

统计学：决策的科学项目说明 – 解答

学生：杨高创

时间：12/08

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？

答：自变量是字体颜色和文字是否一致；
因变量是读出同等数量文字的完成时间。

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）。

答：

文字说明：

零假设：颜色和字体不一致的阅读时间和颜色和字体一致的阅读时间没有区别；

对立假设：颜色和字体不一致的阅读时间将增加阅读时间，或者说比颜色和字体一致的时间更长；

数学说明：

假设 μ_s 代表字体和颜色一致的阅读时间总体均值， μ_d 代表字体和颜色不一致的阅读时间总体均值。

零假设 $\text{null} \rightarrow H_0: \mu_s = \mu_d$

对立假设 $H_a: \mu_s \neq \mu_d$

检验方法选择：

1) 本案例只有两个样本数据，没有总体 population 的均值，标准偏差，无法进行 Z 检验，

因此考虑 t 统计检验；

- 2) 该试验是同一批人在两个不同条件下进行实验，属于相依样本的重复衡量设计。
- 3) 该检验没有方向性，不做预期，因此采用双尾检验。

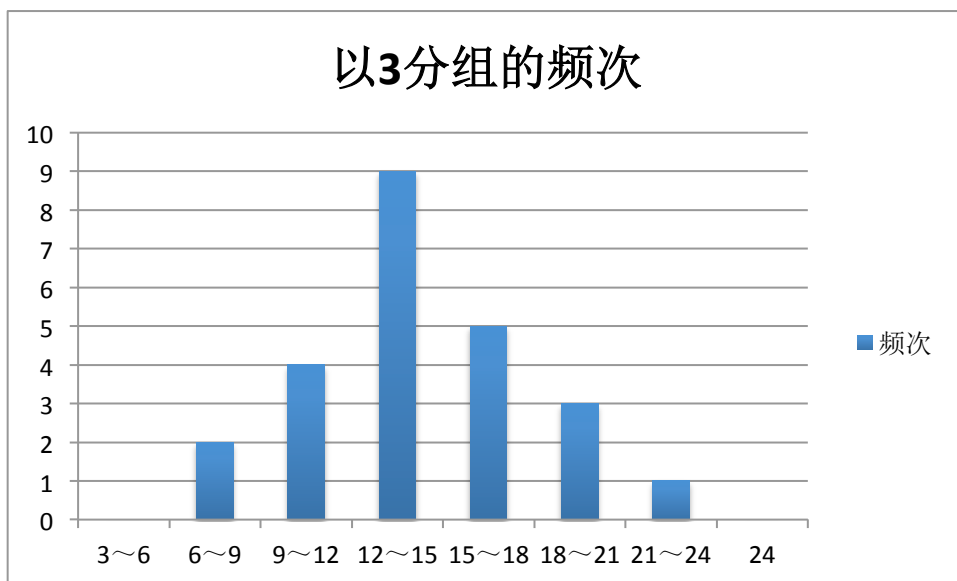
对比 t 检验必须有一定的前提条件：

- 1) 这批人应该是从人群中随机挑选出来，互相之间没有相关性，样本是随机样本；
- 2) 下面对两组样本的分布分析，总体是接近正态分布，因此推测总体样本应该是正态分布；
- 3) 通过样本的方差可以推理总体样本的方差；
- 4) 总体方差应该大致相同 (roughly equal)

3. 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。

对 Congruent 的数据说明：

- 1) 按照 5, 4 和 3 分组进行频次统计，发现按照 3 分组比较能反映数据的分布情况；
- 2) Mean=14.05, Median=14.36, s=3.56, Mean 和 Median 基本接近，符合正态分布数据的特点；
- 3) 通过对直方图的观察，最高 bar 两边的两个 bar 高度基本上接近，因此可以当作正态分布来看；
- 4) 对样本数据排序后，计算分布情况：Q1=11.71, Q2=14.36, Q3=16.40, 所以 IQR=Q3-Q1=4.69, outlier 异常数据应该是 $<Q1 - 1.5 \cdot IQR$ 或者 $>Q3 + 1.5 \cdot IQR$, 从目前的这组数据看，没有异常数据。

