

# 设计研究中样本容量与置信水平关系分析

江滨<sup>1</sup>, 武禹宸<sup>2</sup>

(1. 上海中侨职业技术大学, 上海 201411; 2. 上海工程技术大学, 上海 201620)

**摘要:** 本文针对设计研究领域样本规划科学性问题, 系统探究了样本容量与置信水平之间存在的内在关系及其对设计研究的影响。指出当前设计领域在样本确定过程中存在显著的随意性和非系统性弊端, 制约了研究结论的准确性与结论运用价值。通过综合运用文献研究、抽样分析等方法, 深入剖析了在设计样本容量确定过程中的典型认知误区, 明确了置信区间、抽样误差与样本容量之间的数学关系, 提出了兼顾统计严谨性与实践可行性的样本规划框架。通过实证进一步揭示了三者间存在的非线性平衡关系, 指出盲目扩大样本容量对精度提升的边际效益递减现象, 为设计研究者提供了系统的样本规划框架, 有助于在保证研究信度的同时优化资源配置。文末就样本容量与置信水平的应用前景与后续研究方向提出展望, 为推进设计研究方法的科学化提供了重要参考。

**关键词:** 统计置信度; 设计研究; 样本规划; 样本容量; 置信水平

**DOI:** <https://doi.org/10.71411/-2025-v1i4-1046>

## Analysis of the Relationship Between Sample Size and Confidence Level in Design Research

jiangBin<sup>1</sup>, WuYuchen<sup>2</sup>

(1. Shanghai Zhongqiao Vocational And Technical University, Shanghai 201411; Shanghai University of Engineering Science 201620)

**Abstract:** This article addresses the scientific issues of sample planning in the field of design research, systematically explores the intrinsic relationship between sample size and confidence level and its impact on design research. It points out that there are significant arbitrariness and unsystematic shortcomings in the current sample determination process in the design field, which restrict the accuracy of research conclusions and the value of their application. By comprehensively using methods such as literature research and sampling analysis, the typical cognitive misunderstandings



in the process of determining sample size in design were deeply analyzed, clarifying the mathematical relationship among confidence interval, sampling error, and sample size, and proposing a sample planning framework that balances statistical rigor with practical feasibility. Empirical research further reveals the nonlinear balance among the three, indicating the diminishing marginal returns of blindly increasing sample size for precision improvement, providing design researchers with a systematic sample planning framework that helps optimize resource allocation while ensuring research reliability. At the end of the article, prospects for the application of sample size and confidence level as well as future research directions are discussed, providing important references for advancing the scientific approach of design research methods.

**Key word:** Statistical confidence; research design; sample planning; sample size; confidence level; resource optimization

## 1. 绪论

### 1.1 研究背景与意义

统计学作为当代设计研究领域数据分析的关键工具,在大数据背景下迎来了前所未有的机遇与挑战<sup>[1]</sup>。统计置信度是衡量研究结果可靠程度的重要参考,它直接影响到设计研究结论的精确性,也关系到设计实践中决策是否具备科学依据,不过当前设计研究中的样本规划常常显得随意且缺乏系统性,这在一定程度上削弱了研究成果的科学性和准确性,借助文献回顾的方法对已有研究进行梳理可以发现,统计方法在设计研究中应用广泛,但专门讨论样本规划的内容却不多,很多研究仅仅关注样本数量的计算,而未能充分考虑样本结构的优化以及置信度对最终结果的影响,这种状况不仅限制了设计学科的理论推进,也在实际设计过程中提高了决策失误的可能性。本文对样本容量与置信水平之间的关联展开分析,尝试在设计实验开始阶段就科学地确定所需样本规模,从而避免因样本损失带来的资源浪费或样本不足所引发的问題。明确结论的可信区间也有助于为后续实验划定适用边界,提升实验数据的实际应用价值。

