

项目 1：统计学检验斯特鲁普效应

项目背景分析

斯特鲁普效应（Stroop Effect）展示了人们对事物的认知过程已是一个自动化的历程。当有一个新的刺激出现时，如果它的特征和原先的刺激相似或符合一致，便会加速人们的认知；反之，若新的刺激特征与原先的刺激不相同，则会干扰人们的认知，使人们的所需的反应时间变长。在心理学中，斯特鲁普效应是干扰对处理任务时反应时间的论证。

为了验证这种效应的存在，本文选取了其中一项实验的样本数据，并采用统计学的方法对其进行检验。

实验简介

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。

一致文字条件测试数据列表：

Word Set #1

RED	GREEN	BLUE	YELLOW	PINK
ORANGE	BLUE	GREEN	BLUE	WHITE
GREEN	YELLOW	ORANGE	BLUE	WHITE
BROWN	RED	BLUE	YELLOW	GREEN
PINK	YELLOW	GREEN	BLUE	RED

不一致文字条件测试数据列表：

Word Set #2

RED	GREEN	BLUE	YELLOW	PINK
ORANGE	BLUE	GREEN	BLUE	WHITE
GREEN	YELLOW	ORANGE	BLUE	WHITE
BROWN	RED	BLUE	YELLOW	GREEN
PINK	YELLOW	GREEN	BLUE	RED

表 1 是每名任务参与者完成两组条件的用时统计（单位：秒）。该项统计的样本量为 25，样本的自由度（degrees of freedom, df）为 24。

表 1 参与者完成任务用时数据（表中红色数据为笔者测试数据）

No.	Congruent	Incongruent
1	12.079	19.278
2	16.791	18.741
3	9.564	21.214
4	8.63	15.687
5	14.669	22.803
6	12.238	20.878
7	14.692	24.572
8	8.987	17.394
9	9.401	20.762
10	14.48	26.282
11	22.328	24.524
12	15.298	18.644
13	15.073	17.51
14	16.929	20.33
15	18.2	35.255
16	12.13	22.158
17	18.495	25.139
18	10.639	20.429
19	11.344	17.425
20	12.369	34.288
21	12.944	23.894
22	14.233	17.96
23	19.71	22.058
24	16.004	21.157
25	19.729	32.668

调查问题

1. 自变量与因变量

自变量：文字与油墨颜色是否一致；

因变量：参与者说出同等大小的列表中的墨色名称的时间。

2. 确定统计检验类型

对统计结果使用相依样本 t 检验（Dependent samples t-test）。选择 t 检验是因为：

- 1) 我们是以部分任务参与者的结果作为样本进行统计决策，即总体参数是未知的，且样本量小于 30；
- 2) 用时差集的样本数据近似呈正态分布，满足 t 检验的前提要求；

选择相依样本 t 检验的原因是：

- 1) 该项任务是统计两种不同的处理条件（文字条件是否一致）对同一组人的处理效果（所用时间），属于重复衡量设计（Repeated Measures Design）；
- 2) 样本数据来自的总体是独立的，且总体近似呈正态分布；
- 3) 样本数据可以用于估计总体方差，且总体方差相等；

假设集包含完成一致文字条件测试时间的总体均值（ μ_C ）与完成不一致文字条件测试时间的总体均值（ μ_I ）的差异情况。其中：

- 1) 零假设（ H_0 ）：完成一致文字条件的时间大于或等于完成不一致文字条件的时间，即

$$H_0: \mu_C \geq \mu_I$$

- 2) 对立假设（ H_A ）：完成一致文字条件的时间小于完成不一致文字条件的时间，即

$$H_A: \mu_C < \mu_I$$

尾检验的类型为负向的单尾检验（one tail test in the negative direction），因为根据对立假设，我们希望结果为负，即 $\mu_C - \mu_I < 0$ 。

3. 数据集的描述性统计

基于实验简介中给出的样本数据，可以得到以下描述性统计：

- 1) 集中趋势测量：
 - 均值：-8.16
 - 中位数：-8.134
- 2) 变异测量：
 - 值域：19.97
 - 方差：23.67
 - 标准差：4.87

