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？

自变量是同等大小的文字和油墨颜色一致的样本

及同等大小的文字和油墨颜色不一致的样本

因变量是两种情况下观察值的观察时间

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。

这个任务对假设集分别为：

零假设H0：一致总体平均时间和不一致总体平均时间相等。μc=μi

非零假设Ha：一致总体平均时间和不一致总体平均时间不相等 μc≠μi

执行t假设检验进行统计测试。

选择t假设检验的理由：在数据集中我们并没有得到关于总体的均值，标准偏差的任何信息，就是说，我们对总体并无了解，而只能通过在数据集中两种测试的样本进行统计测试并且来推断总体的情况，两种统计测试z检验需要对总体的分布情况有所了解，而t检验则不需要对总体样本进行了解。与此同时，此检验在变量上面是两个独立分布的变量，可以通过t两个变量间的建设检验来分析两种情况下面有无统计显著差异。

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。

在一致性测试中，总共包含了24个样本，其中最大值为22.328，最小值为8.63

均值为:14.05 中位数为:14.36，把数据进行分组距为5.可以得出下图1，从图中可以看出众数在15左右

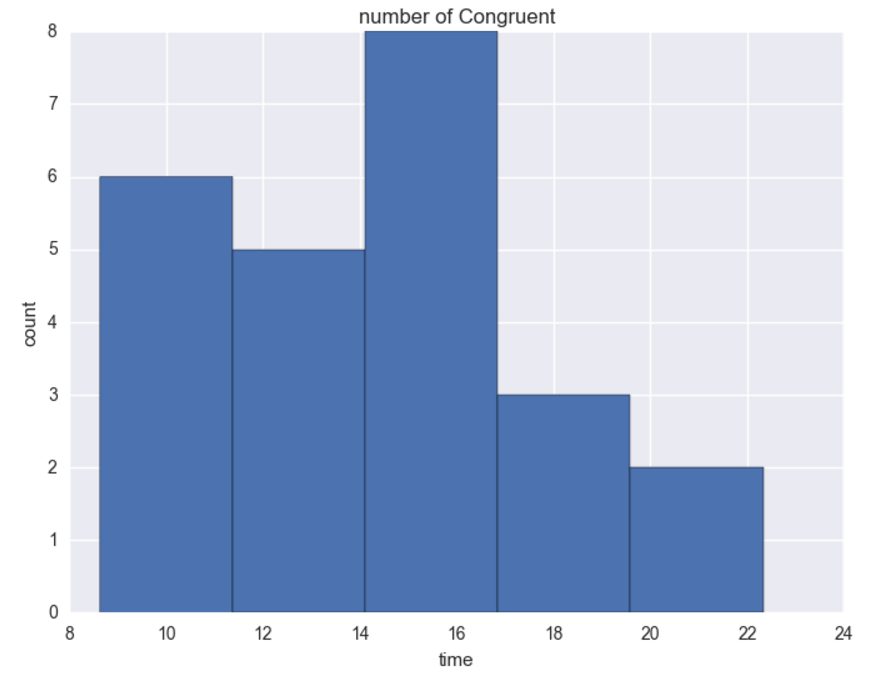


图1

在非一致性测试中，总共包含了24个样本，其中最大值为35.255，最小值为15.687

均值为22.015，中位数为21.0175，把数据进行分组为5可以得出下图，从途中可以看出众数为22-25区间内

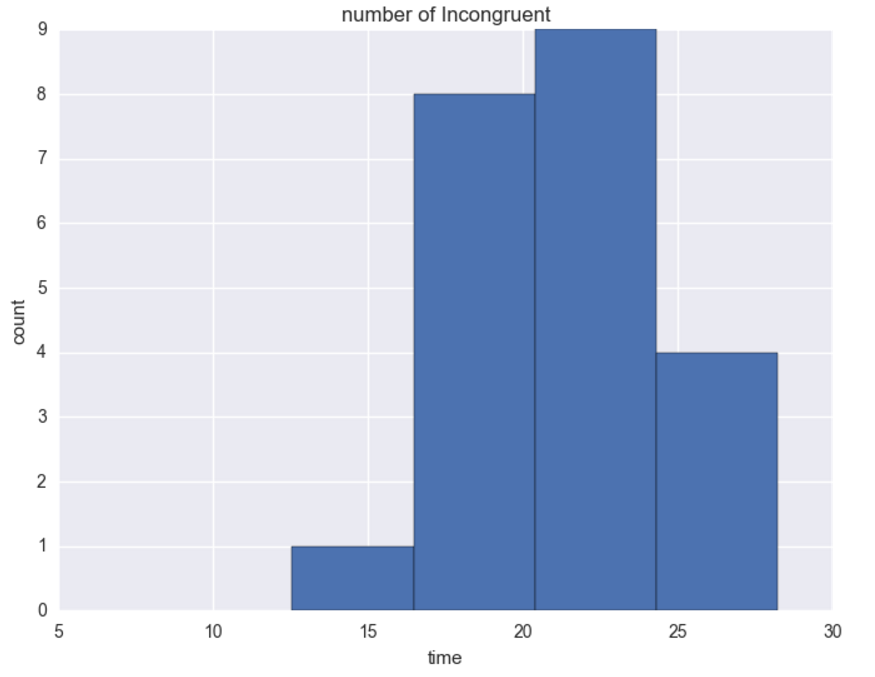


图2

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。

