

斯特鲁普效应调查结果数据分析报告

1. 实验自变量和因变量

自变量：显示的颜色词与它们文字的打印颜色是否匹配

因变量：参与者说出墨色的时间

2. 建立假设

以普通人为总体，设普通人在颜色词与它们文字的打印颜色匹配和不匹配两种情况下，说出墨色名称所用时间的差值（即颜色词和打印颜色不匹配情况下的用时减去匹配的情况）的平均值为 μ_D 。

原假设 (H_0): $\mu_D \leq 0$ （即在颜色词与它们文字的打印颜色不匹配的情况下说出墨色用时不大于匹配的情况）

备择假设 (H_a): $\mu_D > 0$ （即在颜色词与它们文字的打印颜色不匹配的情况下说出墨色用时大于匹配的情况）

选择检验的显著水平 $\alpha=0.05$

通过记录的每位参与者在两种情况下同等大小的列表中的墨色名称所用时间，可以计算出两种情况下每位参与者的用时差值。以每位参与者用时差值为样本，利用 t 检验对提出的假设进行检验，属于配对 t 检验。t 检验的前提是：1. 样本是从目标总体中随机选取的；2. 样本来自的总体近似服从正态分布。

3. 数据集的描述性统计分析

样本数据集为每位参与者在两种情况下的用时差值，描述性统计信息如下：

平均数：7.96

中位数：7.67

众数（将所有时间差四舍五入到整数位计算）：2

样本方差：23.67

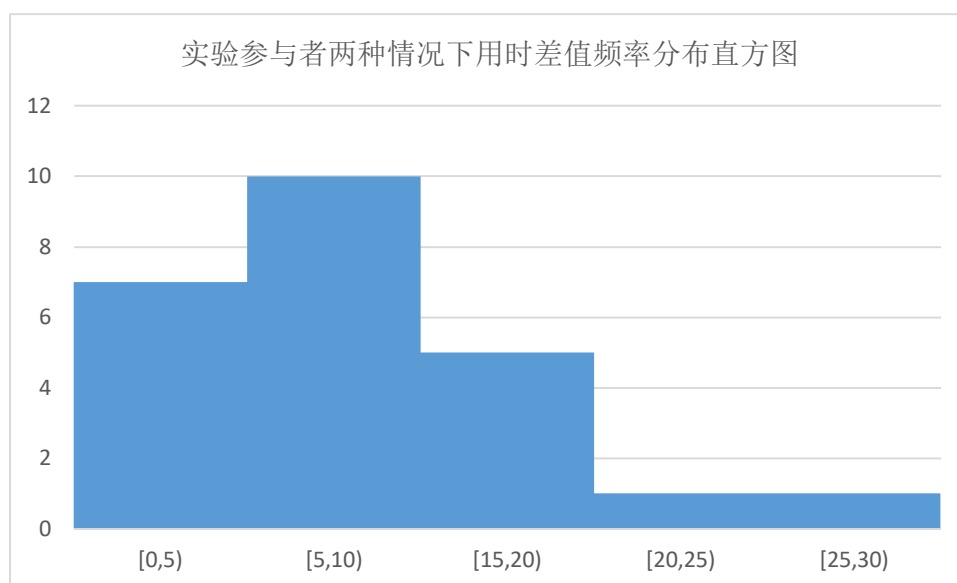
Q1：3.65

Q3：10.26

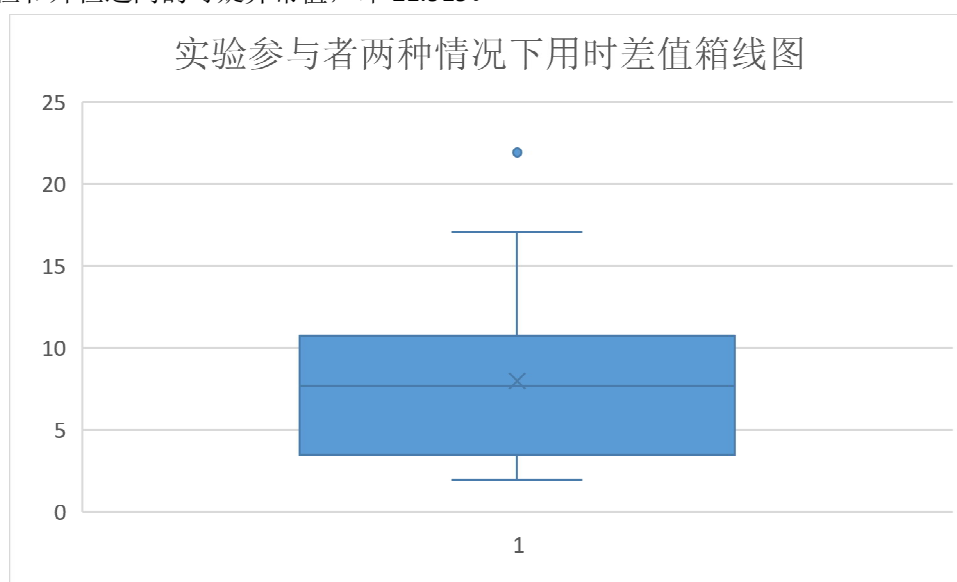
IQR：6.61

4. 样本数据集可视化

样本数据的频率分布直方图如下所示，从图中可以看出，实验参与者在两种下的用时差值呈右偏分布，且用时差值均大于 0。



样本数据的箱线图如下所示，从图中可以看出，实验参与者在两种情况下的用时差值有一个在内栏和外栏之间的可疑异常值，即 21.919。



5. 统计检验

对样本数据进行 t 检验，
$$t = \frac{\bar{X} - 0}{\frac{s}{\sqrt{n}}} = \frac{7.96}{\frac{\sqrt{23.67}}{\sqrt{24}}} \approx \frac{7.96}{0.99} \approx 8.04$$

统计检验结果如下：

$t(23)=8.04$, $p<0.0005$, 单尾

由于 $p<\alpha$ ，在 0.05 显著水平下，拒绝原假设。

在 $\alpha=0.0005$, $df=23$, 单尾的条件下，查表得 $t\text{-critical}=1.714$, $|t\text{-critical}|<|t\text{-statistic}|$ ，同样拒绝原假设。

可以得出结论，当显示的颜色词与它们文字的打印颜色不匹配时，说出墨色名称所用时间大

于匹配的情况，与期望一致。

6. 扩展思考

猜想导致观察效应的原因两种：1、颜色词的词义对普通人说出墨色名称产生干扰；2、文字本身对说出墨色词产生干扰。

由此可以设计两组实验，仍对每组实验中不同条件下参与者的用时差值进行统计检验：

- 1、将实验条件改为不一致文字条件和不相关文字条件。不相关文字条件中，显示的文字是与颜色无关的词语，不一致文字条件与原实验一致。
- 2、将实验条件改为无文字条件和不相关文字条件。无文字条件中，将文字替换成一系列不同的图形，不相关文字条件与原实验 1 一致。