“统计置信度”与“样本容量”的广泛运用,并非简单的方法借用,而是设计学科从依赖直觉、经验与权威的传统范式,向强调证据、可重复性与普遍性的现代科学范式的根本性转变。通过对数据的一致性和稳定性进行可靠性分析,如通过样本容量是否充足、统计检验是否显著,设计研究能够有效控制方向误差,使其研究发现免受小样本偶然性的干扰,反映各因子间的可靠性,从而在学术共同体和产业实践中建立更高的可信度<sup>[2]</sup>。在会展空间布局优化、导视系统效率评估、沉浸式体验效果测量等具体场景中,设计师或研究者可以依据统计检验的客观结果(p值、置信区间),而非个人喜好,选择最优设计方案,实现了设计决策的理性化与透明化。

### 1.2 问题描述与研究目标

本研究通过科学计量学工具 CiteSpace 对来自 CNKI 筛选所得的 500 篇中英文文献进行“样本容量”与“置信水平”关键词共现分析,关键词是对研究内容的高度凝练和精炼阐述,通过对关键词之间关联性的探究,能够在一定程度上解释不同学科领域的内在关联。文章通过对关键词共现分析、关键词频次统计以及关键词聚类



等，帮助研究者识别设计领域如何使用样本容量<sup>[3]</sup>。整理得到的论文关键词关系图谱可知，“统计置信度”与“样本容量”已成为当代设计研究，特别是以实证和量化为导向的研究分支中，不可或缺的方法论支柱。在“统计置信度”与“样本容量”构成的关键词共现图谱网络中，节点的大小代表关键词出现频次，连线的粗细代表共现强度。得出统计往往被用于“设计优化”、“眼动追踪”、“交互体验”等不同子研究群中。这种结构位置证明，统计与设计并非毫无交集，而是深度渗透并串联起了设计实证研究的整个流程。这一结果的确立，为本研究提供了至关重要的理论支撑。任何旨在对设计效果、用户行为或空间性能做出有效推断的研究，都必须严肃对待数据收集的规模与统计分析的严谨性。以经过统计分析检验得出的可信结论作为设计依据，避免将随机波动产生的误判带入后续设计当中。即使在偏重工程与环境的可持续设计领域，对环境影响或用户行为改变的评估，也需建立在具有统计信度的大样本数据之上，以确保决策的可靠性。各行各业的设计研究当使用规范的样本容量大小以及合适的置信度能够为设计优化提供真实可靠的理论支撑。为从宏观层面精准把握“统计”与“设计”交叉领域的知识结构与研究焦点。通过对生成的知识图谱进行算法解读与深度挖掘，一个核心发现得以凸显：“统计置信度”与“样本容量”这两个关键方法论概念，在设计研究领域呈现出异常高频的共现关系与中心性地位，说明以严谨的统计推断为核心特征的实证研究范式，已在设计科学研究中确立其基础性且广泛的应用格局。

在会展与空间设计领域的实际操作中，前期的设计决策和最终效果的呈现，大多还依靠设计师本人的经验积累、直觉判断以及个人偏好，这种依赖经验的设计方式，其局限性已经越来越明显。一方面，设计师的个人感受往往是局部的，

很难全面反映不同用户群体的多样化需要；另一方面，基于感性判断的展示成果不容易用数据来衡量，这给设计经验的沉淀和传递带来了困难，此外，由设计师主观感受主导的设计效果，由于缺乏客观数据的支持，在解释设计缘由和依据时，常常会陷入难以自圆其说的境地。所以，寻找并构建一套系统、客观、能够量化的设计方法，把感性认知和理性数据结合起来，已成为提升会展设计科学水平和专业程度的迫切任务。

本研究正是基于这一背景，旨在通过引入统计抽样与置信区间分析，为空间满意度提供一种超越个人感知直觉的、基于准确数据的决策支持工具。为应对前述依赖主观判断所带来的不确定性，本研究采用概率抽样与推断统计的方法，将‘感觉’转化为‘数据’，把‘猜测’提升为‘估计’。针对量化会展与空间环境设计当中，用户对场馆的期待。通过科学抽样方法，抽取一定数量的、有代表性的观众的感知数据，加上置信区间量化这一估计的可靠程度与误差范围，让设计师与参与者都能清晰地了解决策所依据的数据来源、样本规模及结论的置信水平，从而达到建立共同决策的基础。



