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：文字列表的呈现方式，一致或不一致

因变量：参与者完成每种条件下任务所使用的时间

1. 此任务的适当假设集是什么？你需要以文字和数学符号方式对假设集中的零假设和对立假设加以说明，并对数学符号进行定义。你想执行什么类型的统计检验？为你的选择提供正当理由（比如，为何该实验满足你所选统计检验的前置条件）。

设H0零假设，HA为对立假设，μa为参与者完成一致性条件文字列表所使用的平均时间，μb为参与者完成不一致性条件文字列表所使用的平均时间，

则

H0：μa = μb

HA：μa < μb

由于该任务属于实验研究，对有限同一样本先后应用两种不同处理方法（一致性文字列表和不一致性文字列表），符合t检验的前置条件，因此考虑执行t统计检验。同时，由于对立假设拟设为μa < μb，故采用单尾检验。

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。

集中趋势测量

（1）均数：

一致条件下的均数C = 14.05

不一致条件下的均数I = 22.02

（2）中位数：

一致条件下的中位数 MC = 14.3565

不一致条件下的中位数 MI = 21.0175

变异测量

一致条件下的标准差 SC = 3.48

不一致条件下的标准差 SI = 4.70

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。

上图为一致性条件下使用时间的频次分布图，组距为3。从图中可以看出，参与者在一致性条件下所使用的时间大致呈正态样分布，多数参与者所用时间在13.5左右，该样本中位数和均数都在12～15这一组内。

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？

置信水平α = 0.05，自由度df = n-1 = 23，关键统计值 tcretical = 1.714，单尾

设：两种条件下使用时间差异的均数D = C - I = -7.97

两种条件下使用时间差异的标准差 = 4.76

则：统计量 tstatistic = = = -8.203

因为统计量tstatistic < -tcretical ，p<0.05，参加者在不同条件下所使用的时间有明显差异

所以拒绝零假设H0，接受对立假设HA。

由试验任务可以得出结论：参与者在一致性条件下所使用的时间少于不一致性条件下所使用的时间。

该结论与我的期望一致。

1. 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