

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：文字条件为一致或者不一致；

因变量：说出同等大小的列表中的名称所花的时间；

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）。

零假设： $H_0: \mu_{\text{congruent}} = \mu_{\text{incongruent}}$

对立假设： $H_a: \mu_{\text{congruent}} \neq \mu_{\text{incongruent}}$

零假设：文字条件一致与不一致导致的阅读列表所花时间无明显区别；

对立假设：文字条件一致与不一致导致的阅读列表所花有明显区别；

$\mu_{\text{congruent}}$ ：表示文字条件一致时说出列表所花时间平均值

$\mu_{\text{incongruent}}$ ：表示文件条件不一致时说出列表所花时间的平均值

统计检验类型：相依样本的 t 检验

理由：首先我们并没有总体的数据，因此无法得到总体均值，而样本量小于 30，因此选择 t 检验，又因为我们的样本并非是两组独立样本，而是同一组样本在不同条件下的实验，因此是相依样本；

相依样本的 t 检验的前提：

1. 正态分布的总体（即使不满足正态分布，通常 t 也是有效的，但是只针对分布偏斜不大时）；

2. 针对同一样本不同条件下的观察；

条件都是满足的；

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。

- 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。

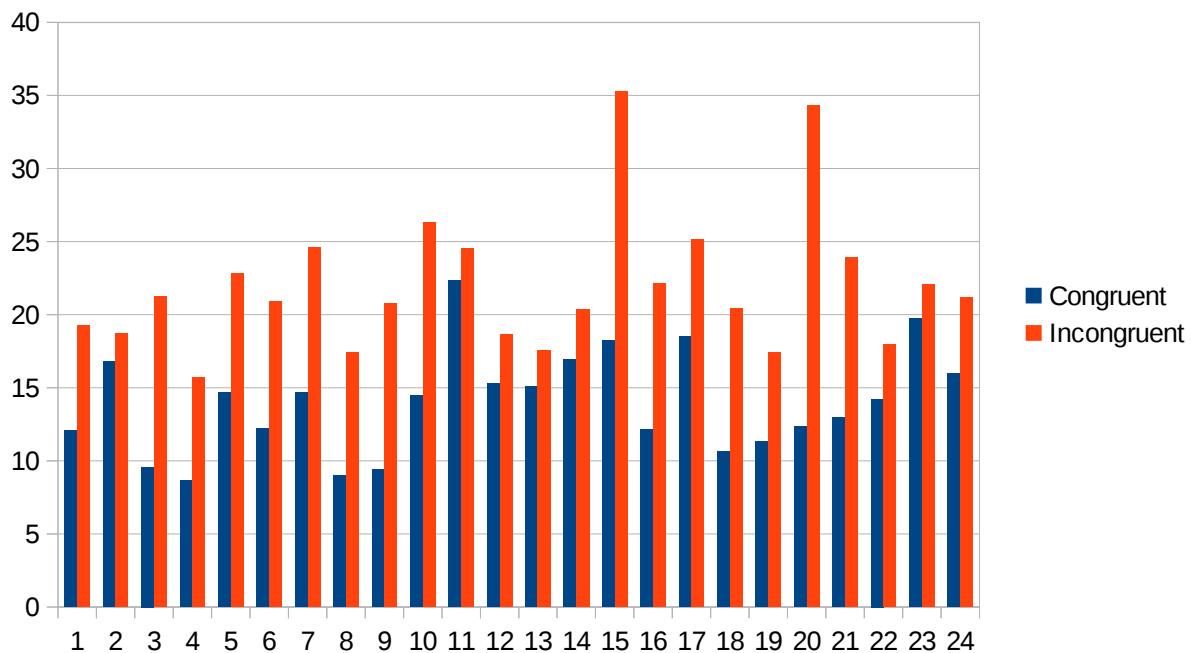
congruent 均值：14.051125

incongruent 均值：22.015917

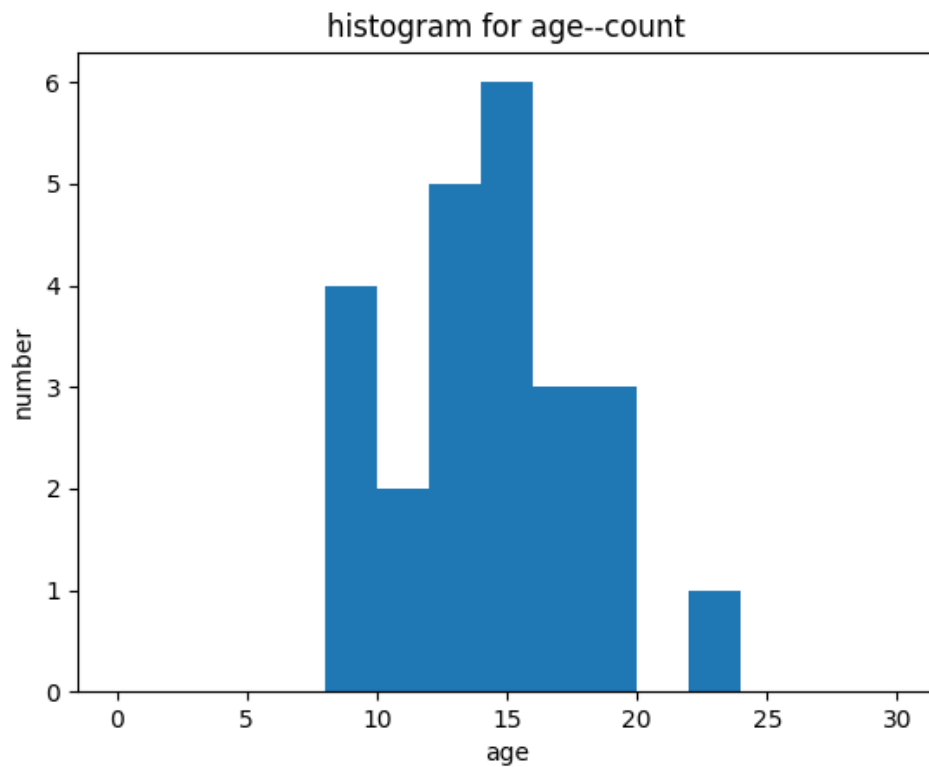
incongruent-congruent 差的均值：7.9648

差的标准差：4.8648

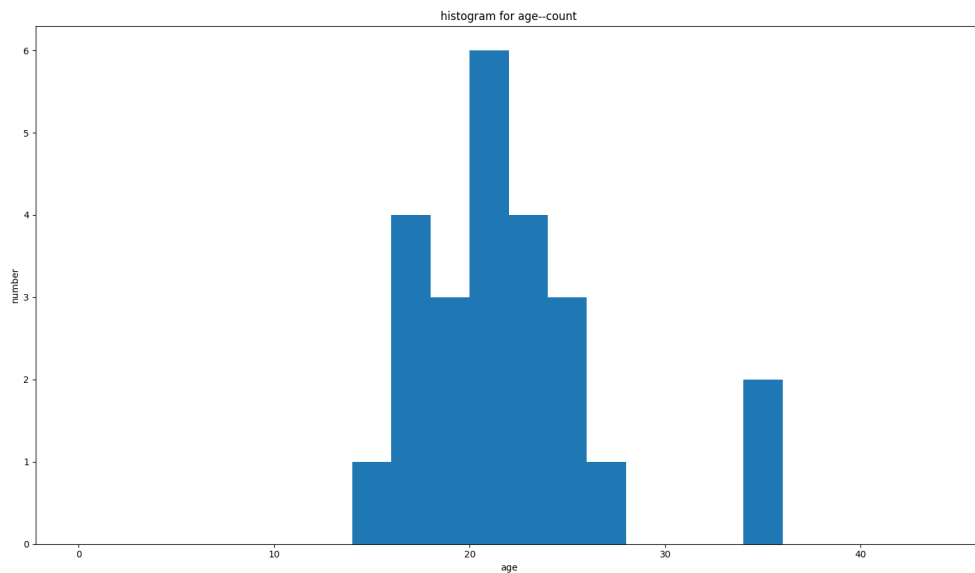
- 提供显示样本数据分布的一个或两个可视化。用一两句话说明你从图中观察到的结果。



