

A/B Test 入门

讲师：吴鸣夏

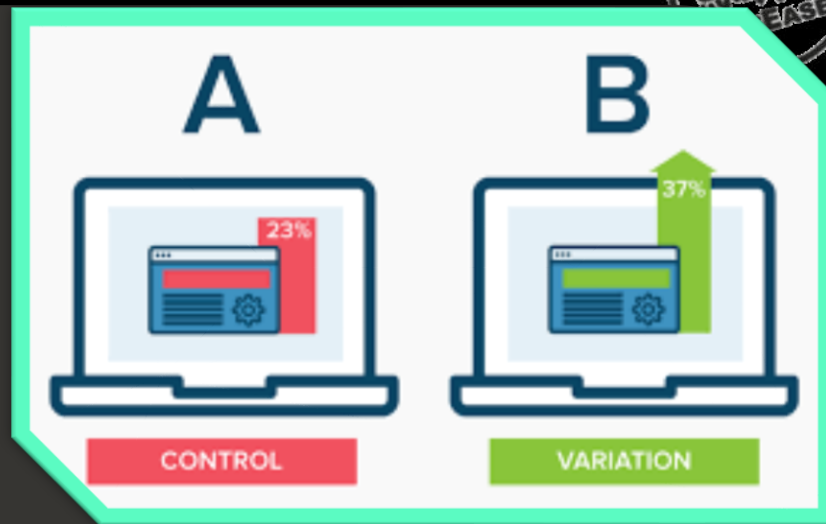
背景介绍

◆ A/B Test的作用:

- 来源于双盲实验
- 互联网应用场景: 算法、功能、设计

◆ A/B Test定义:

A/B Test将不同的用户分成不同的组（实验组、测试组），同时测试不同的方案，基于用户反馈的真实数据和假设检验，来找出哪一个方案更好的过程。



A/B 实验

◆ A/B test基础原理:

- 重要概念: 置信度、随机性、先验性、并行性。
- 运用统计学知识以及基础计算逻辑

◆ A/B test 基本流程:

推测假设 & 预设目标 → A/B test策略 → 指标设计 & 数据收集方案 →
A/B test实现 → A/B test分析 → 扩量/迭代



◆ A/B test基本流程：

Step1. **分组**：用户/流量分流机制：你是什么时候被分到了T/C组；

Step2. **TC实现**：不同用户如何看到不同界面；

Step3. **使用反馈**：对不同策略，用户有哪些关键反馈（业务假设对应的核心路径），同时决定了埋点的内容；

Step4. **数据收集 & 处理**：如何从日志中获取有效数据（分组表&行为数据表）。

A/B test入门

- 随机性检验：AA/ 前置检验

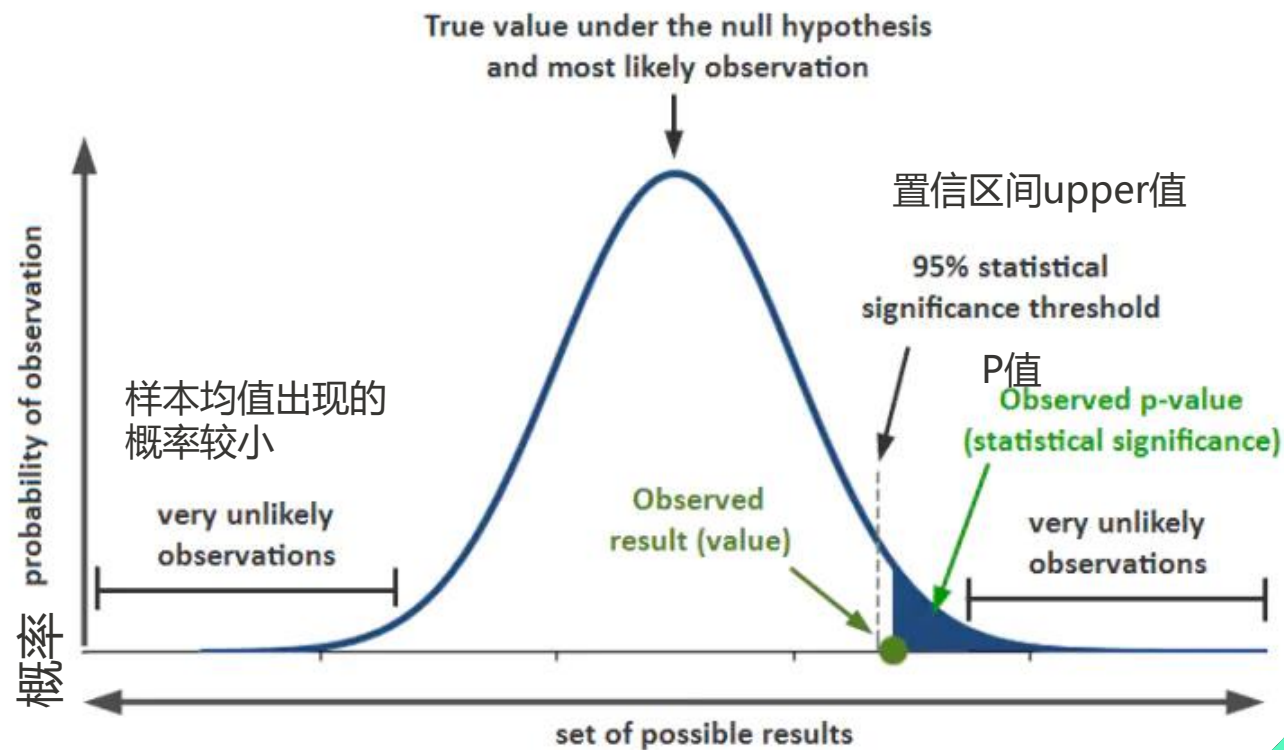
- AA实验：“空转”一段时间，t检验
- 前置检验：invariant/variant metric

- 显著性检验：

- 指标类型：取值型、比例型
- 验证方法：置信区间、假设检验

置信区间 = $\text{delta} \pm 1.96 * \text{标准差}$

Probability & Statistical Significance Explained



Step4. 数据收集 & 处理：如何从日志中获取有效数据（数据流转链路）。

A/B test入门 - 数据分析

1. 实验配置:

实验组 (T, 5%)



对照组 (C, 95%)



2. 实验结果数据

	浏览人数	点击人数
Test	2000	130
Control	38000	2160

3. 实验数据分析 (95% 置信度)

$$CI = \text{delta} \pm 1.96 * SD$$

CI: 置信区间

$$\text{Delta} = P_t - P_c$$

$$SD: \text{标准差} \sqrt{\frac{P_t \times (1 - P_t)}{N_t} + \frac{P_c \times (1 - P_c)}{N_c}}$$