图 1 样本容量研究关键词共现图谱

1.3 本文的结构安排

对被引用文献深入解读发现绝大多数研究将统计方法应用于数据收集之后的分析阶段，却





较少有研究前瞻性地、系统性地关注数据收集本身的科学基础，即如何规划样本容量与确定抽样设计的严谨性。当前研究多在讨论如何分析数据，却相对忽视了如何获取具有代表性的数据这一前提要素。针对上述的问题我们选择会展空间满意度这一典型设计评价指标，采用随机抽样方法，以确保样本在维度上对总体的代表性。抽样框架为采用简单随机抽样抽取预定数量的参与者。

在会展与空间环境设计领域，任何用户体验指标的优劣比较，在调研市场过程中，应当基各类资源性能指标差异，实施资源优化配置，通过适当样本容量收集的数据以实现调研效能的最大化<sup>[4]</sup>。本论文后续针对的实证部分，将严格遵循此范式，在设计调研方案时预先进行样本容量估算与功效分析，并在数据解析环节运用恰当的统计推断方法，以确保研究结论本身的置信度，这正是对本领域核心方法论共识的自觉践行与深化。

## 2. 相关工作

### 2.1 设计研究中的样本容量确定方法

在设计研究当中，分析设计样本的个体差异，建立的最小数据集，计算样本容量，选择适当抽样比例，为设计样本容量的资源收集、保存及设计研究工作提供参考和理论依据<sup>[5]</sup>。样本往往被定义为使用预定的方法从较大的数据集当中选择的较小的数据集。样本容量的确定直接影响到统计效力和统计结论效度。需要保证设计结果的准确性，就要选择合适的样本量。在概率抽样（如简单随机抽样）中，有一个科学的样本量计算公式：

$$n = \frac{[Z^2 \times p \times (1-p)]}{e^2}$$

这个公式看起来复杂，但每个部分都代表了所需样品数量的决策依据：“n”：为取样所需要的“样本容量”。“Z”：代表“置信水平”的量化值，选择不同的置信度进行计算时，也要

正确使用对应的量化值，这体现了你所需结过的可靠性。常见取值有 90% 置信水平对应的 Z 值为 1.64，而最常用的 95% 置信水平对应的 Z 值为 1.96，99% 置信水平下对应的 Z 值为 2.576。

“p”为整体预估比例也就是总体的预期比例或方差。预计有 50% 的人选择某个选项，p 就取 0.5。当设计实验开始前对所需调研情况一无所知，可以选择使用 0.5，因为当 p=0.5 时，p\*(1-p) 的值最大 (0.25)，这样计算出的样本量是最保守、最保险的，能确保任何情况下的样本量都足够。如果有历史数据或预调研数据，可以采用更精确的 p 值，从而可能减少样本量。“e”：可接受的抽样误差范围。例如，实验需要将调查结果的误差控制在 ±3% 以内，e 的取值就是 0.03。在 95% 的置信水平 (Z=1.96)、误差不超过 ±5% (e=0.05)、且对 p 值不确定的情况下进行调查： $n = [(1.96)^2 \times 0.5 \times (1-0.5)] / (0.05)^2 \approx 385$ 。按照上述统计方法确定的所需样本容量，设计者在设计实验当中每组子数据可以设定为 385。

### 2.2 置信水平在统计推断中的运用

在设计研究领域，如何应用统计置信度驱动下的样本规划模型尤为重要。置信水平常被用来估计抽样误差。其亦可称作保证率、可靠度、置信度、置信系数，是指总体参数值落在样本统计值某一区间内的概率，通常用 (1 - α) % 表示<sup>[6]</sup>。核心在于借助科学的样本选取与配置，提升设计研究工作的效率和准确性。设计研究者应用中要明确研究问题所涉及的变量以及它们之间的关系，借助统计和设计的基本理论，先搭建出初步的研究框架，这样能够保证后续样本的针对性与有效性。研究问题的提出是设计研究中的关键环节，设计师需要以问题为导向，随后开展样本需求收集和分析工作，这一阶段需要综合考虑样本的代表性与置信度方面的要求。借助统计学中的置信区间和抽样分布理论，设计师可以计