4. 样本数据分布可视化

令组距为 4，绘制出两种条件用时差集的直方图如图 1 所示。可以看到，用时差集近似呈正态分布，其众数、均值和中位数均位于 (-10, -6) 区间内。

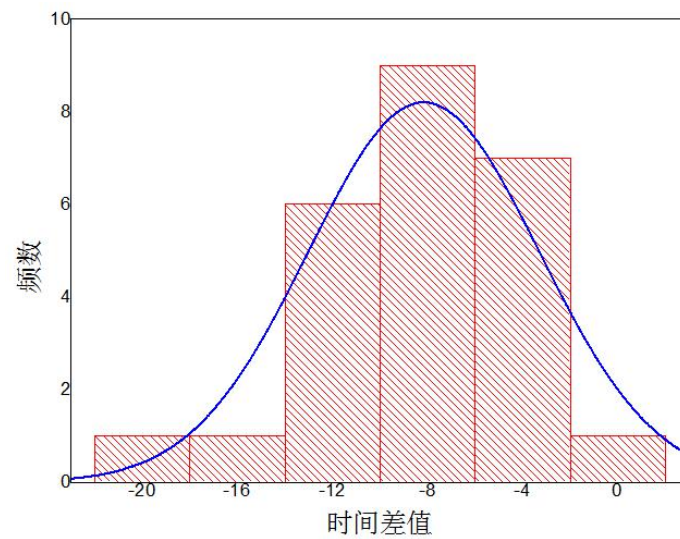


图 1 差集直方图

5. 统计测试

根据上述样本得出的统计数据如下（“//”后内容为注释）：

- 标准误差 (SE) : 0.97
- t 统计量 (t-statistic) : -8.39
- Cohen'd: -1.68 //表示样本均值 (-8.16) 与零假设中均值 (0.00) 间相差-1.68 个标准偏差
- 确定系数 (R^2) : .75 (或 75%) //表示均值差异的 75%是由上述两种测试条件导致的
- P 值 (P-value): 0.00005 //即当零假为真时, 获得上述样本均值的概率为 0.00005

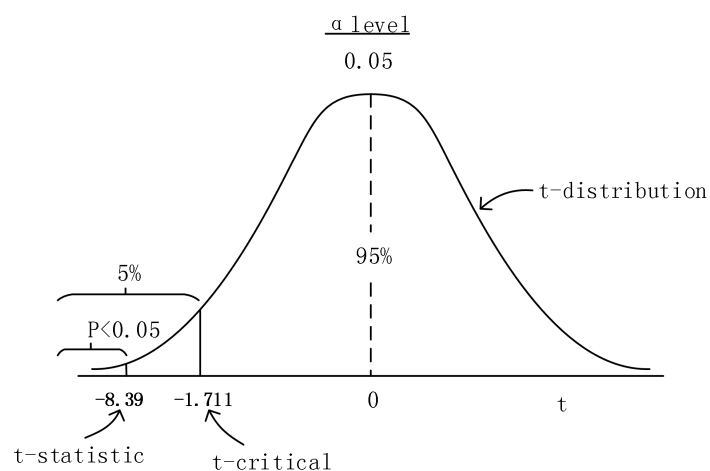


图 2 样本的 t 分布图

若选取显著性水平 α level = 0.05，则：

- t 临界值 (t-critical)：-1.711
- 均值的 95%置信区间 (CI)：(-10.17, -6.16)

如图 2 所示，在显著性水平 α level = 0.05 时，置信水平为 $1-\alpha = 95\%$ ，单尾负向检验时由 t-table 可知，在 $df = 24$ 时，t 临界值 (t-critical) 为-1.711。t 值 (-8.39) 位于临界区内，P 值 <0.05 ，具有统计显著意义，所以拒绝零假设，即可以确认斯特鲁普效应的存在。

6. 结论

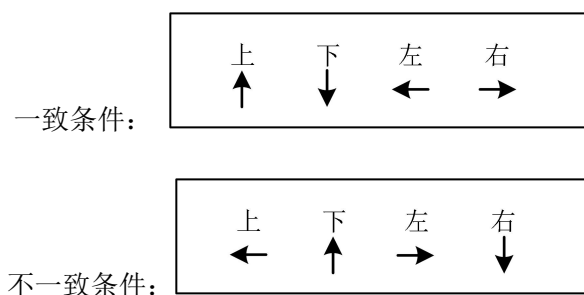
综上，通过对相依样本的 t 检验可知，当油墨颜色与文字不一致时，的确会导致参与者说出同等大小列表中的墨色名称的时间增加。

检验得到的结论完全符合预期，因为斯特鲁普效应 (Stroop Effect) 就是指当有与原有认知不同的情况出现时，人们的反应时间会较长。斯特鲁普效应已通过实验论证，斯特鲁普颜色与文字实验 (SCWT) 是一个在神经心理学领域被广泛运用于临床与调查中的实验。

7. 原因分析与替代任务设计

产生斯特鲁普效应的原因是：人们在认知过程中会受到环境的影响，若果这个环境对认知是干扰的，因为大脑要分神去抑制这个干扰，认知时间自然就会变长。下面给出一个上述任务的替代任务：

该项任务中有两个条件，每个条件均包含由汉字和箭头指出的方向——上、下、左、右。一致条件：汉字字义和箭头指向相同；不一致条件：汉字字义和箭头指向不同。记录参与者读出相同大小列表中箭头方向的时间。



参考资料

<https://zh.wikipedia.org/wiki/%E6%96%AF%E7%89%B9%E9%B2%81%E6%99%AE%E6%95%88%E5%BA%94> 斯特鲁普效应

https://raw.githubusercontent.com/ShiChJ/DAND-Basic-Materials/master/P1/Project_Files/stroopdata.csv Stroop 样本数据

<https://faculty.washington.edu/chudler/java/ready.html> Stroop 任务

<https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg> t-table

<https://www.graphpad.com/quickcalcs/> GraphPad

宋振
2017 年 12 月