

基于斯特鲁普效应(Stroop effect)的分析

#修改处使用蓝色高亮显示，方便复核审验

#本报告系Udacity机器学习纳米学位统计学基础课程，检验心理学现象项目：分析斯特鲁普效应

#项目报告撰写人Udacity学员：张然，2017年6月14日

#统计学数据、引用资料、分析工具等来源优达学城、谷歌和其它平台，备注于报告结尾

一. 斯特鲁普效应Stroop effect (摘自百度百科¹)

斯特鲁普效应(Stroop effect)在心理学中指干扰因素对反应时间的影响。例如当测试者被要求说出某个颜色和其覆盖的文字字面意义不符的词语时，被测者往往会反应速度下降，出错率上升。

准确的表述是当人们对某一特定刺激作出反应时，由于某种因素的干扰，被刺激者难以集中精力对特定刺激作出反应的现象叫“斯特鲁普效应”，因为这个效应是斯特鲁普发现的，故以其名斯特鲁普效应。

二. 项目概述

在该项目中，使用描述统计学和统计检验分析实验心理学中的典型现象——斯特鲁普效应。获得对实验的初步认识，根据结果作出假设，查看其他执行了任务的人的数据，并计算这些统计数据。最后为读者提供直观的数据报告，并根据结果利用统计推断得出结论。

三. 实验样本数据集 (Udacity提供，以下简称本实验)

Congruent	Incongruent	Congruent	Incongruent
12.079	19.278	15.073	17.51
16.791	18.741	16.929	20.33
9.564	21.214	18.2	35.255
8.63	15.687	12.13	22.158
14.669	22.803	18.495	25.139
12.238	20.878	10.639	20.429
14.692	24.572	11.344	17.425
8.987	17.394	12.369	34.288
9.401	20.762	12.944	23.894
14.48	26.282	14.233	17.96
22.328	24.524	19.71	22.058
15.298	18.644	16.004	21.157

数据解释

- 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。

- b. 计量单位为秒，为方便计算数据省略单位

四. 实验数据的统计学描述和假设分析

1. 确认实验中的变量

- a. 本实验中的自变量是：文字条件是否一致，即显示文字和打印颜色是否匹配
- b. 本实验中的因变量是：说出同等大小的列表中的墨色名称的时间

2a. 建立假设

- a. 零假设符号： H_0 ，对立假设符号： H_A
- b. 文字条件一致的总体时间均值符号： μ_1 ，文字条件不一致的总体时间均值符号： μ_2
- c. 零假设文字描述：文字条件一致时说出同等大小的列表中的墨色名称的时间和文字条件不一致时没有区别，零假设符号表示 $H_0 : \mu_1 - \mu_2 = 0$
- d. 对立假设文字描述：文字条件一致时说出同等大小的列表中的墨色名称的时间和文字条件不一致时有区别，对立假设符号表示 $H_A : \mu_1 - \mu_2 \neq 0$

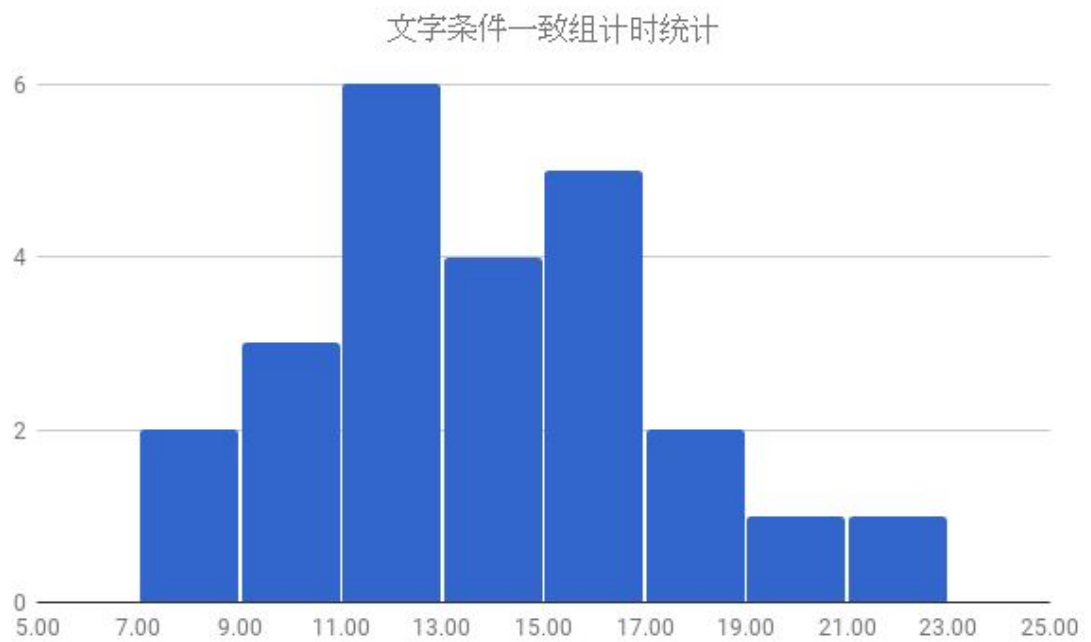
2b. 建立统计检验 (参考优达学城²、网易公开课³、互动百科⁴)