算出特定置信水平下所需的样本容量，从而确保研究结果具备统计显著性。

### 2.3 现有研究的局限性

虽然样本容量与置信水平在医学与社会科学领域中被广泛运用，但是与工科相比，设计本质是一个非线性的、探索性的、高度情境化的创造性过程，如果将设计与高度结构化的统计检验生硬地嵌入，容易导致为了测量而过度简化或扭曲真实的设计情境与用户交互，在设计过程当中仍会出现一下两方面问题：

#### (1) 样本容量方面

研究中最常见的问题是样本量普遍不足且计算方式不科学。一方面，部分研究者仅凭经验或主观判断确定样本量，未以科学性角度对整体样本量进行分析，选择合适的样本容量达到检测差异所需的样本量，导致假阴性结果或无法反映总体特征。另一方面，在确定样本容量之前需要使用正确的计算公式准确计算出总体方差。

在探究多因素设计、重复测量等复杂设计时，对样本量的要求往往更高，不能仅按单组样本容量得出的结果进行研究，要确保样本容量与样本质量失衡。在部分设计研究当中，即便样本量达标，也因在数据清洗与整理过程当中未对样本进行筛选，导致样本同质性差等问题，使得抽选样本失去代表性。未把控样本纳入标准，盲目混入大量不符合研究条件的样本追求样本容量而盲目扩大样本也是不可行的，这样会导致数据噪音增加，降低输出结果的质量。

#### (2) 置信水平方面

在设计当中我们通常使用 95% 置信水平作为标准，但是并非每个设计都是以 95% 为标准，需要结合研究场景，内容和目的调整。在一些对安全性和准确性要求极高的研究中，未采用 99% 或者更高置信度水平进行研究，可能增加结果风险或者在初步市场探索性研究中，使用

95% 置信水平，会导致置信区间过宽，很难快速捕捉市场核心需求。很多研究者将置信区间与预测区间误认为是同一概念，进而错误运用置信水平。置信区间是针对总体参数估计，而预测区间是针对具体观测值估计，二者适用场景不同。若将二者视为同一方式进行研究，错误地认为高置信水平就等同于高准确性，会导致结论偏离真实情况。置信水平与样本量紧密关联，较高置信水平需要更大样本量支撑。但在实际的设计研究当中置信水平与样本量往往会出现不匹配的情况，例如在样本量较小的研究中采用 99% 的高置信水平，会导致置信区间过宽，预测值有更大的概率落在所规定的置信区间外。错误衡量置信水平与样本容量会导致研究结果的误判，进而影响研究的决策。99% 置信水平虽包含真实值概率高，但置信区间较宽会降低估计精度；而低置信水平虽区间较窄，却增加了真实值落在区间外的风险。

### 3. 样本容量与置信水平关系模型

#### 3.1 基本理论框架

统计置信度量化了样本估计值与总体真实值之间的一致性程度，为研究结论的可靠性提供量化依据。在研究线性校正模型置信区间方面，大多数是对校正数据进行单重使用，即用于估计单个未知参数的范围<sup>[7]</sup>。其理论根基植根于概率论和数理统计，置信度实际上描述的是在重复抽样过程中，样本统计量落入某个特定范围的可能性，从原理上看，统计置信度的形成主要借助中心极限定理和大数定律。中心极限定理提到，样本容量足够大的情况下，样本均值的分布会接近正态分布，这为置信区间的建立打下了基础，大数定律则保证随着样本量的上升，样本均值会逐渐逼近总体均值，借此提升置信区间的稳定程度。

置信区间是统计学中的一个核心概念，用于估计总体参数的范围。在设定样本容量区间的



时候, 我们通常使用 Delta 方法构造置信区间, Delta 方法的原理是利用数学上的泰勒展开取得多复杂统计量的均值和方差等特征, 是统计中应用性很强的计算随机变量函数渐进分布的一种方法。<sup>[8]</sup> 然而, 在实际的设计工作当中, 我们往往很难对全部个体进行调查, 于是, 我们便从总体中抽取一定数量的样本进行对总体参数的估计。包括点估计和区间估计两种方法, 点估计是用相应的样本统计量直接作为总体参数的估计值, 区间估计是指从点估计值和抽样标准误差出发, 按预先给定的概率建立包含总体参数的一个区间范围。通常我们将置信区间范围设定为 95% 或者 99%。

