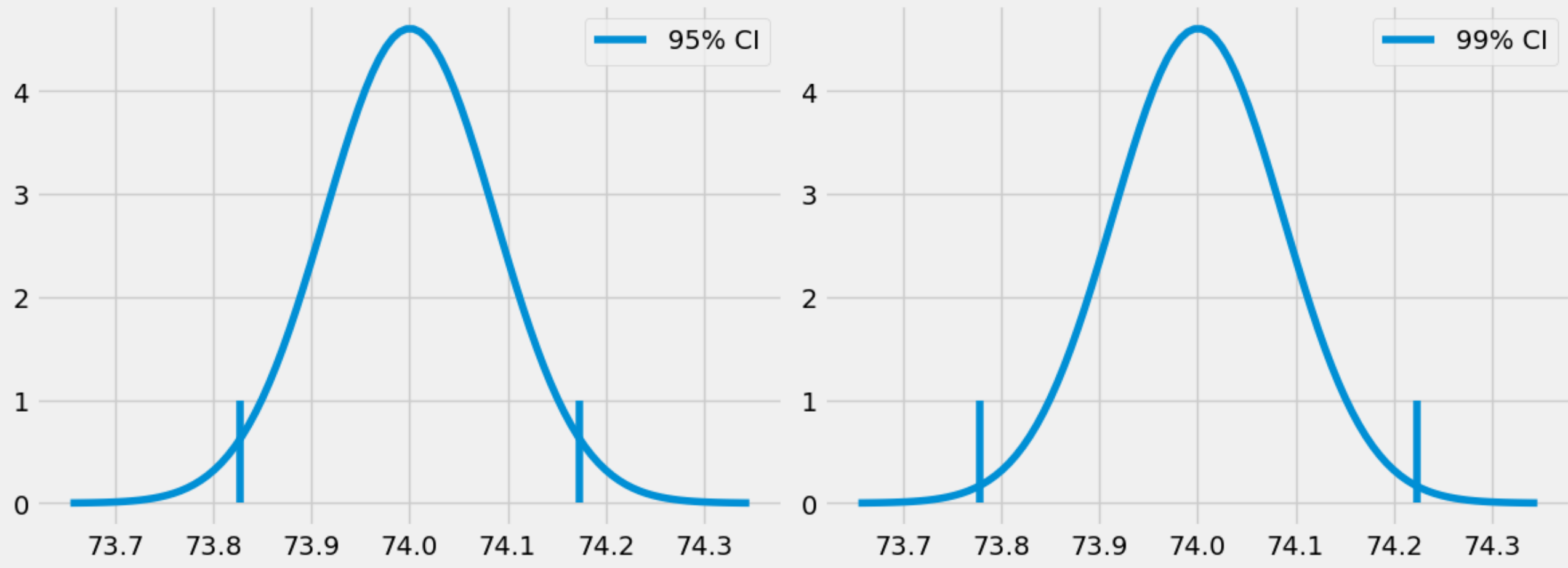
## 置信区间的定义

置信区间是基于样本数据对总体参数进行估计的范围，通常表示为  $\bar{x} \pm 1.96 \hat{\sigma}$ ，其中  $\bar{x}$  是样本均值，1.96是标准正态分布的临界值  $\hat{\sigma}$  是标准误差。

## 95%置信区间的计算

以95%置信区间为例，临界值 ( $z$ ) 为1.96。通过图表展示置信区间的含义，解释其在统计推断中的作用，帮助我们判断样本均值是否显著偏离总体均值。

## 置信区间的构建



95% CI for Online: (70.56094429049804, 76.7095818797147)  
95% for Face to Face: (76.80278229206951, 80.29218687459715)