N: 样本数

P: 百分比

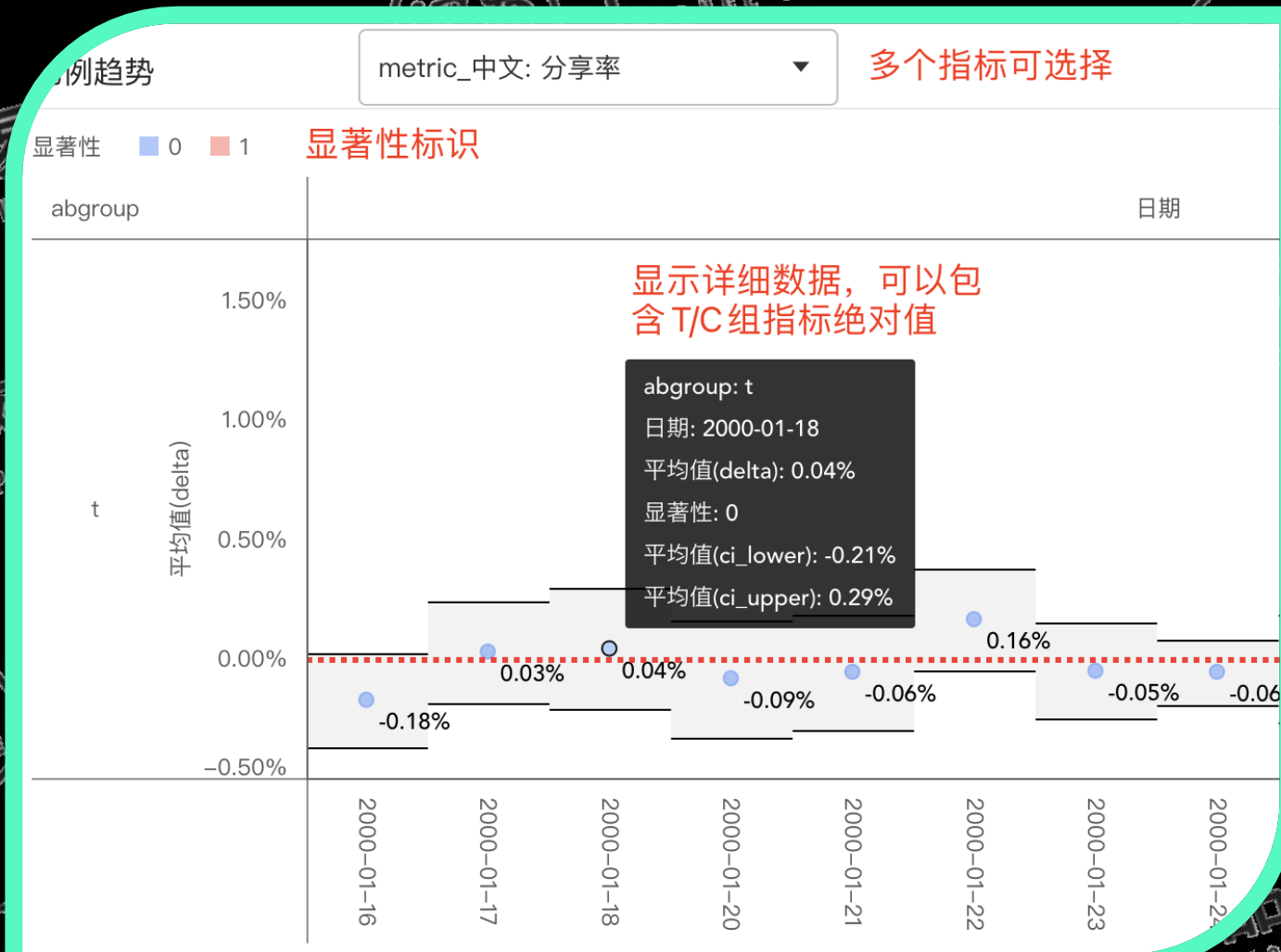


A/B test入门 – 可视化

数据可视化：一图胜千言

分享按钮实验处理数据

Metric	Value
N_t	2000
N_c	38000
P_t	6.50%
P_c	5.68%
Delta	0.82%
CI_upper	1.92%
CI_lower	-0.29%



