

## Stroop(斯特布鲁)任务调查报告

1.此任务中，我们的自变量是文字条件，它可能是一致文字条件，也可能是不一致文字条件。因变量是参与者完成说出列表中墨色名称的时间。

2.我们用 $\mu_A$ 表示总体在文字颜色一致时完成任务使用的时间均值，用 $\mu_B$ 表示总体在文字颜色不一致时完成任务使用的时间均值。

我们的假设集如下：

零假设：两种文字条件下完成任务需要的时间不存在显著差异。用数学符号表示如下：

$$H_0: \mu_A = \mu_B$$

对立假设：两种文字条件下完成任务需要的时间存在显著差异。用数学符号表示如下：

$$H_A: \mu_A \neq \mu_B$$

因为我们并不知道不一致文字条件对比一致文字条件会造成的效应的方向，所以我们选择 $\mu_A \neq \mu_B$ 这种双尾检验（非方向性检验）。除非是我们的目的是要清楚的判断两种情况其中一种是否优于(或劣于)另一种，才适合选择单尾检验，选择 $\mu_A < \mu_B$  或  $\mu_A > \mu_B$  作为对立假设。

选择 t 检验还是 z 检验：z 检验适用于知道总体的平均值和方差，然后条件不变的情况下，通过抽取 n 个样本，计算这些样本平均值和总体平均值的关系（即 z 值）。比如总体平均值和方差我们并不知道，而且计算 z 值也和我们的情况不匹配。我们这个实验是要给人提供两种条件下，观察对阅读时间的影响，这两种条件为对照关系，计算的是两组样本之间的区别有多大（t 值），这并不是总体中取一部分样本、计算样本和总体的区别（z 值）。

至于是相依样本和独立样本的选择，这里应该选择相依样本。因为独立样本要求参与两组测试的人应该没有交集。而我们实验中是让所有参与者都完成了两种条件下的任务，获取两种条件分布使用的时间，没有做到保持受试者的独立性。

另外，总体是近似于正态分布，通过样本数据可以估算总体方差，总体方差大概相等，这些 t 检验适用的条件都符合。

综上，建议选择相依样本的双尾 t 检验。

3. 从数据集中我们可以得到以下信息：

共 24 人参与了测试，每人均统计了文字颜色一致时使用的时间和文字颜色不一致使用的时间，结果汇总成两组数据，我们称文字颜色一致条件下的数据为 A 组，称文字颜色不一致条件下的数据为 B 组。

集中趋势测量数据：

两组数据中均没有出现重复的数字，没有众数。

经计算 A 组的平均值为 14.05，中位数为 14.3565，B 组的平均值为 22.02，中位数为 21.0175。能看出 B 组(不一致文字条件)整体上花费的时间更长。

