

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(dependent variable): 文字和颜色是否一致，

因变量(independent variable): 说出颜色的时间

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。

1) 假设集

零假设(null hypothesis): 识别文字和颜色一致(congruent)时间的样本均值 \bar{x}_c 等于 识别颜色和文字不一致(incongruent) 时间的样本均值 \bar{x}_i

$$H_0: \bar{x}_c - \bar{x}_i = 0$$

对立假设 (alternative hypothesis): 识别文字和颜色一致时间的样本均值 \bar{x}_c 小于 识别颜色和文字不一致的时间的样本均值 \bar{x}_i 这是单尾测试

$$H_1: \bar{x}_c - \bar{x}_i < 0$$

2) 统计测试的类型

相依样本 t 检验 (Dependent-samples t-test)

选用 t 检验而不是 z 检验的原因：

不知道全体的标准差，只有样本数据，无法使用 z 检验，只能选用 t 检验

参考 <http://www.statisticshowto.com/when-to-use-a-t-score-vs-z-score/>

执行相依样本的原因：

对同样的观察对象在不同的条件下做了重复的测量，这些前后测量的数据是有关联的，所以是两个相依样本而不是独立样本

样本呈正态分布

参考 <http://www.psychology.emory.edu/clinical/bliwise/Tutorials/TOM/meanstests/assump.htm>

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

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。

以下是样本数据信息：

n	Congruent	Incongruent	Diff
1	12.079	19.278	-7.199
2	16.791	18.741	-1.95
3	9.564	21.214	-11.65
4	8.63	15.687	-7.057
5	14.669	22.803	-8.134
6	12.238	20.878	-8.64
7	14.692	24.572	-9.88
8	8.987	17.394	-8.407
9	9.401	20.762	-11.361
10	14.48	26.282	-11.802
11	22.328	24.524	-2.196
12	15.298	18.644	-3.346
13	15.073	17.51	-2.437
14	16.929	20.33	-3.401
15	18.2	35.255	-17.055
16	12.13	22.158	-10.028
17	18.495	25.139	-6.644
18	10.639	20.429	-9.79
19	11.344	17.425	-6.081
20	12.369	34.288	-21.919
21	12.944	23.894	-10.95
22	14.233	17.96	-3.727
23	19.71	22.058	-2.348
24	16.004	21.157	-5.153

1) Calculate the mean and standard deviation of the differences gives

$$\bar{d} = -7.96$$

$$S = 4.86$$

2) Calculate standard error of the mean difference,

$$SEM = S / (n^{0.5}) = 4.86 / (24^{0.5}) = 0.99$$

3) Calculate the t-statistic

$$t\text{-stat} = \bar{d} / SEM = -8.02$$

$$df = n - 1 = 23$$

查询 t-table 可知 $p < 0.0005$, p 小于 $\alpha = 0.05$

4) Calculate the t-critical

$df = n - 1 = 23$, $\alpha = 0.05$ 单尾, 查询 t-table 得知

$$t\text{-critical} = -1.714$$

5) **t-test** 统计值远大于临界值, 说明识别文字和颜色一致时间的样本均值小于识别颜色和文字不一致的时间样本均值, 有统计学意义

6) Confidence interval for the true mean difference

双尾 t 临界值 $\alpha = 0.025$, $df = 23$, 查询 t 表得知

$$t\text{-critical } 2 = 2.069$$

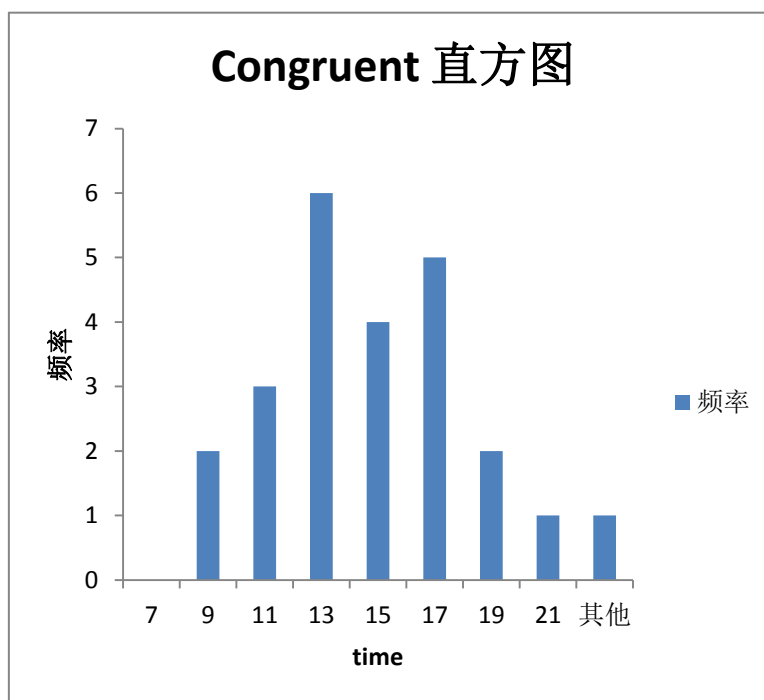
$$\text{置信区间 } CI = (\bar{d} - t\text{-critical } 2 * SEM, \bar{d} + t\text{-critical } 2 * SEM)$$