高级玩法：复杂实验

◆ 交叉实验的实现，层/域：

- 分层 Layer：每层独立分组，依赖关系

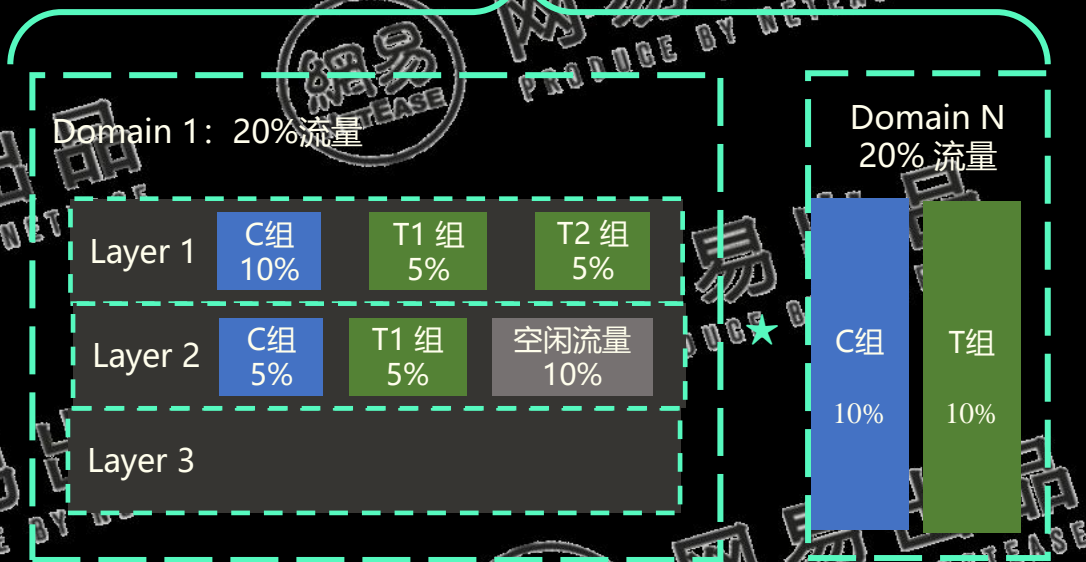
的实验之间同层不同实验

- 分域 Domain：流量完全分隔，一个请求只分配到一个domain

◆ 小流量实验：

- 实验最小样本数
- 小流量实验处理方式

Domain1 + Domain 2+ ...+ Domain N = 100% 流量



AB 实验：回顾 & 思考

回顾

Step1. 分组 → Step2. TC实现 →
Step3. 使用反馈 → Step4. 数据收集 & 处理 → Step5. 可视化 & 报告
呈现

思考

- ◆ AB实验的结果是显著正/负，是否能判断 要不要 全量实验？
- ◆ 哪些场景不适合AB 实验？

课程回顾

内容/服务

主动获取
被动推荐

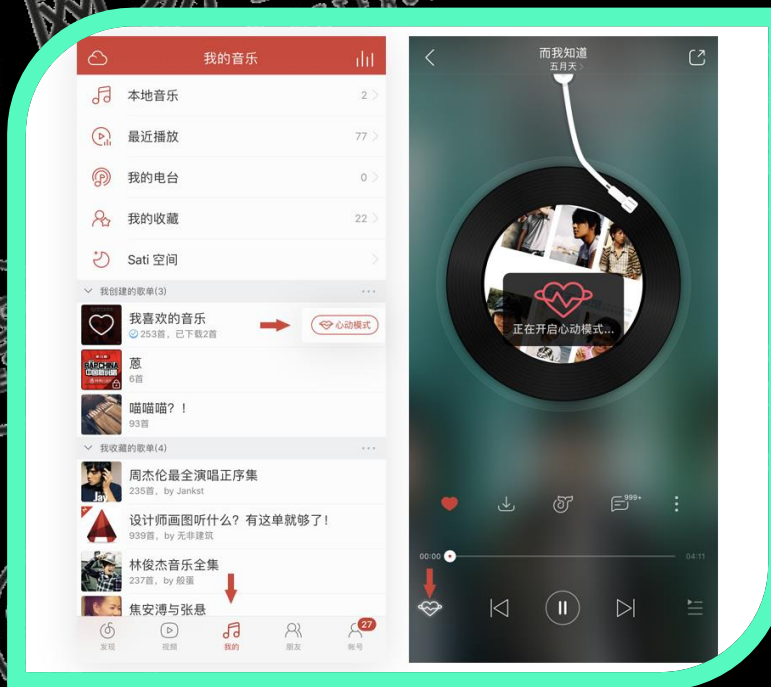
变现

发商业
流量分

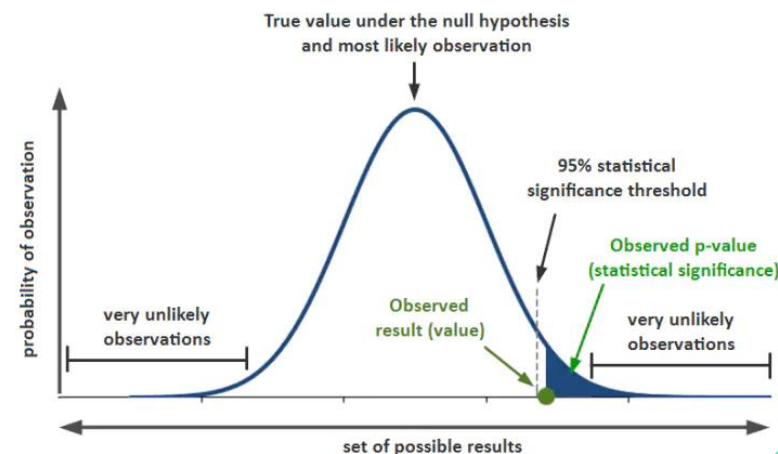
提供内容

消费反馈

平台



Probability & Statistical Significance Explained



谢谢观看