变异测量数据：

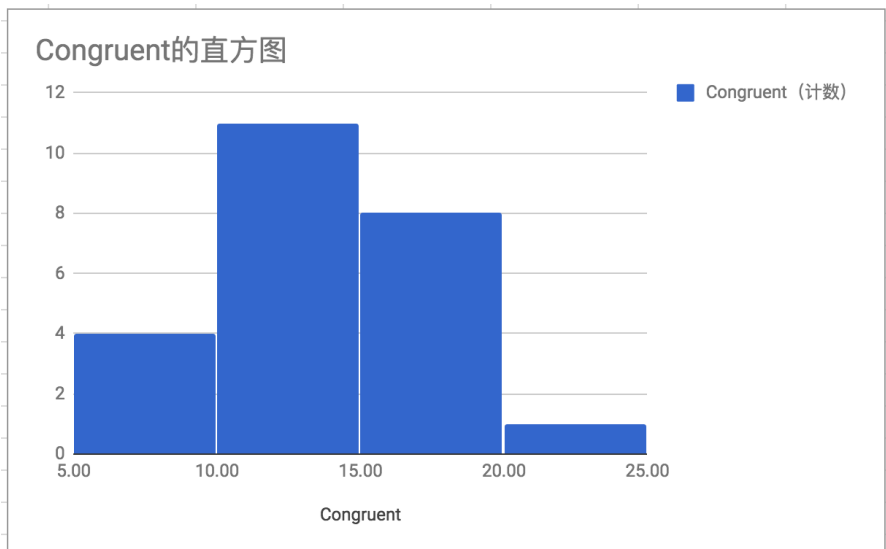
A 组的样本方差为 12.67，样本标准差为 3.56，离散系数为 0.25。

B 组的样本方差为 23.01，样本标准差为 4.80，离散系数为 0.22。

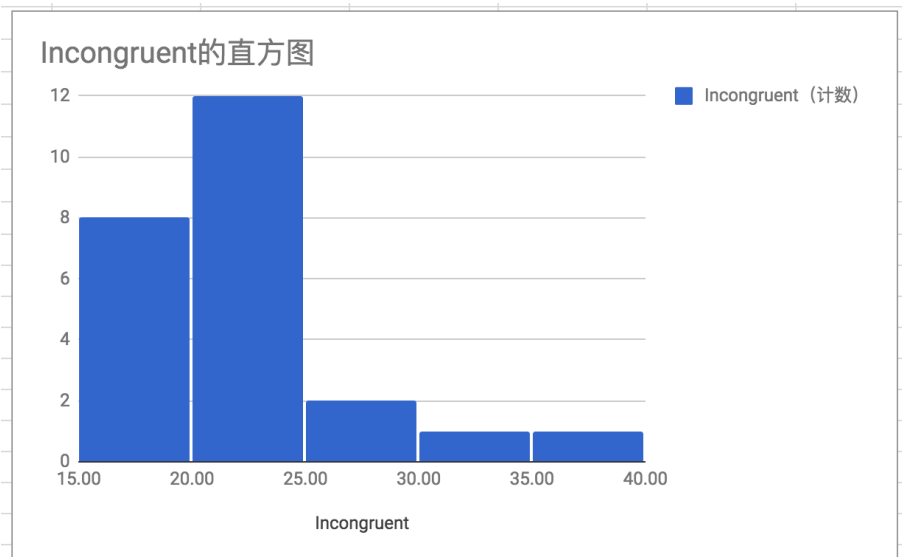
能看出 B 组数据离散程度比 A 组略小一些。

4.

A 组的直方图如下：



B 组的直方图如下：



从图中能看到，B 组（不一致文字条件）所用的时间整体上都比 A 组要多，B 组大量人所用时间集中在 20-25 秒，而 A 组集中在 10-15 秒。

5.为了表示方便，我们把一致文字条件下测试的数据简称 A 组，把不一致文字条件下测试的数据简称 B 组。根据实验数据可以看出，有 24 名参与者均参与了两种测试，所以自由度 df 为  $24-1=23$ 。通过计算可以获得 A 组的样本均值  $\bar{x}_A=14.051$  ,B 组的样本均值  $\bar{x}_B=22.016$ 。

目前通过样本数据获得样本之间的差值  $\bar{x}_A - \bar{x}_B = -7.965$ ，接下来，计算两组数据样本均值的标准偏差  $s = 4.865$ ，均值的标准误差  $sem = s/\sqrt{n} = 4.865/\sqrt{24} = 0.993$ ，然后计算 t 统计量  $t = (\bar{x}_A - \bar{x}_B) / sem = -7.965/0.993 = -8.021$ 。

自由度 df 为 23，对于  $\alpha$  水平为 0.05 来说，每个尾部是 0.025，查询 t 表格，t 的临界值为  $\pm 2.069$ 。

通过比较发现，t 统计量的值 -8.021 比临界值的最小值 -2.069 要小得多。

接下来计算一下 p 值（对于双尾，p 表示的是小于负的 t 统计量及大于 t 统计量的概率）。根据 t 表格，df 自由度为 23，正数 t 统计量为 8.021，能看到它比 p 值为 0.0005 是 t 值的 3.768 还要大，所以上尾的 p 值  $< 0.0005$ ，即  $\frac{1}{2}p < 0.0005$ ，得到  $p < 0.001$ ，能看出概率很低。

另外一种方法，通过 graphpad 网站，输入 t 和 df，得出 p 小于 0.001。

按照传统的标准，这种差异被认为具有极大的统计学意义。

因此我们拒绝零假设，即认为两种文字条件产生的实验结果具有显著差异。

这一结果符合自己的期望，因为正常情况下不一致的文字条件会使人思维产生逻辑混淆，影响人准确的判断，造成完成任务时间更长。通过实验验证这一结果也在情理之中。

6.我认为导致观察到的效应的原因，是当文字和颜色不一致时，人的潜意识会取思考正确应该对应的颜色，而他和看到的颜色是不一致的，由此就会感觉到混乱，影响了认知能力。类似的情况，比如把电脑键盘的键帽按乱序重新安上，虽然很多人都能盲打，知道每个键的正确位置是在哪个位置，但如果看到显示的是别的字母，还是会有一些干扰作用，会影响打字效率。