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？

因变量是参与者说出墨色名称的时间，自变量是墨色和文字是否一致的类型

1. 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。

零和假设(H0):墨色和文字是否一致对参与者使用的时间没有影响，即μ1 = μ2

备选假设(Ha): 墨色和文字是否一直对参与者使用时间有影响，即μ1 ≠ μ2

因为在本试验中初始时是不确认前后两次处理方式下时增加了还是减少了，固采取双尾检验。执行的统计测试的类型是Dependent test，因为测试者是测定了不同条件的反应时间，而且结果都是成对的

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 | Incongruent | difference |
| count | 24 | 24 | 24 |
| mean | 14.051125 | 22.015917 | -7.964792 |
| std | 3.559358 | 4.797057 | 4.864827 |
| min | 8.63 | 15.687 | -21.919 |
| 25% | 11.89525 | 18.71675 | -10.2585 |
| 50% | 14.3565 | 21.0175 | -7.6665 |
| 75% | 16.20075 | 24.0515 | -3.6455 |
| max | 22.328 | 35.255 | -1.95 |

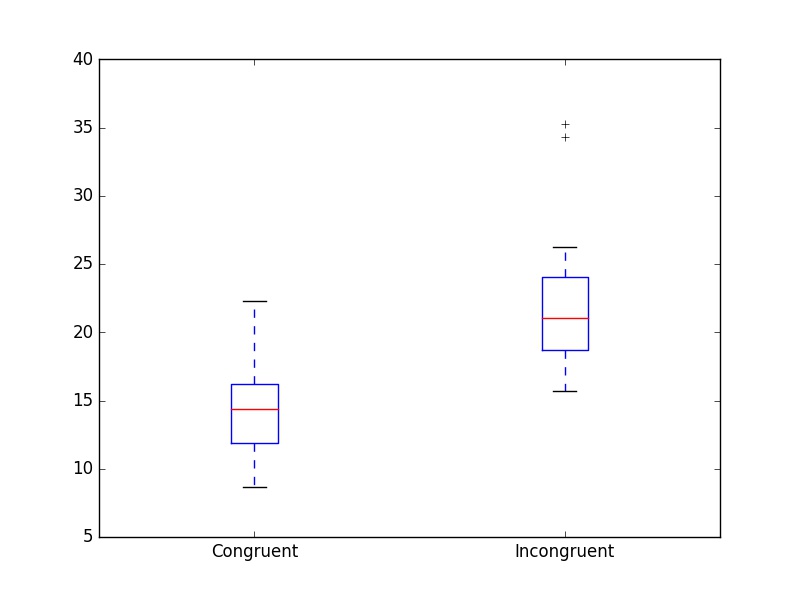
一致条件:平均值为14.051，中位数是14.357；标准差是3.559

不一致条件:平均值为22.016，中位数21.018；标准差是4.797

两者差异的平均值为-7.965， 中位数为-7.667；标准差为4.865

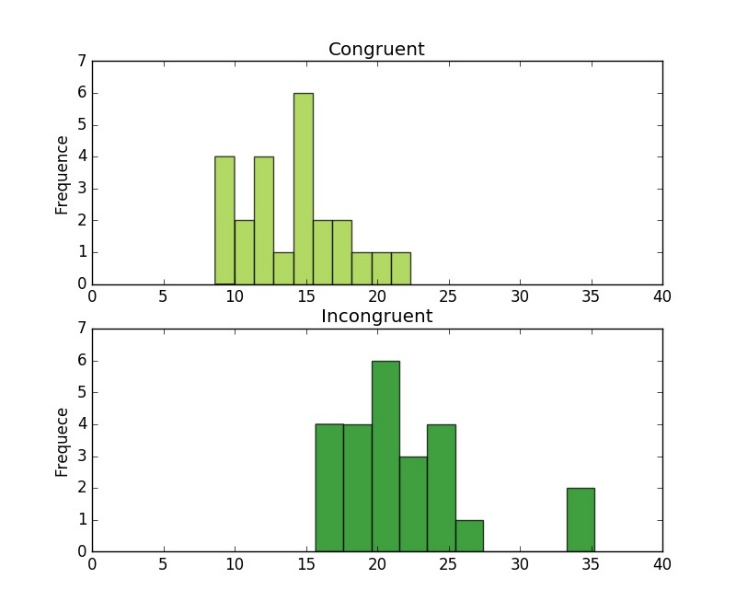
根据结果来看，初步确认文字和颜色一致的条件的使用时间比不一致的时间要少，相对来说更集中与均值，变异性较小

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。



a.文字和颜色一致的测试用时比不一致的要少，而且1/4位和3/4的距离也较小

b.不一致的处理存在异常值(IQR的值为5.33475，异常值大于29.019625)



a.文字和颜色一致的测试中，测试时间在15左右频率最高；而不一致的测试中，测试时间在20左右的数量比较集中，且有两个极大值超过了32

b.两个分布图都不是正太分布

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？

|  |  |  |
| --- | --- | --- |
|  | Congruent | Incongruent |
| 平均 | 14.051125 | 22.01591667 |
| 方差 | 12.66902907 | 23.01175704 |
| 观测值 | 24 | 24 |
| 泊松相关系数 | 0.351819527 |  |
| 假设平均差 | 0 |  |
| df | 23 |  |
| t Stat | -8.020706944 |  |
| P(T<=t) 单尾 | 2.0515E-08 |  |
| t 单尾临界 | 1.713871528 |  |
| P(T<=t) 双尾 | 4.103E-08 |  |
| t 双尾临界 | 2.06865761 |  |

因为该试验类型为不同条件下的Dependent test。此试验的自由度df=23，差异的标准差s=4.865，差异点估计μC– μI = -7.965，统计值t=-8.021；另本试验中采用置信水平α=0.5的双尾试验，t的临界值为±2.069。因为统计值t远小于临界值t，固推测试文字和颜色是否一致对测试者的反应时间表现出显著差异，同样拒绝零和假设H0。

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！

该试验中观察到的效应可能的原因:a.两次测试时间间隔短，前次测验影响测试者对第二次测试的反应时间b.测试者对文字的识别能力强于对颜色的判断c.人脑不能同时协调处理颜色判断和文字识别。

可能具有相似效应的任务:动物的条件反射，例如每个测试的动物都经过对两种处理的不同反应

注:本项目中使用了Excel的数据分析工具，python的matplotlib包的绘图工具以及pandas，运行条件是windows64位系统

优达学城

2016年9月