

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：文字和颜色是否一样。

因变量：说出同等大小的列表中的颜色名称的时间。

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）。

a. 任务的假设集

$\mu_{\text{congruent}}$ ：文字和颜色一致的条件下，识别出颜色时间的总体均值

$\mu_{\text{incongruent}}$ ：文字和颜色不一致的条件下，识别出颜色时间的总体均值

H_0 （零假设）： $\mu_{\text{incongruent}} = \mu_{\text{congruent}}$ ，文字和颜色不一致的条件下，识别出颜色时间的总体均值相等；

H_a （对立假设）： $\mu_{\text{incongruent}} \neq \mu_{\text{congruent}}$ ，文字和颜色不一致的条件下，识别出颜色时间的总体均值不相等；

预测结论：

	Reject H_0	Retain H_0
H_0 (true)	I 型错误	保持零假设
H_0 (false)	拒绝零假设	II 型错误

b. 执行的统计检验类型

使用 **Paired T-Test**，t 检验的双尾检验， α level 为 0.05

c. 执行该统计检验的理由：

1、由 z 检验（知道总体参数）和 t 检验（不知道总体参数）的概念可以得知，由于此实验不确定总体参数，且样本量 <30 ，所以选择 t 检验；

2、假设验证是否两个总参相等或不相等，所以选择双尾检测；

3、由于是同一个测试者参与的不同测试，所以是相依样本；

4、t 检验假设的前提：

a、自变量的两样本是配对的，观察值得数目相同、顺序不可随意更改

b、连续因变量

c、样本接近正太分布

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。

集中趋势测量结果：

	均值	中位数
一致的文字条件	14.051125	14.3565
不一致的文字条件	22.01591667	21.0175

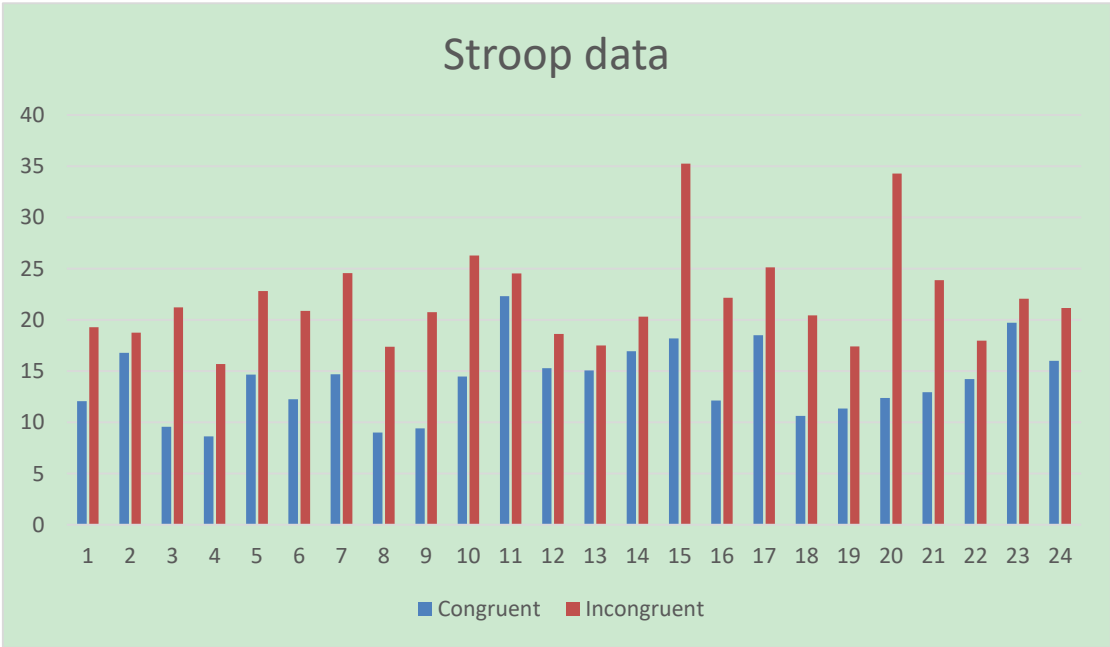
IQR（变异测量）结果：

	IQR
一致的文字条件 $IQR_{congruent}$	4.686
不一致的文字条件 $IQR_{incongruent}$	5.52

标准偏差结果：

	使用 Excel 中 STDEV.S 公式
一致的文字条件	3.5594
不一致的文字条件	4.7971

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。



综上所述，由上图得到 **Incongruent** 的数据全部高于 **congruent** 条件的阅读时间。

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？

1、计算两组数据的均值差异：

x: 代表 congruent

y: 代表 Incongruent

$$\text{公式: } \bar{d} = \frac{\sum_{i=1}^n (y_i - x_i)}{n}$$

计算结果: $\bar{d} = 7.96479167$

2、计算 t 临界值：

$\alpha = 0.05$, 统计类型: 双尾检验, 自由度 $df = n-1 = 23$

查 t 表格, 得到结果: $t_{\text{critical}} = 2.069$

3、计算差异的样本标准偏差：

$$\text{公式: } S_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

计算结果: $S_d = 4.86482691$

4、计算均值的标准误差：

$$\text{公式: } SEM = \frac{S}{\sqrt{n}}$$

计算结果: $SEM = 0.993029$

5、计算 t 统计量：

$$\text{公式: } t_{\text{statistic}} = \frac{\bar{d} - 0}{SEM}$$

计算结果: $t_{\text{statistic}} = 8.02070694$

发现: $t_{\text{statistic}} > t_{\text{critical}}$

$p < 0.05$ 统计结果显著, 得出结论:

成功拒绝零假设情况, 即: 文字和颜色不一致的条件下, 使得正确识别文字颜色的时间变长。