# Part. 04

假设检验与P值

---

CHAPTER



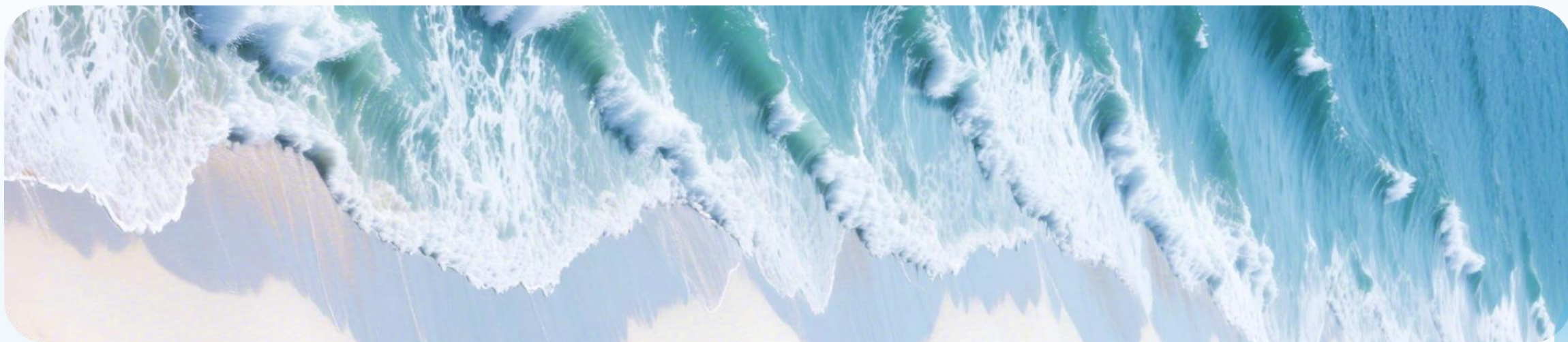
# 假设检验的基本概念

## 01 原假设与备择假设

假设检验的基本概念包括原假设和备择假设。原假设通常表示没有差异或效应，而备择假设则表示存在差异或效应。通过假设检验，我们可以判断数据是否支持备择假设。

## 02 $z$ 统计量的应用

假设检验中常用 $z$ 统计量来衡量样本均值与总体均值的差异。通过计算 $z$ 值并将其与标准正态分布的临界值进行比较，可以判断结果的显著性。



# P值的计算与解释



## P值的定义

P值是在零假设成立的条件下，观察到当前或更极端结果的概率。它用于判断结果的显著性，通常以0.05为显著性水平。

## P值的计算方法

通过实际案例展示P值的计算过程。例如，在一个样本中，计算得到的z值为2.5，对应的P值为0.012，表示在零假设成立的情况下，观察到当前或更极端结果的概率为1.2%。