### 3.2 关键参数定义

科学性原则构成模型构建的基础, 在跨学科研究样本规划过程中, 需要严格遵循科学性与系统性的要求, 借此来保证数据结论的有效和实用。不同学科内容的结合应当建立在科学合理的基础之上, 科学性原则指出, 跨学科教学要符合认知发展规律, 内容难度控制在可接受范围, 必须依托严谨的统计学理论和数据分析方法, 确保每个环节都经得起逻辑推敲和实证检验。科学性原则还提出跨学科教学内容应具备系统性和连贯性。不同学科知识的融合不是简单拼凑, 而是借助科学手段进行整合, 使学习者在跨学科学习过程中形成有序的知识框架。科学性能帮助跨学科教学做到目标清晰、方法合理<sup>[9]</sup>。样本量确定时, 要依据预设的置信区间和误差范围, 借助中心极限定理开展精确计算, 从而保证样本数据的代表性。实用性原则关注模型在实际应用中的可行与便捷, 这意味着实验设计不仅要理论成立, 还要考虑具体操作中的便利性和资源消耗的合理性, 具体来说, 研究结论要能适应不同规模和类型的设计研究, 提供灵活的参数调整方式, 方便研究人员按实际情况优化配置。借助模块化设计, 模

型能在不同研究背景下快速适配, 扩大应用范围。

模型构建原则	模型验证方法
置信水平	交叉验证
	Bootstrap 重抽样
样本量	数据收集方法
	研究假设与变量定义
误差范围	案例分析
	案例结果分析与讨论

表格 1 样本参数分析流程表格

### 3.3 关系模型的数学表达

统计置信度驱动设计研究样本规划模型构建时 模型参数设定属于关键环节 这直接影响模型的精确程度和可靠性。参数选择能依据统计学原理和设计研究具体需求来推进, 置信水平作为模型核心参数之一 设定标准要兼顾研究严谨性和实际操作可行性, 多数情况下置信水平取 95% (CL=0.95) 借此保证结果具备较高可信度, 样本量确定要结合参数优化法展开。借助敏感性分析可以评估不同样本量对模型输出结果带来的影响, 假设随机变量 X 符合:

$$\varphi_{\sigma, \mu} = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in (-\infty, +\infty)$$

则认为 X 服从正态分布, X 为正态变量, 记作  $X \sim N(\mu, \sigma^2)$ , 其中  $(-\infty < \mu < +\infty)$ 。

## 4. 实验设计与案例分析

### 4.1 实验场景设定

在设计领域, 尤其是针对特定产品, 这种方法可以帮助设计者深入了解用户的使在实证研究设计的框架下, 收集数据的方法选择关系到研究结果的可靠性与有效性。用户行为研究作为一种调研方法, 旨在通过估计均值或比例来描述特定群体的一个或者多个特征理解用户在特定环境或系统内的行为路径和决策过程或者通过回归分析来探索变量之间用习惯、偏好, 以及面临的问题, 从而指导设计过程, 创造更加人性化和满足用户需求的产品<sup>[10]</sup>。本次研究在实证研究设计框架下, 数据收集方法的选择会直接影响研究结



果是否可靠和有效。用户行为研究作为一种调研手段，可以借助估计均值或比例来描述特定群体的某些特征，理解用户在特定环境或系统内的行为路径和决策逻辑，或者通过回归分析探索变量之间的关联。本研究采用了问卷调查和实验数据记录两种主要数据采集方法，以保证数据的全面和多维。问卷调查作为常用的定量研究工具，优势在于能快速收集大量样本数据，方便后续统计分析，问卷设计基于李克特 5 点量表，内容涵盖信任度、态度、评分等影响要素，确保答案标准化和可比性<sup>[1]</sup>，通过对设计研究领域的专家和实践者进行系统抽样，问卷内容涉及样本背景信息、设计经验以及对统计置信度在样本规划中应用的认知与态度，但问卷调查也存在局限，比如可能出现样本偏差或回答者主观性干扰，因此研究进行了信度和效度检验，以保证数据内部一致性和结构效度。