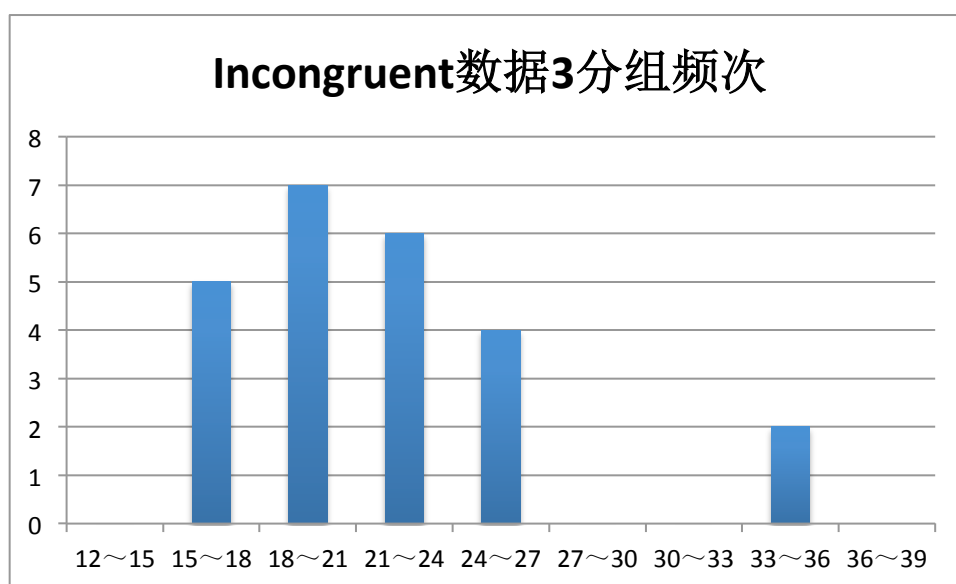


对 Incongruent 的数据说明：

- 1) 按照 5, 4 和 3 分组进行频次统计，发现按照 3 分组比较能反映数据的分布情况；
- 2) 把数据排序后进行计算：Q1=18.69, Q2=21.02, Q3=24.21, 所以 IQR=Q3-Q1=5.52, outlier 异常数据应该是 $<Q1 - 1.5 \cdot IQR = 10.41$ 或者 $>Q3 + 1.5 \cdot IQR = 32.49$, 从目前的这组数据看，有两个数据是异常数据 outlier: 34.288 和 35.255。
- 3) Mean=22.02, Median = 21.02, s=4.80, Mean, Median 基本接近，从数据上看

可以符合正态分布特点；

- 4) 去掉两个异常值，通过对直方图的观察，最高 bar 两边两边的 bar 告诉基本接近，基本上可以当作正态分布来进行统计检验。



4. 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。

见第三题答案。

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？

统计数据 and 关键值如下：

名称	公式	数值
X_1 mean		14.05
X_2 mean		22.16
X_1 mean - X_2 mean M_d		-7.96
X_1 - X_2 的差值的标准偏差 S_d		4.86
SE(d)标准误差	S_d / \sqrt{n}	0.99
t-statistic	$M_d / SE(d)$	-8.02
对应的 P 值		<.0001
自由度 df	n-1	23
t-critical value($\alpha = 0.05$) two tailed	在 t-table 上找 $\alpha = 0.025$, df=23 的值	± 2.069
CI ($\alpha = 0.05$)	$M_d \pm t^* S_d / \sqrt{n}$	(-10.01, -5.91)
t-critical value($\alpha = 0.01$) two tailed	在 t-table 上找 $\alpha = 0.005$, df=23 的值	± 2.807
CI ($\alpha = 0.01$)	$M_d \pm t^* S_d / \sqrt{n}$	(-10.74, -5.18)
Correlation measures r^2	$t\text{-statistic}^2 / (t\text{-statistic}^2 + df)$	0.74
Cohen's d	M_d / S_d	-1.64

结论：

按照 α 是 0.05 进行总结，t-statistic(23)的 P 值(小于.0001))远远小于 t-critical value 的 P 值.05, 就说明这两组数据有显著性差异。因此 reject H_0 假设，对立假设 H_a 成立。
差不多不一致的阅读时间和一致的阅读时间相差 7 到 9 秒。
有 74%的差异是由于显示文字与打印颜色不一致造成的。

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！

答：

原因是由于字体和颜色不一致干扰了大脑的认知，使反应速度变慢，因此增加了阅读时间。

如果增加干扰强度可能会增加阅读时间，比如增加灯光的变化，旁边有音乐或者电视的干扰等等。