根据样本对数据分布绘制了图1及图2两个相应对直方图分别对于两个样本对直方图，对比图1 及图2中我们可以看出一致性测试中，最小值，最大值，均值，中位数都小于非一致性测试。从直方图对分布可以看出一致性测试中分布众数小于非一致性的分布众数。

使用了python绘制了线箱图如图3所示，从图中我们可以看到两者对中位数有明显差异。

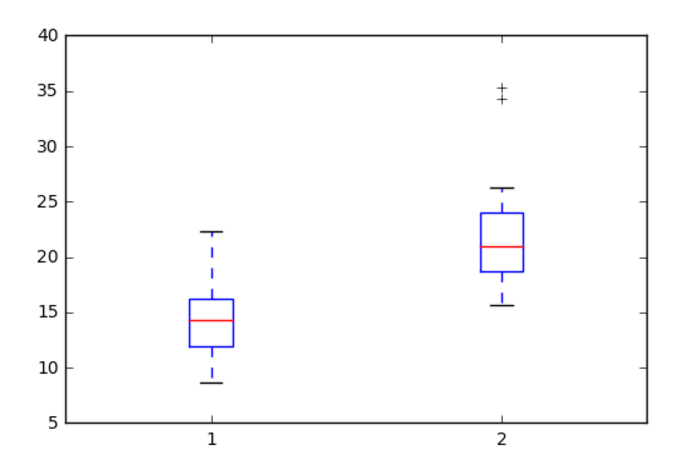


图3

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？

首先两个统计值是相依样本的t检验

一致性测试：NC=24 (样本数量) XC=14.05（均值），Sc=3.48 （标准方差）

非一致性测试：NI=24(样本数量) XI=22.016(均值) 方差SI=4.70（标准方差）

总体均值差异点估计μc-μi=14.05-22.016=-7.97

差异的标准偏差S=4.86

t 值==-7.97/4.86/=-8.03

置信水平选择了 95%的双尾检验,自由度=24-1 =23

查找t表格对α=0.5，自由度为46的双尾检验临界值为+-2.069

这里t<-2.069

根据<http://www.graphpad.com/>网站P值对计算

双尾检验时候P值小于0.0001

零假设为H0：一致性测试和非一致性测试无统计显著性差异Xc=XI

非零假设为Ha:一致性测试和非一致性测试有统计显著性差异XC

可以看出t>2.009对，于是成功拒绝了零假设。

从试验任务可以得出，Stroop试验中，一致性测试和非一致性测试具有统计显著差异，两种方法测试时间均值上面具有显著差异，从已有样本数据对总体对估计中可以看出一致性测试平均时间和非一致性测试平均时间显著不同

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！

我认为导致观察到这个效应的原因是人在观察文字颜色一致对时候，辨认时间比不一致要简单容易，因为人对颜色文字对对应有已有的认知积累在大脑，所以辨别时间要短。这种类似效应可以调查英语母语人群和非英语母语人群（学习过英语）人在阅读英语文章后，阅读时间对统计行差异。

优达学城

2016年9月