$$\text{计算结果为 } CI = (-10.02, -5.91)$$

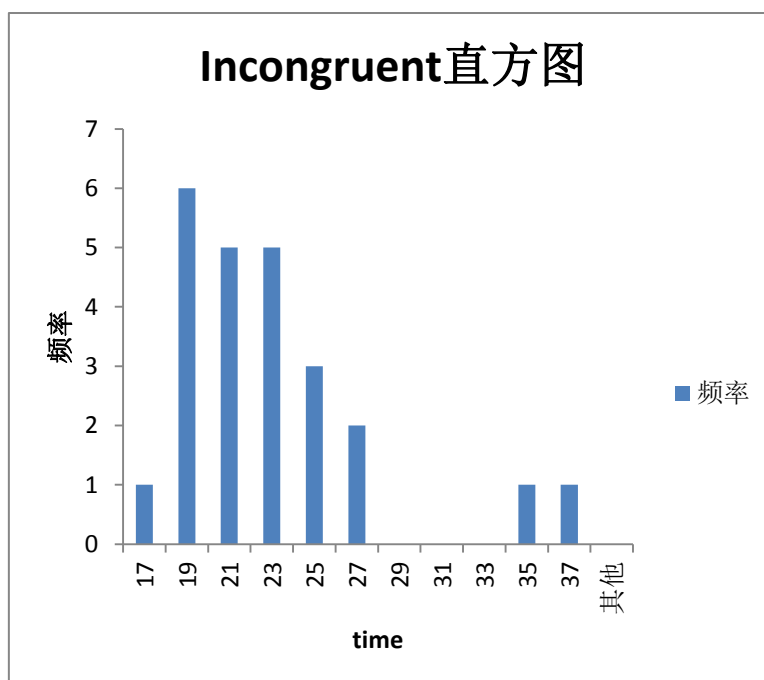
相关计算过程代码见同目录下的 testsa.py

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。

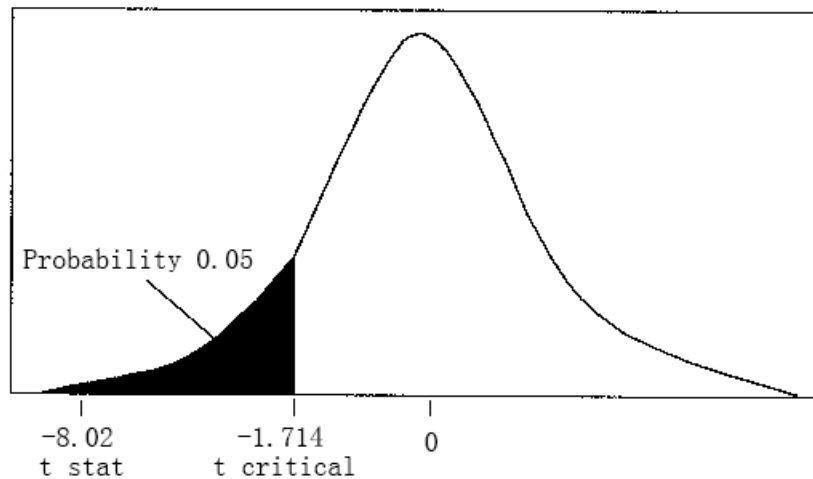
T 检测的前提条件是样本是正态分布。下图是 Excel 绘制的两个样本的频率直方图 Congruent 的直方图接近正态分布



Incongruent 的直方图，分布类似 Positively skewed，当样本较大时可以认为其接近正态分布



下图粗略表示 t-test 的结果



由上图可知，t 统计值落在临界区中，reject 零假设，对立假设成立，即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

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？

置信等级： 95%

t 临界区： -1.714

t 统计值： -8.02

reject 零假设，对立假设成立，即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

试验结论：视觉对颜色的判断和文字识别不一致时，会延长决策时间
结果与期望一致

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！

可能原因：视觉对颜色的判断和文字识别不一致时，会干扰大脑判断，使之延长决策时间

还可以添加两个对照组。

1) 只识别颜色而不识别文字，即用无关颜色的文字来表示颜色。

2) 只识别文字而不识别颜色，用同样的颜色的文字显示

识别结果如下，

文字识别的速度 > 文字和颜色一致 > 颜色识别的速度 > 文字和颜色不一致

可能原因，文字可以直接映射抽象概念，而颜色需要匹配的时间，当颜色和文字不一致时，大脑需要逻辑判断