

语言学量化研究方法导论

第一讲：概述、理论与基本概念

(guoshujian@tongji.edu.cn)

Tongji University

2025 年 11 月 5 日

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

本讲大纲

- 1 导论：为什么需要量化？
- 2 核心概念
- 3 常见量化研究方法
- 4 学以致用：课堂练习
- 5 总结与展望

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

什么是语言学研究？

传统的语言学研究

- 侧重于语言的结构、历史和本质。
- 方法论：通常是定性的 (Qualitative)，依赖于内省、观察和个案分析。
- 例子：音系学规则的推导、句法结构的分析（如生成语法）、词义的演变。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

什么是语言学研究？

传统的语言学研究

- 侧重于语言的结构、历史和本质。
- 方法论：通常是定性的 (Qualitative)，依赖于内省、观察和个案分析。
- 例子：音系学规则的推导、句法结构的分析（如生成语法）、词义的演变。

一个问题

“这个语法结构是‘可以接受的’吗？”

- 传统方法：语言学家（或母语者）的语感判断。
- 局限性：主观性、个体差异、缺乏代表性。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

为什么语言学需要量化？

- 客观性 (Objectivity)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

为什么语言学需要量化？

- 客观性 (Objectivity)
 - 使用数据和统计来验证假设，减少研究者的主观偏见。
- 可重复性 (Replicability)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

为什么语言学需要量化？

- **客观性 (Objectivity)**
 - 使用数据和统计来验证假设，减少研究者的主观偏见。
- **可重复性 (Replicability)**
 - 清晰的研究设计和数据分析过程，使得其他研究者可以重复实验并验证结果。
- **可推广性 (Generalizability)**

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

为什么语言学需要量化？

- **客观性 (Objectivity)**
 - 使用数据和统计来验证假设，减少研究者的主观偏见。
- **可重复性 (Replicability)**
 - 清晰的研究设计和数据分析过程，使得其他研究者可以重复实验并验证结果。
- **可推广性 (Generalizability)**
 - 通过对样本的统计分析，将结论推广到更大的群体（总体）。
- **处理大规模数据**

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

为什么语言学需要量化？

- **客观性 (Objectivity)**
 - 使用数据和统计来验证假设，减少研究者的主观偏见。
- **可重复性 (Replicability)**
 - 清晰的研究设计和数据分析过程，使得其他研究者可以重复实验并验证结果。
- **可推广性 (Generalizability)**
 - 通过对样本的统计分析，将结论推广到更大的群体（总体）。
- **处理大规模数据**
 - 计算机技术的发展使得处理数百万甚至数十亿词的语料库成为可能。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

定性研究 vs 定量研究

定性研究 (Qualitative)

- 目的：深入理解、探索意义。
- “为什么？”“怎么样？”
- 数据：文本、访谈、观察。
- 分析：归纳、诠释、话语分析。
- 样本：小，有目的性。

定量研究 (Quantitative)

- 目的：检验假设、测量关系。
- “多少？”“频率如何？”
- 数据：数字、可量化的编码。
- 分析：统计、模型。
- 样本：大，随机。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

定性研究 vs 定量研究

定性研究 (Qualitative)

- 目的：深入理解、探索意义。
- “为什么？”“怎么样？”
- 数据：文本、访谈、观察。
- 分析：归纳、诠释、话语分析。
- 样本：小，有目的性。

定量研究 (Quantitative)

- 目的：检验假设、测量关系。
- “多少？”“频率如何？”
- 数据：数字、可量化的编码。
- 分析：统计、模型。
- 样本：大，随机。

核心观点

定性与定量并非对立，而是互补的。

- 定性研究可以帮助形成假设（探索性）。
- 定量研究可以用来检验假设（验证性）。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

① 提出研究问题 (Research Question)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

- ① 提出研究问题 (Research Question)
- ② 形成研究假设 (Hypothesis)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

- ① 提出研究问题 (Research Question)
- ② 形成研究假设 (Hypothesis)
- ③ 操作化：定义变量 (Variables)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

- ① 提出研究问题 (Research Question)
- ② 形成研究假设 (Hypothesis)
- ③ 操作化：定义变量 (Variables)
- ④ 收集数据 (Data Collection)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