在设计实验开始阶段，借助向多个年龄段和不同职业人群发放问卷的手段来收集实验所需的样本数据，问卷内容涉及被调查者对场馆的满意程度，以及他们对科技如何提升场馆功能的期待值，借此来量化在食品安全受到重视的背景下，消费者对于以食品溯源为主题的虚拟场馆的实际需求，为之后场馆建设提供可靠的数据基础。经过计算，理论上的最小样本数量为 385 份，考虑到问卷可能回收不全以及存在无效回答的情况，实际发放问卷 588 份，最终得到有效问卷 420 份，有效回收的比例是 71.4%，随后将有效问卷信息进行分类并展开整体分析，根据预先设定的置信水平和允许的误差范围，科学地确定样本容量，同时讨论不同样本量对统计推断准确度的具体影响，从而让宏观的统计方法能够在具体操作中真正落地。

## 4.2 不同置信水平下的样本容量计算与分析

### 4.2.1 样本容量选择是否越大越好

在一般的抽样实验当中，通常认为样本容量越大越好。这种假设的前提为以下两方面，首先是当样本容量趋于样本总量时，由于我们是通过抽取的样本对整体目标进行特征估计，那么显然，增大样本容量可以增加预测准确度，但是这并不能说明样本容量越大越好，从公式得出的函数关系可知，当样本容量在 385 左右时，可以预测绝大多数设计研究当中的目标总体，继续增加样本容量也不会得出更有价值的实验结果数据。其次，当样本总量很小时，如果选取数量很小的抽样目标作为样本容量，则选取的样本对整体特征性的预测可信程度不高。按照上述思路，普查是针对样本数量不大时的方法选择，但是在当下的设计研究当中，所需要的调研数量庞大，由于人力，时间，经济等多重限制因素，使得普查的运用日渐减低。样本的大小对设计研究结果的可靠性有着重大影响，样本容量太小无法展示正确的结果，合适且足够的样本容量可以降低结果偏差的可能性。

### 4.2.2 样本容量是否需要与整体成整比

由于样本容量是从整体样本中进行抽样得到的集合，那么在规定样本容量时是否需要与抽样整体成  $1/n$ （ $n$  为正整数）的比例呢？按照这样的理论，假设抽样比例为 5%，样本总数为 100，样本容量大小为 5，对于整体数量而言，5 个样本数目得出的统计效力和统计结论效度并不能达到预测总体的作用。若总体数目为 1000000，则抽样数量达到 50000，重复进行如此大规模的抽样取样显然也是不可行的。从中可知，不同置信度水平下的抽样误差并非成等比，而是与样本容量的平方呈反比。因此，样本容量与置信水平达到某个定值时，减少抽样误差需要抽取更多的样本。如表格 2 所示在 99% 置信度下，误差范围从 2% 降低至 1% 需要增加 12481 个抽样样本，而置信度 10% 降低至 9% 只需增



加 39 个抽样样本。在设计中需要平衡抽样误差带来的样本量增加与设计结果之间的关系。避免造成抽样数量过多而造成的资源浪费或者抽样数量过少造成的实验结果误差。

确定所需样本容量时，需考虑多个因素，包括目标总体的规模、允许的误差范围、以一定的置信水平代表总体以及调研预算等。通常，可通过统计公式或专业软件来计算所需的样本量。此外，还需注意样本的代表性，确保样本能够充分反映目标总体的特征。

抽样误差	置信水平		
	0.9	0.95	0.99
±1%	6806	9604	16641
±2%	1702	2401	4160
±3%	756	1067	1849
±4%	425	600	1040
±5%	272	384	666
±6%	189	267	462
±7%	139	196	340
±8%	106	150	260
±9%	84	119	205
±10%	68	96	166

表格 2 抽样、置信水平与所需最小样本容量的关系

注：表格 2 引自朱雁“教育研究中的样本容量”

#### 4.3 固定样本容量的估计精度分析

在置信区间内抽取的样本容量是对样本总体参数进行估计的参照，表明抽样样本有一定几率落在测试范围内。在规定样本置信度的前提下，当样本容量足够大是，样本图形会呈现标准正态分布图样，即样本的两端更为平缓。如图 2 为样本容量与抽样误差的关系图。置信区间的变化程度并不完全与样本量成比例关系。如表所示，样本容量从 800 增加至 1600 时，间隔也并非减少为 1/2，因此在设计取样时，当样本容量达到一定数量的时候便不再增加样本量。起到收支平衡的效果。样本容量与置信区间的变化由公式得到。置信水平，置信区间，抽样误差三者关系可归结