## P值的解释

P值越小，表示结果越显著，越有理由拒绝零假设。在实际应用中，P值可以帮助我们判断数据是否支持备择假设，从而做出科学的决策。

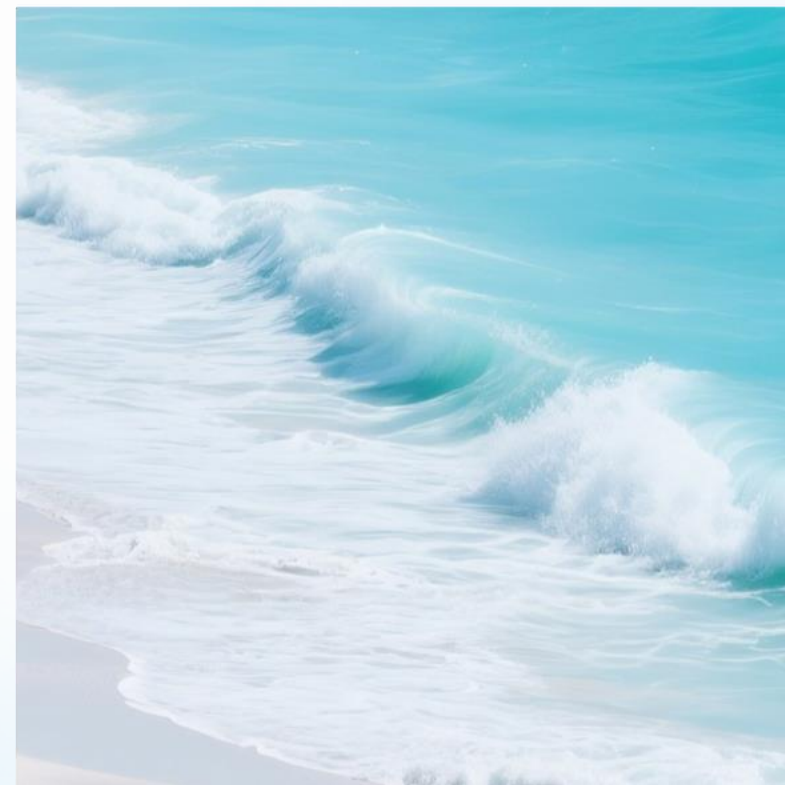




## 置信区间与假设检验的关系

### 两者的关系

置信区间与假设检验密切相关。当两个置信区间不重叠时，差异显著；即使置信区间有重叠，仍可能存在显著差异。通过图表展示这种关系，帮助理解两者的联系。





# Part. 05

实际应用案例

---

CHAPTER





## 在线课程与面对面课程的比较

### 数据比较

通过实际数据比较在线课程与面对面课程的成绩差异。计算两组的置信区间和P值，展示统计分析的过程。

### 统计学意义的解释

解释结果的统计学意义，强调小样本量可能导致的误导性结论。在实际应用中，需要谨慎对待样本量不足的情况，以避免错误的决策。

# A/B测试的统计分析

1

## A/B测试的介绍

A/B测试是一种常见的统计分析方法，用于比较两个版本的效果。通过实际代码和数据展示A/B测试的分析过程。

2

## 置信区间与z统计量的计算

展示如何计算两组的置信区间、z统计量和P值。通过实际案例解释这些统计量的含义和应用。

3

## 结果的解释与应用

解释A/B测试结果的统计学意义，帮助理解如何根据结果做出科学的决策。强调在实际应用中考虑样本量和不确定性的必要性。