- ① 提出研究问题 (Research Question)
- ② 形成研究假设 (Hypothesis)
- ③ 操作化：定义变量 (Variables)
- ④ 收集数据 (Data Collection)
- ⑤ 分析数据 (Data Analysis)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

量化研究的基本流程

- 1 提出研究问题 (Research Question)
- 2 形成研究假设 (Hypothesis)
- 3 操作化：定义变量 (Variables)
- 4 收集数据 (Data Collection)
- 5 分析数据 (Data Analysis)
- 6 得出结论 (Conclusion)

示例

- 问题：双语者在切换语言时反应更慢吗？
- 假设：是的，存在“语言切换成本” (Switch Cost)。
- 变量：IV (自变量) = 任务类型 (切换 vs. 不切换), DV (因变量) = 反应时。
- 数据：实验中测得的反应时 (毫秒)。
- 分析：t 检验，比较两种条件下的平均反应时。

研究问题 (Research Question)

好的研究问题应该是...

- **明确的 (Specific)**: 不模糊, 清楚界定研究范围。
- **可测量的 (Measurable)**: 可以被量化和收集数据。
- **有意义的 (Significant)**: 对理论或实践有贡献。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

研究问题 (Research Question)

好的研究问题应该是...

- **明确的 (Specific)**: 不模糊, 清楚界定研究范围。
- **可测量的 (Measurable)**: 可以被量化和收集数据。
- **有意义的 (Significant)**: 对理论或实践有贡献。

例子

- **不好的问题**: 社交媒体对语言有影响吗?(太模糊)
- **好的问题**: 在微博平台上, “yyds” 的使用频率在 2021 年至 2023 年间是否发生了显著变化?(明确、可测量)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

假设 (Hypothesis)

定义

假设是对研究问题的一个可检验的预测 (Testable Prediction)。

- **零假设 (H_0):** 假设没有效应或没有差异。(e.g., A 组和 B 组的反应时没有差异)
- **备择假设 (H_1):** 假设存在效应或差异。(e.g., A 组的反应时显著快于 B 组)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

假设 (Hypothesis)

定义

假设是对研究问题的一个可检验的预测 (Testable Prediction)。

- **零假设 (H_0)**: 假设没有效应或没有差异。(e.g., A 组和 B 组的反应时没有差异)
- **备择假设 (H_1)**: 假设存在效应或差异。(e.g., A 组的反应时显著快于 B 组)

统计推断的目标

我们的目标是看收集到的数据是否有足够的证据来**拒绝**零假设 (H_0)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

变量 (Variables)

变量的类型

- **自变量 (Independent Variable, IV)**
 - 研究者操纵或观察的变量，被认为是“原因”。
 - e.g., 语言背景（双语 vs. 单语）、刺激词类型（高频词 vs. 低频词）。
- **因变量 (Dependent Variable, DV)**
 - 被测量的变量，被认为是“结果”。
 - e.g., 反应时 (RT)、准确率、F0 (基频)、某个词的使用频率。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

变量 (Variables)

变量的类型

- **自变量 (Independent Variable, IV)**

- 研究者操纵或观察的变量，被认为是“原因”。
- e.g., 语言背景 (双语 vs. 单语)、刺激词类型 (高频词 vs. 低频词)。

- **因变量 (Dependent Variable, DV)**

- 被测量的变量，被认为是“结果”。
- e.g., 反应时 (RT)、准确率、F0 (基频)、某个词的使用频率。

练习

“研究者想知道不同的教学方法 (A vs. B) 是否会影响学生的语法测试成绩。”

- IV 是什么？ (教学方法)
- DV 是什么？ (语法测试成绩)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

数据类型 (Types of Data)

分类变量 (Categorical)

- **定类 (Nominal)**

- 类别之间没有顺序。
- e.g., 词性 (名词, 动词), 母语 (英, 中, 法)。

- **定序 (Ordinal)**

- 类别之间有顺序, 但间隔不均等。
- e.g., 李克特量表 (非常同意, 同意, ...), 语言水平 (初, 中, 高)。

连续变量 (Continuous)