为。在抽样中固定置信水平，需要将误差范围缩小则需要将置信区间扩大。如图可见将样本容量 1000 扩大至 2000 时，抽样误差并未显著降低。也就是说，在设计实验当中可以选择适当的样本容量对在保证实验结果准确的前提下达到降低损耗的效果。

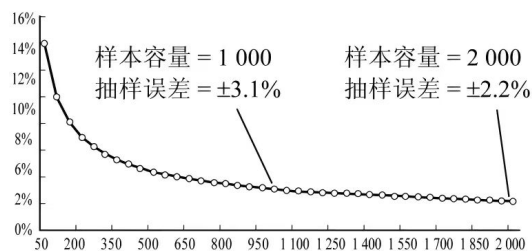


图 2 样本容量与抽样误差的关系图

注：图 2 引自黄文仙的“决定样本的大小”

### 5. 结果与讨论

#### 5.1 关系模型的量化结果

使用 Origin 软件进行数据的清洗和分析。并且分析计算 G“均值”（“G”为有效问卷中，每个参与问卷调查的受访者，所有问题相加得到的分数除以总题数）的置信区间，得出如图 3 所示，判断在有效问卷的人群中，对于场馆打分的值为 3.3。通过设定的 95% 置信区间计算可得，下一位有效问卷用户的问卷分数有 95% 的概率落在区间 [2.38,4.23] 内。在后续调研当中，若置信区间重叠面积较小，说明差异显著。置信区间有助于评估样本在各产品类别中的代表性。若某类别的置信区间较宽，则需要增加该类别的样本量。根据数据走向趋势从市场宏观经济、竞品政策调整、消费者偏好技术进步等市场环境变化会对市场规模和产品消费规模产生影响。结合数据走向趋势变化的多维度业务包括应对市场就业新机会制定相应的市场拓展策略。通过分析产品消费规模的占比变化。用户群体消费规模的精分占比变化制定差异化的营销策略。通过置信区间样本确保研究结果的可靠性和有效性，规划制定科学合理的策划，为后续实验的决策提供有力的



数据支持。

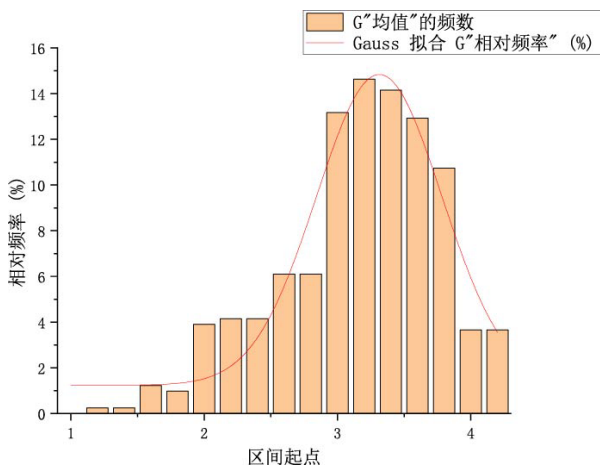


图3 抽样样本均值的正态分布

## 5.2 对设计实践的启示

本研究的实践意义启示在于，通过在“引入样本容量”以及“置信区间”的概念证明了在设计初期调研与后期评估中，引入相对规模大小的用户测量数据是可行且高效的。对于后来的设计实践者而言，本研究说明：在设计的关键决策点，可以借鉴本研究的方法，快速收集数目标用户的量化反馈，并依据统计置信度做出风险更可控的选择。有助于在行业内部培育一种“用数据说话”的理性决策观念。

## 5.3 研究的局限性

首先从测量维度角度出发，本研究采用的量化量表虽保证了效率与客观性，但可能无法完全捕捉用户体验的复杂性与情境性。未来研究可采用混合方法，通过量化调查与用户访谈等方法结合，解析调研对象打分背后原因的深度理解。其次，在研究设计上，单一的截面数据限制了我们对于用户体验动态过程的理解，后续研究在观展前、中、后多个时间点进行纵向追踪设计的数据采集方法，或利用移动设备进行体验抽样，以绘制更完整的用户体验数据。最后，在技术应用上，伴随着人工智能技术的不断成熟以及大数据的广泛应用，以5G、大数据等技术为代表的人工智能技术不断发展发展，尤其是深度学习算法设计