「LOGO」

# Part. 06

总结与展望

---

CHAPTER



# 统计学的核心要点

## 棣莫弗方程的重要性

棣莫弗方程揭示了样本量与不确定性之间的关系，是理解统计学基础的重要工具。

01

## 标准误差的作用

标准误差帮助我们评估估计值的精确性，是统计推断中的关键指标。

02

## 置信区间的构建

置信区间用于估计总体参数的范围，帮助我们判断结果的可靠性。

03

## 假设检验与P值

假设检验和P值用于判断结果的显著性，是统计推断的重要方法。

04



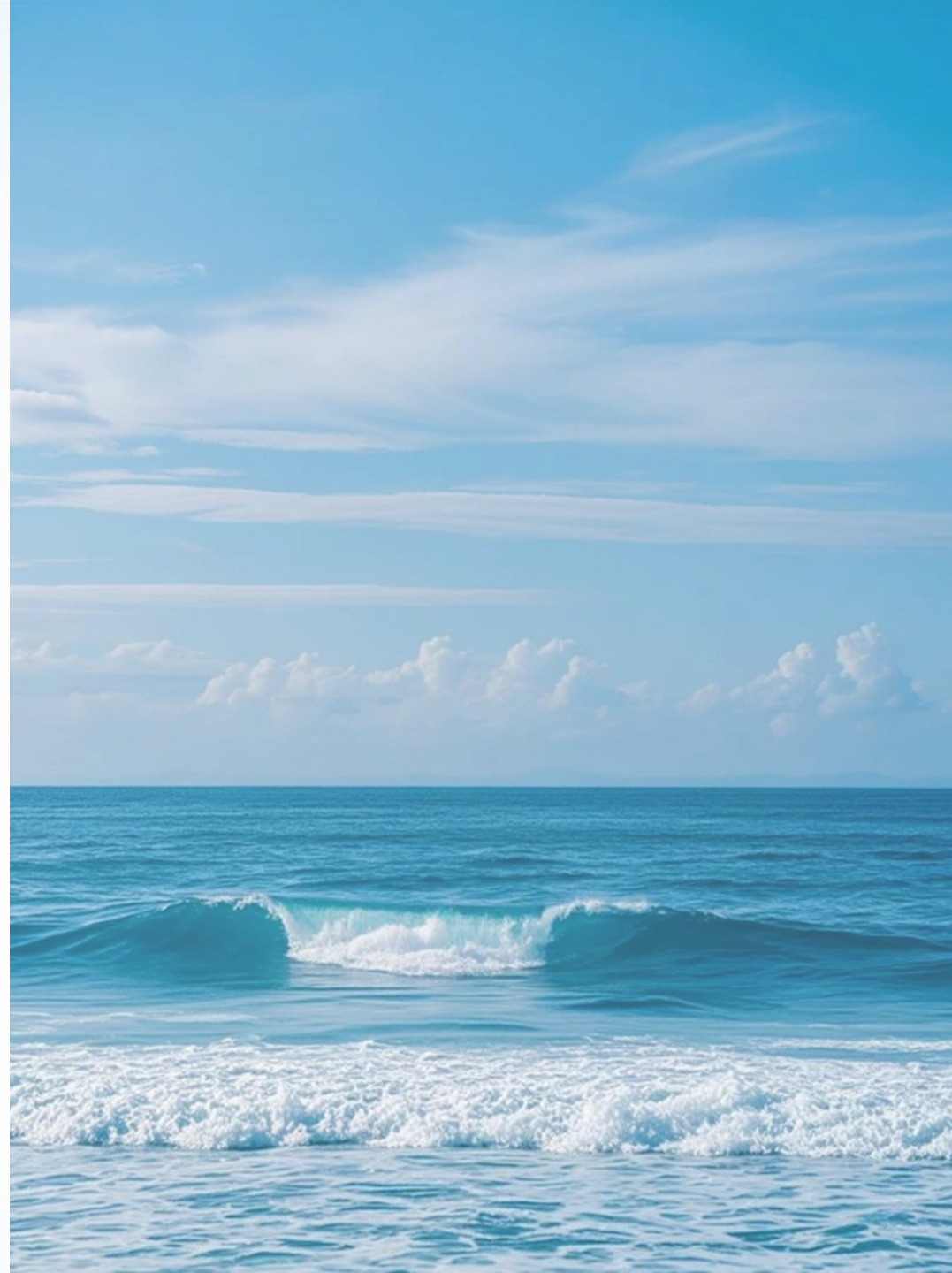
# 对数据分析的启示

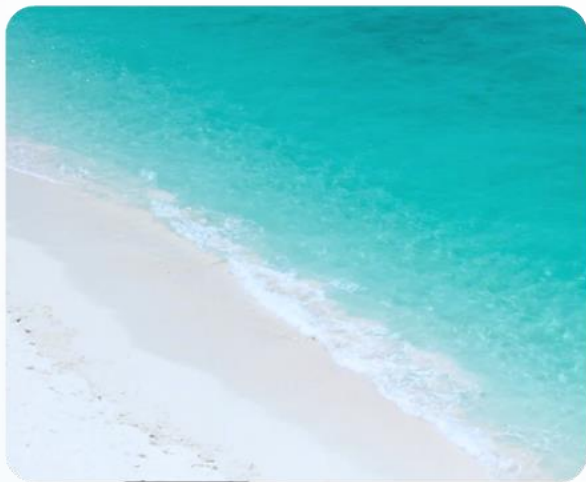
## 避免小样本量的误导

在数据分析中，应警惕小样本量带来的误导性结论。小样本可能导致极端结果，从而影响决策的准确性。

## 正确应用统计学方法

在实际决策中，应正确应用统计学方法，考虑样本量和不确定性，以确保分析结果的可靠性和科学性。





## 未来研究方向

### 改进统计学方法

展望未来研究方向，探讨如何进一步改进统计学方法以更好地处理小样本数据。可能的研究方向包括贝叶斯统计的应用和更复杂的统计模型。



谢谢  
Thank you for listening