- **定距 (Interval)**

- 有顺序, 间隔均等, 但没有“真零点”。
- e.g., 温度 (摄氏度), 标准化测试分数。

- **定比 (Ratio)**

- 有顺序, 间隔均等, 有“真零点”。
- e.g., **反应时** (0ms = 没有时间), **词频**, 年龄, 元音时长。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

数据类型 (Types of Data)

分类变量 (Categorical)

- **定类 (Nominal)**

- 类别之间没有顺序。
- e.g., 词性 (名词, 动词), 母语 (英, 中, 法)。

- **定序 (Ordinal)**

- 类别之间有顺序, 但间隔不均等。
- e.g., 李克特量表 (非常同意, 同意, ...), 语言水平 (初, 中, 高)。

连续变量 (Continuous)

- **定距 (Interval)**

- 有顺序, 间隔均等, 但没有“真零点”。
- e.g., 温度 (摄氏度), 标准化测试分数。

- **定比 (Ratio)**

- 有顺序, 间隔均等, 有“真零点”。
- e.g., **反应时** (0ms = 没有时间), **词频**, **年龄**, **元音时长**。

为什么这很重要?

不同的数据类型决定了你能使用哪种统计分析方法!

- 分类数据 → 卡方检验 (Chi-Square)

描述性统计 (Descriptive Stats)

目的：总结和描述数据的基本特征

- **频数 (Frequency)**

- 某个值或类别出现了多少次？
- e.g., 在语料库中, “happy” 出现了 500 次。

- **集中趋势 (Central Tendency)**

- **平均数 (Mean)**: 所有值的总和除以值的数量。(受极端值影响大)
- **中位数 (Median)**: 排序后位于中间的值。(不受极端值影响)
- **众数 (Mode)**: 出现次数最多的值。

- **离散程度 (Dispersion)**

- **标准差 (Standard Deviation)**: 数据偏离平均数的程度。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

推断性统计 (Inferential Stats)

目的：从样本数据推断总体的特征

我们几乎永远无法测量总体 (e.g., 所有英语母语者), 我们只能测量一个**样本** (Sample)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

推断性统计 (Inferential Stats)

目的：从样本数据推断总体的特征

我们几乎永远无法测量总体 (e.g., 所有英语母语者)，我们只能测量一个**样本** (Sample)。

核心问题

我们在样本中观察到的差异 (e.g., A 组比 B 组快 50ms) 是真的差异，还是仅仅由**随机抽样** (Random Chance) 造成的？

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

推断性统计 (Inferential Stats)

目的：从样本数据推断总体的特征

我们几乎永远无法测量总体 (e.g., 所有英语母语者)，我们只能测量一个**样本** (Sample)。

核心问题

我们在样本中观察到的差异 (e.g., A 组比 B 组快 50ms) 是真的差异，还是仅仅由**随机抽样** (Random Chance) 造成的？

P 值 (P-value)

- **P 值**是在假设 H_0 (零假设) 为真的前提下，获得当前观察结果（或更极端结果）的概率。
- **通俗理解**：P 值越小，说明我们的发现由“纯粹的运气”造成的可能性越小。
- **阈值 (Alpha)**：通常我们设置 $\alpha = 0.05$ 。
- 如果 $p < 0.05$ ，我们称结果是“统计显著的” (Statistically Significant)，我们**拒绝** H_0 。

方法一：语料库语言学

什么是语料库 (Corpus) ?

- 语料库是大规模、有组织、存储在计算机中的**真实语言数据**集合。
- 它们是“观察性” (Observational) 数据，反映语言的实际使用。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

方法一：语料库语言学

什么是语料库 (Corpus) ?