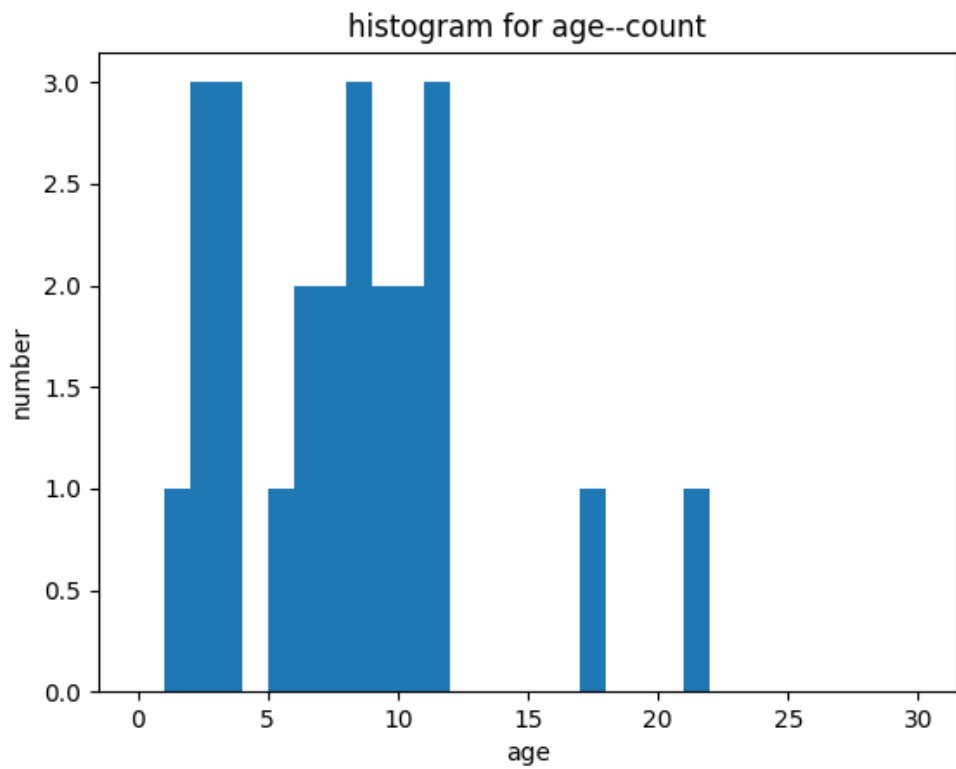
该图描述了每个样本在一致和不一致的情况下分别所花的时间，可以看到所有样本均为不一致所花时间比一致多；



