统计学：决策的科学项目说明

**说明：**[**点此查看此文档的英文版本**](https://s3.cn-north-1.amazonaws.com.cn/static-documents/nd002/StatisticsTheScienceofDecisions-ProjectInstructions.pdf)**。**

**背景信息**

在一个 Stroop （斯特鲁普）任务中，参与者得到了一列文字，每个文字都用一种油墨颜色展示。参与者的任务是将文字的打印颜色大声说出来。这项任务有两个条件：一致文字条件，和不一致文字条件。在一致文字条件中，显示的文字是与它们的打印颜色匹配的颜色词，如“红色”、“蓝色”。在不一致文字条件中，显示的文字是与它们的打印颜色不匹配的颜色词，如“紫色”、“橙色”。在每个情况中，我们将计量说出同等大小的列表中的墨色名称的时间。每位参与者必须全部完成并记录每种条件下使用的时间。

**调查问题**

作为一般说明，请确保记录你在创建项目时使用或参考的任何资源。作为项目提交的一部分，你将需要报告信息来源。

1. 我们的自变量是什么？因变量是什么？

自变量(dependent variable) :文字和颜色是否一致，

因变量(independent variable): 说出颜色的时间

1. 此任务的适当假设集是什么？你想执行什么类型的统计测试？为你的选择提供正当理由。

零假设(null hypothesis): 识别文字和颜色一致(congruent)的时间 大于或等于 识别颜色和文字不一致(incongruent)的时间

对立假设 (alternative hypothesis): 识别文字和颜色一致的时间 小于 识别颜色和文字不一致的时间

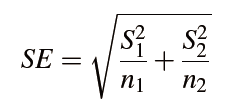
相依样本 t 检验 (Dependent-samples t-test) 单尾测试

同样的人在条件改变先后的测试，符合依赖性

只需要样本数据，符合t测试

现在轮到你自行尝试 Stroop 任务了。前往[此链接](https://faculty.washington.edu/chudler/java/ready.html)，其中包含一个基于 Java 的小程序，专门用于执行 Stroop 任务。记录你收到的任务时间（你无需将时间提交到网站）。现在[下载此数据集](https://s3.cn-north-1.amazonaws.com.cn/static-documents/nd002/stroopdata.csv)，其中包含一些任务参与者的结果。数据集的每行包含一名参与者的表现，第一个数字代表他们的一致任务结果，第二个数字代表不一致任务结果。

1. 报告关于此数据集的一些描述性统计。包含至少一个集中趋势测量和至少一个变异测量。
2. 求出均值标准差



均值标准差 SE = 1.2193

t-test 统计值 = ）/ SE = -6.53

1. t-test 临界值 alpha = 0.05 , 单尾， df = (n1 + n2 - 2) = 46

查询t-table得知，对应的临界值为 -1.684 到 -1.676

1. t-test 统计值远大于临界值，说明识别文字和颜色一致的时间 小于识别颜色和文字不一致的时间, 有统计学意义

相关计算过程代码见同目录下的testsa.py

1. 提供显示样本数据分布的一个或两个可视化。用一两句话说明你从图中观察到的结果。

图片来自 <http://www.math.uah.edu/stat/apps/MeanTestExperiment.html>

该图粗略表示t-test的结果

1. 现在，执行统计测试并报告你的结果。你的置信水平和关键统计值是多少？你是否成功拒绝零假设？对试验任务得出一个结论。结果是否与你的期望一致？
2. 可选：你觉得导致所观察到的效应的原因是什么？你是否能想到会取得类似效应的替代或类似任务？进行一些调查研究将有助于你思考这两个问题！

优达学城

2016年9月