- 语料库是大规模、有组织、存储在计算机中的**真实语言数据**集合。
- 它们是“观察性” (Observational) 数据，反映语言的**实际使用**。

常见的语料库

- **BNC** (British National Corpus): 英国英语, 1 亿词。
- **COCA** (Corpus of Contemporary American English): 当代美国英语, 10 亿词 (持续更新)。
- **BCC** (北京大学汉语言研究中心语料库): 现代汉语。
- **LIVAC** (香港城市大学 LIVAC 汉语共时语料库)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

语料库的主要分析方法

- 词频 (Frequency)

- 最基本但最强大的工具。
- e.g., 比较 “very” 和 “really” 在不同年代 (如 COCA 1990s vs 2010s) 的使用频率。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

语料库的主要分析方法

- **词频 (Frequency)**

- 最基本但最强大的工具。
- e.g., 比较 “very” 和 “really” 在不同年代 (如 COCA 1990s vs 2010s) 的使用频率。

- **搭配 (Collocation)**

- 经常一起出现的词对。
- e.g., “strong” 经常搭配 “coffee”, “powerful” 经常搭配 “engine”。(MI, t-score)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

语料库的主要分析方法

- **词频 (Frequency)**

- 最基本但最强大的工具。
- e.g., 比较 “very” 和 “really” 在不同年代 (如 COCA 1990s vs 2010s) 的使用频率。

- **搭配 (Collocation)**

- 经常一起出现的词对。
- e.g., “strong” 经常搭配 “coffee”, “powerful” 经常搭配 “engine”。(MI, t-score)

- **关键词 (Keyword)**

- 比较两个语料库 (e.g., 新闻 vs. 学术), 找出在 A 中出现频率异常高 (或低) 的词。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

语料库的主要分析方法

- **词频 (Frequency)**

- 最基本但最强大的工具。
- e.g., 比较 “very” 和 “really” 在不同年代 (如 COCA 1990s vs 2010s) 的使用频率。

- **搭配 (Collocation)**

- 经常一起出现的词对。
- e.g., “strong” 经常搭配 “coffee”, “powerful” 经常搭配 “engine”。(MI, t-score)

- **关键词 (Keyword)**

- 比较两个语料库 (e.g., 新闻 vs. 学术), 找出在 A 中出现频率异常高 (或低) 的词。

- **索引行 (Concordance / KWIC)**

- 查看一个词在真实上下文中的所有实例。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

研究示例 (语料库)

研究问题

在学术写作中，男性和女性作者在第一人称代词（“I”，“we”）的使用上是否存在差异？

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

研究示例 (语料库)

研究问题

在学术写作中，男性和女性作者在第一人称代词 (“I”，“we”) 的使用上是否存在差异？

研究设计

- **语料库：** 建立两个子语料库 (Sub-corpora)
 - 语料库 A： 由男性学者撰写的语言学论文。
 - 语料库 B： 由女性学者撰写的语言学论文。
- **变量：**
 - IV (自变量)： 作者性别 (男 vs. 女)。
 - DV (因变量)： 第一人称代词的标准化频率 (e.g., 每万词)。
- **分析：**
 - 计算两个语料库中 “I” 和 “we” 的频率。
 - 使用 **卡方检验** (Chi-Square test) 或 **Log-Likelihood** (对数似然比) 来比较频率差异是否显著。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

方法二：实验语言学

什么是实验？

- 实验是研究者在高度控制的条件下，通过操纵自变量 (IV)，来观察其对因变量 (DV) 影响的研究方法。
- 目标：建立因果关系 (Causality)。
- 常见领域：心理语言学、实验语音学、神经语言学 (fMRI, EEG)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

方法二：实验语言学

什么是实验？

- 实验是研究者在高度控制的条件下，通过操纵自变量 (IV)，来观察其对因变量 (DV) 影响的研究方法。
- 目标：建立因果关系 (Causality)。
- 常见领域：心理语言学、实验语音学、神经语言学 (fMRI, EEG)。

常见范式 (Paradigms)

- 词汇判断任务 (Lexical Decision Task, LDT)
 - 呈现一个字母串 (e.g., "doctor" 或 "norsup")。
 - 被试尽快判断这是“真词”还是“假词”。
 - DV: 反应时 (RT) 和 准确率。
- 启动任务 (Priming)
- 眼动追踪 (Eye-tracking)
- 自我阅读 (Self-paced Reading)

实验的关键：控制

为什么实验需要控制？

我们需要确保因变量 (DV) 的变化 只 是由自变量 (IV) 引起的，而不是其他无关因素。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

