

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](#)。

背景信息

在一个 Stroop（斯特鲁普）任务中，参与者得到了一系列文字，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？

自变量是文字意义与颜色的一致性，分为一致与不一致两种情况。因变量是被试说出所有文字颜色所花费的时间，简称时间。

2. 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）。

1) 此任务的适当假设集是：

a) 零假设 (H_0)：文字意义与颜色一致性对于被试说出所有文字颜色所花费的时间

没有影响或者不一致情况会导致花费时间变少，也就是说 $\mu_c - \mu_i \geq 0$ 。其中， μ_c 表示文

字意义与颜色一致情况下，所有被试说出所有文字颜色所花费的平均时间， μ_i 表示文字意义与颜色不一致情况下，所有被试说出所有文字颜色所花费的平均时间。

b) 非零假设 (H_1)：文字意义与颜色不一致会使人说出所有文字颜色所话费时间增

多，即 $\mu_c - \mu_i < 0$ 。

2) 此处应使用 t 检验。因为我们无法获得总体（所有人）完成这个任务的情况，也就无法统计出总体（所有人）完成这个任务所花费时间的均值之类的统计量，只能获得样本（参与实验的人）完成这个任务所花费时间的均值、标准差之类的统计量，这种情况适合使用 t 检验开展假设检验。其次这个实验是比较在两种情况下（一致与不一致）人们完成任务所花费时间的均值差异情况，更适合用 t 检验。再次此项实验是一个重复

衡量设计，是在两种不同情况下开展同样的实验任务，这种实验适合使用相依样本 t 检验来进行假设检验。最后，想要验证的结论应该是“当文字颜色与文字意义不一致时，人说出文字颜色所花费的时间较一致情况下长，即 $\mu_c - \mu_i < 0$ ”，因此采用负向单尾 t 检验更合适。

现在轮到你自行尝试 Stroop 任务了。前往[此链接](#)，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](#)，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。

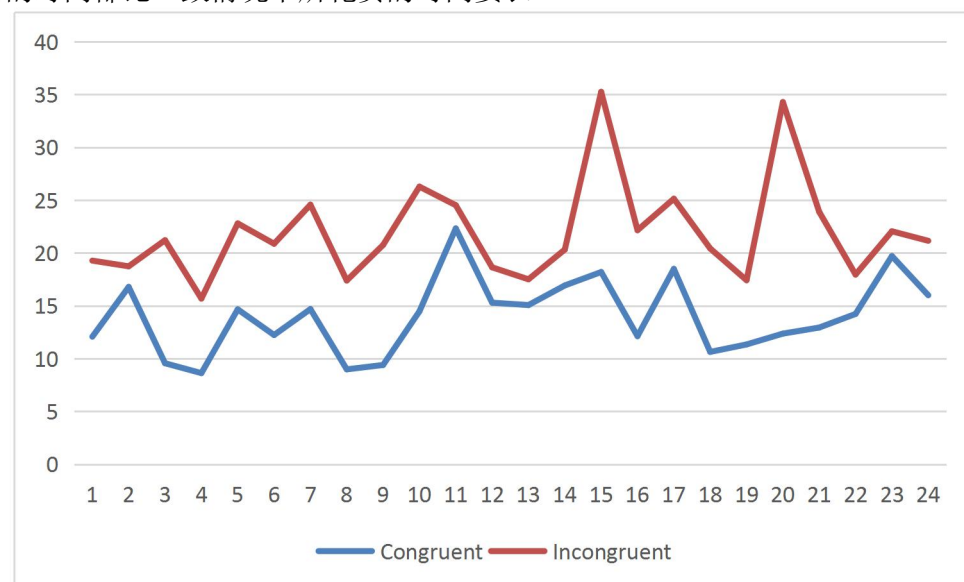
3. 报告关于此数据集的一些描述性统计。包含至少一个集中趋势测量和至少一个变异测量。

1) 文字颜色与文字意义一致情况下，所有被试者说出文字颜色所花费时间的平均值 $\mu_c = 14.05$ ，中位数 $m_c = 14.36$ ，标准偏差 $\sigma_c = 3.56$ ；

2) 文字颜色与文字意义不一致情况下，所有被试者说出文字颜色所花费时间的平均值 $\mu_i = 22.02$ ，中位数 $m_i = 21.02$ ，标准偏差 $\sigma_i = 4.797$ ；

4. 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。

1) 如下图所示，所有被试者在文字颜色与文字意义不一致情况下说出文字颜色所花费的时间都比一致情况下所花费的时间要长。



2) 如下图所示，一致情况与不一致情况下说出文字颜色所花费的时间都是正偏斜的正太分布，并且一致情况下所花费时间的均值比不一致情况下所花费时间的均值要小。



5. 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？

在置信水平 $\alpha = 0.01$ ，自由度 $df = n_c + n_i - 2 = 46$ ，执行负向单尾独立样本 t 检验的情况下，关键统计量 $t_{critical} \approx -2.403$ 。而 $\mu_c - \mu_i = 14.05 - 22.02 = -7.97$ 。又

$$SS_c = \frac{\sum (x_j - \mu_c)^2}{df_c} = \frac{291.39}{23} = 12.67, \quad SS_i = \frac{\sum (y_j - \mu_i)^2}{df_i} = \frac{529.27}{23} = 23.01, \quad \text{因此}$$

$$S_p^2 = \frac{SS_c + SS_i}{df_c + df_i} = \frac{35.68}{46} = 0.78, \quad \text{由此得到标准误差 } SE = \sqrt{\frac{S_p^2}{n_c} + \frac{S_p^2}{n_i}} = 0.25。最后，得到 t 统计$$

$$\text{量为 } t = \frac{\mu_c - \mu_i}{SE} = \frac{-7.97}{0.25} = -31.88, \quad \text{在关键区内，具有显著性，因此成功拒绝零假设，得出}$$

结论：文字颜色与文字意义不一致情况下，人正确读出文字颜色需要花费更多时间。这个结论与我的期望是一致的。另外计算出的 p 值小于 0.0001，也就是说不一致情况下花费时间更少或者与一致情况下花费时间差不多的情况有极小概率才会出现。

6. 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！

参考资料列表：

- 1、[斯特鲁普效应](#)，[百度百科](#)