

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：任务的两个条件，一致文字条件 / 不一致文字条件。

因变量：说出同等大小的列表中的墨色名称的时间。

2. 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。

$\mu_{\text{一致}}$: 不一致文字条件下说出时间的样本平均值

$\mu_{\text{不一致}}$: 一致文字条件下说出时间的样本平均值

零假设：文字条件对说出同等大小的列表中的数据没有影响。

对立假设：不一致文字条件会增加说出时间。

$H_0: \mu_{\text{一致}} = \mu_{\text{不一致}}$

$H_1: \mu_{\text{一致}} < \mu_{\text{不一致}}$

由于同一测试者参加了两次测试且不知道总体，因此我们选择相依样本 t 检验。

相依样本 t 检测的假设前提是：

1.Interval or ratio scale of measurement

2.Random sampling from a defined population

3.Samples or sets of data used to produce the difference scores are linked in the population through repeated measurement, natural association, or matching.

4.Scores are normally distributed in the population; difference scores are normally distributed

3. 总体数据服从正态分布，且不同的数据服从不同的正太分布。

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。

样本均值：

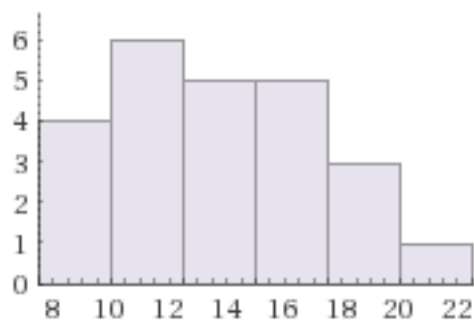
$$\bar{X}_{\text{一致}} = 14.05, \bar{X}_{\text{不一致}} = 22.02$$

差异点估计： $\mu_{\text{一致}} - \mu_{\text{不一致}} = -7.96$

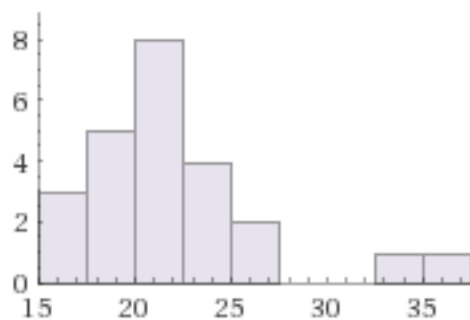
差异点标准偏差： $S = 4.86$

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。

一致文字条件下样本数据分布直方图：



不一致文字条件下样本数据分布直方图：



从上面两个图可以看出一致条件下样本数据更集中，不一致文字条件下样本数据分布更分散，不一致条件下说出的数据的样本平均值要大于一致条件下说出的样本平均值。

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？

t- statistic = -8.02

df = 23

t- critical for one-tailed test at $\alpha = 0.05$:

t-critical = -1.714 \rightarrow reject H_0

CI : (5.52, 10.42)

统计报告如下：

$t(23) = -8.02, p < 0.001$, one-tailed in the negative direction

95%CI(-9.66,-6.26)

结果和期望的一致。

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！

参考的网站：

<http://www.graphpad.com/quickcalcs/pValue1/>

<http://www.wolframalpha.com/>