持续优化，人工智能所依赖的数据和算法两个必要条件已经具备，大数据的广泛运用也为设计调研获取样本提供了新思路。在未来研究中可以结合生理数据（如眼动、心率）、行为轨迹数据与主观量表数据，构建一个更立体、更客观的会展空间体验评估框架，这也是本领域一个富有前景的前沿方向。

## 6. 结论与展望

设计研究中的样本容量和置信度水平是本文重点讨论的内容，文章借助对置信度水平与样本容量关系的梳理，提到当前设计研究在样本规划环节存在一些不足，进而说明构建一个兼顾统计严谨性与实践可行性的样本规划框架很有必要。在此基础上，本文进一步阐明模型构建需遵循的科学性原则，为设计研究搭建理论支撑。围绕置信水平、样本容量等量化理论展开分析，能看出统计学在设计领域具有重要价值。设计研究者可以借助收集和分析有效的用户数据，把“用户体验”这类抽象概念转变成能够可视化的量化指标。借助科学软件处理大规模、多维度的有效数据，有助于揭示数据背后的普遍性规律与结构性特征，帮助设计师依据可靠的数据结论把握不同人群在会展空间中的实际需求，提升设计方案的说服力与可信度。引入统计方法之后，会展与空间设计的研究成果能够以标准化、可复现、可验证的方式呈现，这不仅增强了设计学科的严谨性和可信度，也促进了设计学科与其他学科之间的融合，推动不同领域系统化知识体系的共建。

统计学方法所搭建的评估框架，能基于有效数据结果准确识别问题所在，借此为下一轮设计或同类项目提供具体、可操作的优化方向，促进设计知识的持续积累，通过上述探讨，本文为统计置信度驱动下的设计研究样本规划模型的构建与应用提供了较为全面的理论支持。



## 参考文献

- [1] 张伟娜, 毕忠勤, 王宝楠. 面向数据科学的应用统计学创新教学体系探索 [J]. 计算机教育, 2024, (05): 175-179.
- [2] 金辞宜. 城市公园老年人休闲动机与满意度研究——以杭州市湖滨公园为例 [J]. 园林, 2025, 42(11): 87-96.
- [3] 吴涵. 新时代文化强国研究述评——基于CiteSpace的文献计量与可视化分析 [J/OL]. 昆明理工大学学报(社会科学版), 1-8.
- [4] 茆美琴, 葛星, 杨铖, 等. 计及用户偏好的电动汽车集群参与调频市场优化控制 [J]. 电力系统自动化, 2025, 49(14): 130-140.
- [5] 郭晓亮, 范荫荫, 黄浩, 等. 重齿毛当归表型性状筛选与样本量确定 [J/OL]. 分子植物育种, 1-12.
- [6] 何笑, 贺欢, 田茂再, 等. 二项分布下相对风险的置信区间构造方法比较 [J]. 数学的实践与认识, 2025, 55(10): 259-275.
- [7] 兰晓君. 探索应用统计学在大数据背景下的应用与创新 [J]. 内蒙古统计, 2025, (03): 48-51.
- [8] 赵源璐. 基于组合信息的单变量线性校正模型多重使用置信区间研究 [D]. 北方工业大学, 2025.
- [9] 葛丽妍. 基于素质教育的小学体育跨学科教学模式研究 [J]. 教师, 2025, (11): 116-118.
- [10] 李光辉, 宋武, 陈振益. 基于用户行为及Kano模型的流水茶台设计研究 [J]. 包装工程, 2025, 46(06): 142-150.
- [11] 常鑫, 廖语嫣, 汪林, 等. 城市空中交通公众接受度研究 [J]. 公路交通科技, 2025, 42(03): 34-44.

## 作者简介:

江滨 (1962—), 男, 博士, 教授, 研究方向为可持续设计。

武禹宸 (2003—), 男, 上海工程技术大学硕士研究生, 研究方向为可持续设计。