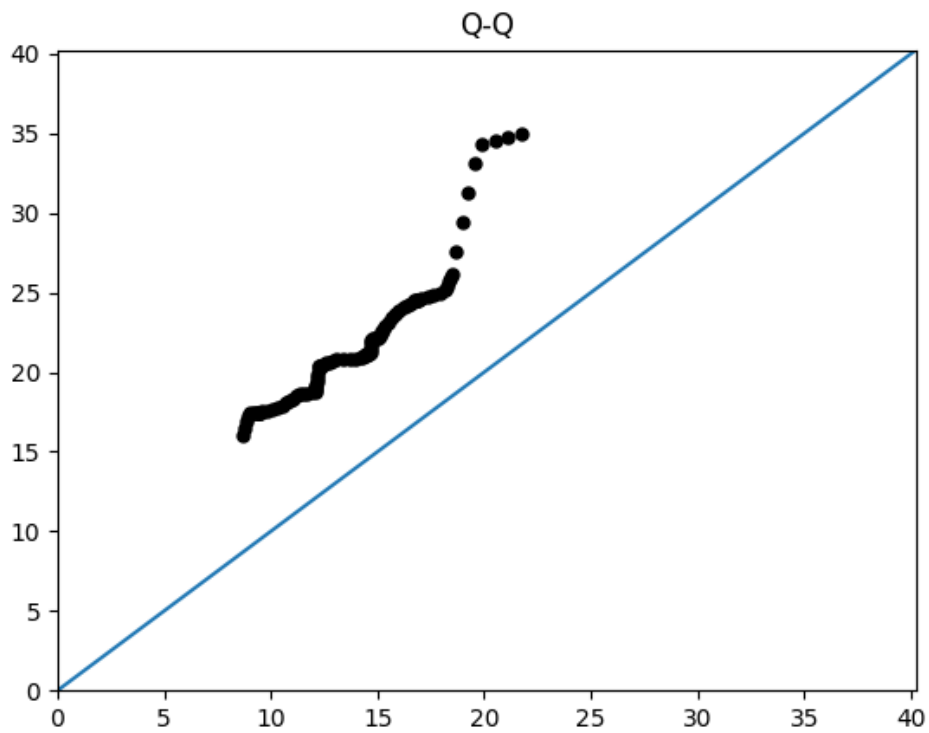
该图描述了当字体条件为一致时，组距为 2 的分布图，横坐标代表所花时间，纵坐标代表对应的人数，可以看到大部分样本集中在 8~18 的区间；



该图描述了在字体条件为不一致时，组距为 2 的分布图，横坐标代表所花时间，纵坐标代表对应的人数，可以看到大部分样本集中在 17~25 之间，对比上图也能看到所花时间有一个明显的增多；



该图描述了不一致所花时间减去一致所花时间的分布图，组距为 1.8，可以看到大部分数据集中在 3~10 之间，这代表着字体是否一致对样本的影响普遍有 6,7；



该图描述了两组样本绘制的 Q-Q 图，可以看出离 $y=x$ 线较远，可以估计两组样本并不来自同一个分布；

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？

mean_diff(incongruent-congruent 差的均值):7.9648

样本量相同 n:24

检验方向:双尾检验

检验方式:相依样本 t 检验

自由度 df:n-1=23

偏差平方和 ss:544.3304

标准误差 se:4.8309

置信水平:95%

α :0.05

t 临界值:±2.069

置信区间:-2.0304~17.9599

t_statistic:8.077

r^2 :0.7393

r^2 表示 mean_diff 有 73.93% 可以有文字条件是否一致来解释；

$t_statistic > 2.009$, reject H_0 , 该统计结果具有显著性，统计结果差异中有约 73.93% 来自于条件的一致与否；

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！

原因：当人类在识别一个物体的时候（该实验是说出文字颜色），是通过多个渠道来获知的，比如该实验，有两个渠道取获取，一个是本身的颜色，一个是文字所描述的颜色，当该文字条件一致时，由于有两个渠道会得到同一个结果，因此获取时间更短，也更准确，而文字条件不一致的情况下，当这两个渠道的信息到达大脑时大脑会花一点时间去判断哪一个是正确的，这就影响了速度和准确性。

替代/类似任务：

类似任务：比如可以同时播放两个声音，一个是人声“猫”，一个是猫叫声“喵喵”，同样的还有狗-“汪汪”，牛-“哞哞”等等，条件一致为猫-“喵喵”，条件不一致为猫-“汪汪”，记录样本说出叫声对应的动物，记录时间；

优达学城

2016 年 9 月