- a. 通过建立科学的统计检验方法，参考Interactive Stroop Effect Experiment⁵
- b. 本实验的抽样是相依样本，使用相依样本的统计学方法
- c. 基于我们建立的假设和样本数据，我们计算出相应的统计学数据
- d. 参考T临界值，对样本做 $\alpha = 0.05$ 的双尾t统计值检验
- e. 使用相依样本t检验而不是使用z检验的原因：如果我们知道总体的标准偏差，样本数量在30以上我们可以使用z检验，否则使用t检验
- f. 执行相依样本t检验满足的前提条件：样本来自于正态分布或者近似正态分布的总体，两样本是配对的，观察值数目相等顺序不能改变
- g. 根据统计检验结果进行决策，辩证所提出的假设

3. 描述性统计分析 (使用Google表格文档⁶)

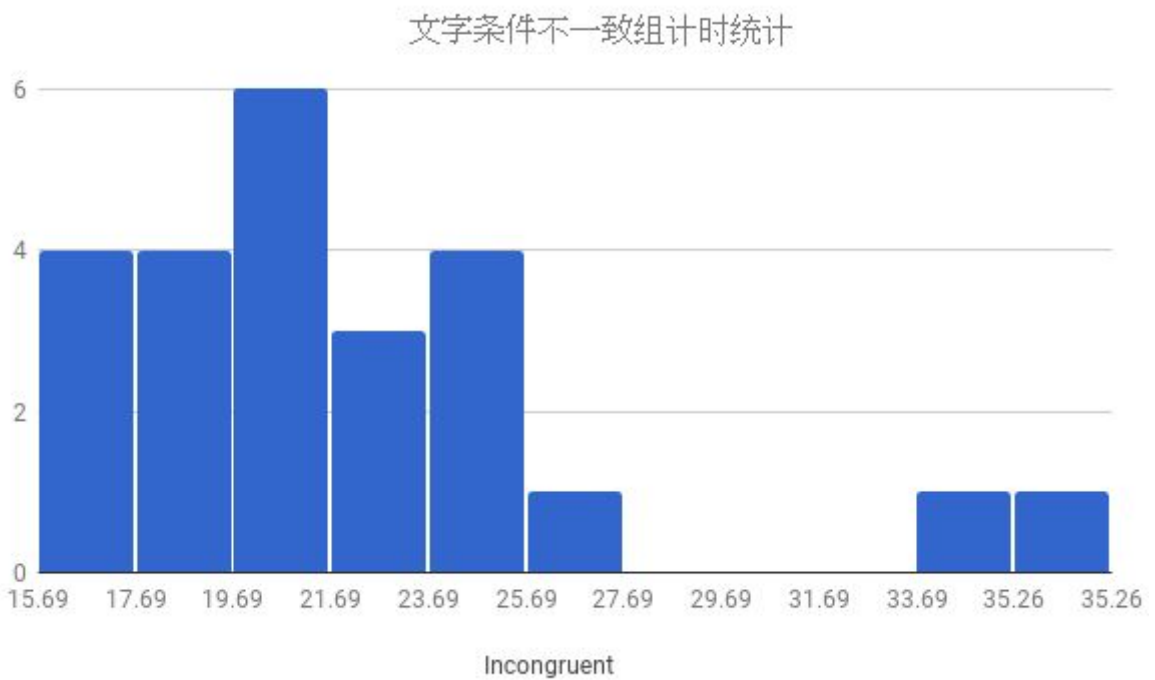
- a. 一致文字条件组的样本数量 $n_1 = 24$ 不一致文字条件组的样本数量 $n_2 = 24$
- b. 一致文字条件组的自由度 $df_1 = 23$ 不一致文字条件组的自由度 $df_2 = 23$
- c. 文字条件的两种情况为一对数据的两组
- d. 两组数据的差的均值为 $\bar{d} = -7.9648$
- e. \bar{d} 的标准差 $S_d = 4.8648$ ，标准误差 $SE(\bar{d}) = 0.9930$
- f. $t\text{-statistic} = 8.0209$
- g. $r^2 = 0.74$

4. 绘制数据图



直方图解释

此图反映了文字条件一致的样本组的用时统计分布，可以看出用时为13秒的样本数量最多，用时分布大概成正态分布，低于9秒和高于19秒的样本较少



直方图解释

此图反映了文字条件不一致的样本组的用时统计分布，大部分样本用时在15到25秒之间，集中度较高，用时特别高35秒的样本很少

5. 执行统计检验解释结果

- a. $\alpha = 0.05$, 自由度为23 , 双尾检验的 t 临界值为 ± 2.069
- b. 样本 t 统计值 = 8.0209 , $P < 0.05$ 存在显著差异 , 拒绝零假设 , 即零假设不成立
- c. 对立假设成立 , 受到文字条件不一致的干扰后 , 说出同等大小的列表中的墨色名称的时间增加。
- d. 根据 r^2 的数据可以得出 : 自变量(文字条件是否一致 , 即显示文字和打印颜色是否匹配)引起的差异在整体差异的比例为 74%

参考资料列表 (点击蓝色文字转跳)

1. [斯特鲁普效应Stroop effect](#) , 百度百科 , 20170612
2. [推论统计学](#) , Udacity , 20170613
3. [可汗学院：统计学](#) , 网易公开课 , 20170613
4. [统计检验](#) , 互动百科 , 2017013
5. [Interactive Stroop Effect Experiment](#) , Eric H. Chudler , 20170613
6. [Google表格文档](#) , Google , 20170613