实验的关键：控制

为什么实验需要控制？

我们需要确保因变量 (DV) 的变化 **只** 是由自变量 (IV) 引起的，而不是其他无关因素。

混淆变量 (Confounding Variables)

- 假设：我们想知道“词频”（高频 vs. 低频）是否影响词汇判断 (LDT) 的反应时。
- 操纵 IV = 词频。测量 DV = 反应时。
- **潜在混淆**：如果我们选择的高频词 (e.g., "house") 碰巧比低频词 (e.g., "pylon") 更短，那怎么办？
- 被试反应快，是因为“高频”还是因为“词更短”？
- **控制**：必须确保两组词在 **词长 (Length)** 上匹配 (没有显著差异)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

研究示例 (实验)

研究问题

启动效应 (Priming Effect): 看到一个词 (e.g., "doctor") 是否会加速对相关词 (e.g., "nurse") 的识别?

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

研究示例 (实验)

研究问题

启动效应 (Priming Effect): 看到一个词 (e.g., "doctor") 是否会加速对相关词 (e.g., "nurse") 的识别?

研究设计 (LDT)

- **被试**: 30 名母语为英语的大学生。
- **设计**: 2x2 组间设计 (Within-subject design)
- **IV (自变量)**: 启动词与目标词的关系
 - 条件 1: 相关 (e.g., 启动词 doctor → 目标词 nurse)
 - 条件 2: 不相关 (e.g., 启动词 tree → 目标词 nurse)
- **DV (因变量)**: 对目标词 (e.g., "nurse") 的反应时。
- **假设 (H_1)**: 条件 1 的反应时会显著快于条件 2。
- **分析**: 配对样本 t 检验 (Paired-samples t-test)。

社会语言学 (Sociolinguistics)

- 研究语言使用与社会因素（年龄、性别、阶层、地域）的关系。
- 方法：访谈、调查问卷、变异分析 (Variationist analysis)。
- e.g., Labov 的纽约市 /r/ 音的研究。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

其他方法简介

社会语言学 (Sociolinguistics)

- 研究语言使用与社会因素（年龄、性别、阶层、地域）的关系。
- 方法：访谈、调查问卷、变异分析 (Variationist analysis)。
- e.g., Labov 的纽约市 /r/ 音的研究。

问卷调查 (Surveys)

- 收集关于语言态度 (Language Attitudes)、方言可接受度等主观数据。
- 常用：李克特五点或七点量表 (Likert Scale)。
- 数据类型：通常是 **定序 (Ordinal)** 数据。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 1：识别变量类型

情景

一位研究者想知道以下因素分别是什么类型的变量：（定类，定序，定距，定比）

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 1：识别变量类型

情景

一位研究者想知道以下因素分别是什么类型的变量：（定类，定序，定距，定比）

- (1) 说话人的基频 (F0)，单位是赫兹 (Hz)。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 1：识别变量类型

情景

一位研究者想知道以下因素分别是什么类型的变量：（定类，定序，定距，定比）

- (1) 说话人的基频 (F0)，单位是赫兹 (Hz)。
 - 答案：定比 (Ratio)。(0 Hz 是有意义的真零点)
- (2) 学习者对某个语法结构的可接受度判断（1=完全不可接受, 5=完全可接受）。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 1：识别变量类型

情景

一位研究者想知道以下因素分别是什么类型的变量：（定类，定序，定距，定比）

- (1) 说话人的基频 (F0)，单位是赫兹 (Hz)。
 - 答案：定比 (Ratio)。(0 Hz 是有意义的真零点)
- (2) 学习者对某个语法结构的可接受度判断（1=完全不可接受, 5=完全可接受）。
 - 答案：定序 (Ordinal)。(1 和 2 的差距 $\neq$ 4 和 5 的差距)
- (3) 实验中呈现的图片类型（动物，工具，食物）。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 1: 识别变量类型

情景

一位研究者想知道以下因素分别是什么类型的变量：(定类, 定序, 定距, 定比)

