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：文字是否和颜色一致

因变量：说出同等大小列表中的墨色名称的时间

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）。

其中x代表文字条件一致，y代表文字条件不一致。

Ux:x的总体均值（第一组实验的总体均值）

Uy:y的总体均值（第二组实验的总体均值）

H0假设：读文字打印颜色，文字条件一致和文字条件不一致所用时间相同，即Ux=Uy

H1假设：读文字打印颜色，文字条件颜色不一致所花费的时间要比文字颜色一致所花费的时间长。即Uy>Ux

我准备用配对单边t检验：

1. 两组实验的自变量是双变量，两次实验所参与的人是相同的，并非是两个独立样本，配对 t 检验计算每对前后测量值之间的差异，确定这些变化的均值，并报告差异均值在统计上是否显著。
2. 两组样本数据（因变量）未严重违反正态假设。
3. 两组数据的因变量是连续的。
4. 因变量的每一个观察独立于因变量的其他观察，即因变量之间互不影响。
5. 只需证明第二组总体均值大于第一组总体均值，所以用单边正向检验即可。
6. t分布是z分布的小样本分布，即当总体符合z分布时，从总体中抽取的小样本符合t分布，而对于符合t分布的变量，当样本量增大时，变量数据逐渐向z分布趋近。

t检验主要用于样本含量较小（例如n<30），总体标准差σ未知的正态分布资料。z检验是一般用于大样本(即样本容量大于30)平均值差异性检验的方法。所以t检验比较适合该测验。

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。

Congruent的相关数据：

中心性测量：

中位数=14.3565

众数：由于数据不存在一致情况，所以不存在众数。

平均数=14.05113

可变性测量：

IQR=4.686

极差=13.698

样本方差= 12.14115

样本标准偏差=3.559358

Incongruent的相关数据：

中心性测量：

中位数：21.0175

众数：由于数据不存在一致情况，所以不存在众数。

平均数：22.01592

可变性测量：

IQR=5.5165

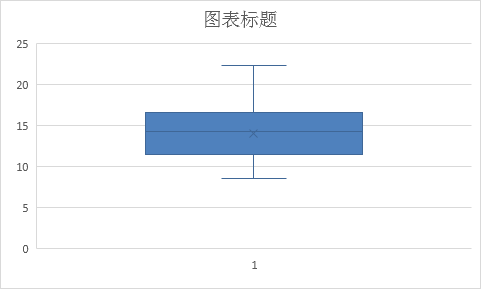
极差=19.568

样本方差= 22.05293

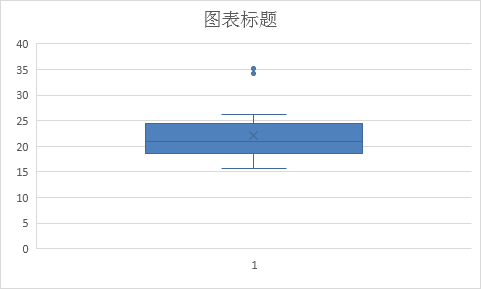
样本标准偏差：4.797057

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。

X轴的数字代表不同的人，Y轴代表每人所用的时间，由图表得出，该样本中，文字条件不一致说出墨色时间高于条件一致时说出墨色时间。



此图为congruent的箱装图中位数大概为14左右，IQR为4左右，最大值为23左右，最小值为8左右



此图为incongruent的值，中位数位于22左右，IQR大概为5左右，偏离的点为异常值。除去异常值的最大值为26左右，最小值为15.5左右

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？

点估计差异= 7.964792

差异的标准偏差= 4.864827

t统计量= 8.020707

a=0.05

t临界值=1.714

结论：由于t统计量位于拒绝域内，所以检验的P值小于临界P值，原假设不成立，即文字条件一致读取时间小于文字条件不一致读取时间。结论与我观察图表所得出的结论一致。

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！

原因：关于斯特鲁普百度百科的解释是：优势反应对非优势反应的干扰，当测试者被要求回答有颜色意义的字体的颜色时，回答字本身的意义为优势反应，而回答字体颜色为非优势反应，若字体颜色与自意不同，被测者往往会反应速度下降，出错率上升。

而我个人所理解的是：当对某一件事物的认知存在冲突时，选取合适的认知会消耗一定时间，所以会存在文字与颜色条件不一致时会用时较长，当把文字换成英文或者其他小语种的语言时，所用时间不会存在显著性差异。

类似的任务：选取一些几何图形，正方形，等边三角形，等腰三角形，圆形，椭圆，长方形，在图形里面写上图形名称。

A组：图形名称和真实图形名称对应

B组：图形名称和真实名称不对应。

让两组人员念出图形里面的字，记录两组人员的反应时间。

类似下图（可以多放一些图形）

A组：

三角形

三角形三角形

正方形

B组：

正方形

圆形

优达学城

2016年9月