- (1) 说话人的基频 (F0), 单位是赫兹 (Hz).
 - 答案: 定比 (Ratio)。 (0 Hz 是有意义的真零点)
- (2) 学习者对某个语法结构的可接受度判断 (1=完全不可接受, 5=完全可接受).
 - 答案: 定序 (Ordinal)。 (1 和 2 的差距 $\neq$ 4 和 5 的差距)
- (3) 实验中呈现的图片类型 (动物, 工具, 食物).
 - 答案: 定类 (Nominal)。 (类别之间没有顺序)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

请大家思考 (做中学)

- 你会选择什么研究方法？(实验？语料库？问卷？)

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

请大家思考 (做中学)

- 你会选择什么研究方法？(实验？语料库？问卷？)
- 方法：实验。
- 自变量 (IV) 是什么？

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

请大家思考 (做中学)

- 你会选择什么研究方法？(实验？语料库？问卷？)
- 方法：实验。
- 自变量 (IV) 是什么？
- IV：阅读媒介（手机 vs. 纸质书）。
- 因变量 (DV) 是什么？

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

请大家思考 (做中学)

- 你会选择什么研究方法？(实验？语料库？问卷？)
- 方法：实验。
- 自变量 (IV) 是什么？
- IV：阅读媒介（手机 vs. 纸质书）。
- 因变量 (DV) 是什么？
- DV：阅读理解测试的得分。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

- 你需要控制哪些混淆变量？

练习 2：设计一个简单的研究

研究问题

“相比于纸质书阅读，在手机上阅读是否会降低阅读理解能力？”

请大家思考 (做中学)

- 你会选择什么研究方法？(实验？语料库？问卷？)
- 方法：实验。
- 自变量 (IV) 是什么？
- IV：阅读媒介（手机 vs. 纸质书）。
- 因变量 (DV) 是什么？
- DV：阅读理解测试的得分。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

- 你需要控制哪些混淆变量？

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

统计分析 (独立样本 t 检验)

$t(98) = 2.45, p = 0.016$

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

统计分析 (独立样本 t 检验)

$t(98) = 2.45, p = 0.016$

问题：我们能得出什么结论？

- (1) 两组之间有差异吗？

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

统计分析（独立样本 t 检验）

$t(98) = 2.45, p = 0.016$

问题：我们能得出什么结论？

- (1) 两组之间有差异吗？
 - 有。纸质书组 (79.8) 平均分高于手机组 (75.5)。

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

统计分析（独立样本 t 检验）

$t(98) = 2.45, p = 0.016$

问题：我们能得出什么结论？

- (1) 两组之间有差异吗？
 - 有。纸质书组 (79.8) 平均分高于手机组 (75.5)。

练习 3：解读一个（虚构的）结果

情景

我们做了练习 2 的实验，比较“手机组”和“纸质书组”的阅读理解得分（满分 100）。

组别	平均分 (Mean)	标准差 (SD)
手机组 (n=50)	75.5	8.2
纸质书组 (n=50)	79.8	7.9

统计分析（独立样本 t 检验）

$t(98) = 2.45, p = 0.016$

问题：我们能得出什么结论？

- (1) 两组之间有差异吗？
 - 有。纸质书组 (79.8) 平均分高于手机组 (75.5)。

本讲总结

- 为什么量化：客观性、可重复性、可推广性。
- 核心概念：
 - 研究问题 与 假设 (H_0 / H_1)
 - 变量 (IV / DV)
 - 数据类型 (分类 / 连续)
 - 统计 (描述性 / 推断性) 与 P 值
- 主要方法：
 - 语料库 (观察性)：词频、搭配、关键词。
 - 实验 (控制性)：LDT、启动，寻找因果关系。
- 做中学：量化思维是关于如何设计、测量和比较。

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

接下来我们将深入学习...

- **统计基础**：t 检验, 卡方检验, ANOVA (方差分析)
- **相关与回归** (Regression)：预测一个变量如何影响另一个变量。
- **R 语言入门**：用于统计分析和可视化的强大工具。
- **混合效应模型** (Mixed-effects Models)：现代语言学研究的标配。
- **如何阅读/撰写一篇量化研究论文。**

(C) <https://languages.lol/> | guoshujian@tongji.edu.cn

谢谢 Q & A (guoshujian@tongji.edu.cn)