

# 统计学：决策的科学项目说明

说明：[点此查看此文档的英文版本](#)。

## 背景信息

在一个 Stroop（斯特鲁普）任务中，参与者得到了一系列文字，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？

自变量：文字条件是否一致性；

因变量：说出正确墨色名称的时间。

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）。

假设  $\mu_{\text{congruent}}$  墨色与颜色词匹配上的总体均值， $\mu_{\text{incongruent}}$  为墨色与颜色词匹配不上的总体均值。

根据 wiki 的背景知识，以及数据集中的“不匹配” - “匹配”  $> 0$ ，所以可以预期 颜色-字义不匹配 比 颜色-字义匹配 需要更多的时间，所以采用右侧单尾检验。

零假设  $H_0$  为：说出正确墨色名称的时间 与 颜色词是否匹配上对应的油墨颜色 无关。

$\mu_{\text{congruent}} = \mu_{\text{incongruent}}$  ；

对立假设  $H_A$  为：说出正确墨色的时间 与 颜色词是否匹配上对应的油墨颜色 有关,并且 “颜色-字义匹配”的时间比 “颜色-字义不匹配”的时间段。  $\mu_{\text{congruent}} < \mu_{\text{incongruent}}$  。

确定假设的显著水平  $\alpha$ ，根据“小概率原理”，本实验采用  $\alpha=0.05$

检验采用配对 t 检验，理由如下

1. 数据集中只有样本数据，缺乏总体数据的标准偏差（t 检验前提）；
2. 数据量小于 30，呈现大概正态分布；（t 检验前提）
3. 由于测试中强调“每位参与者必须全部完成并记录每种条件下使用的时间”，是同一测试者在两种条件下的不同测试时间。所以适合用于配对 t 检验。（配对 t 检验前提）

现在轮到你自行尝试 Stroop 任务了。前往[此链接](#)，其中包含一个基于 Java 的小程序，专门用于执行 Stroop 任务。记录你收到的任务时间（你无需将时间提交到网站）。现在[下载此数据集](#)，其中包含一些任务参与者的结果。数据集的每行包含一名参与者的表现，第一个数字代表他们的一致任务结果，第二个数字代表不一致任务结果。

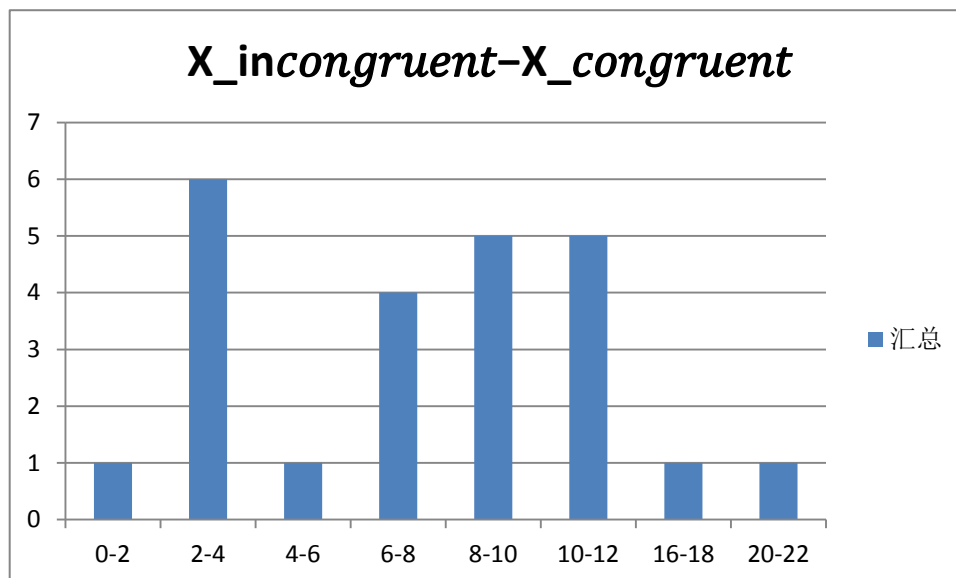
3. 报告关于此数据集的一些描述性统计。包含至少一个集中趋势测量和至少一个变异测量。

样本均值  $\bar{X}$  :  $\bar{X}_{congruent} = 14.05$  ,  $\bar{X}_{incongruent} = 22.02$

样本标准偏差 S:  $S_{congruent} = 3.56$  ,  $S_{incongruent} = 4.80$

样本量 n:  $n_{congruent} = 24$  ,  $n_{incongruent} = 24$  ;

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。



上图是  $X_{incongruent} - X_{congruent}$  的频数直方图  
可见图形正态分布。

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？

两列样本数据相减后，求出以下信息，此处用 匹配 减去 不匹配 ，作为新样本进行检查：

均值差异  $\bar{X}_D$  :  $\bar{X}_{incongruent} - \bar{X}_{congruent} = 7.96$

均值差异样本数 n :  $n = 24$

自由度 df :  $df=n-1=23$

均值差异标准差  $S_D$  :  $S_{\text{incongruent-congruent}} = 4.86$

均值差异标准误 SEM:  $\frac{S_D}{\sqrt{n}} = \frac{4.86}{\sqrt{24}} = 0.99$

t 统计量:

$$t\text{-statistic} = \frac{\bar{X}_D}{SEM} = \frac{7.96}{0.99} = 8.02$$

t 临界区:

设  $\alpha=0.05$ , 由于采用单尾检验, 则  $p=0.05$ , 依据 t 表, 查出  $df=23, p=0.05$ , 对应的临界区为:

t-critical= 1.714

t 分布图如下:

Table B		$t$ distribution critical values											
	Tail probability $p$												
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005	
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6	
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60	
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92	
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610	
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869	
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959	
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408	
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041	
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781	
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587	
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437	
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318	
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221	
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140	
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073	
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015	
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965	
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922	
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883	
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850	
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819	
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792	
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768	
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745	
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725	

由于 t 统计量 8.02 > t 临界值 1.714, 所以可以拒绝零假设(说出正确墨色名称的时间 与 颜色词是否匹配上对应的油墨颜色 无关), 说明显著水平为 0.05 时, 存在显著差异。

置信区间:

由于实验对立假设是 “颜色-字义不匹配” 比 “颜色-字义匹配” 的时间长, 均值差异是采用 “颜色-字义不匹配” 减去 “颜色-字义匹配” 获得, 所以采用单向右侧检验, 采用 95% 置信水平, 所以对应的置信空间为:

$$CI : \text{区间值为 } (-\infty, \bar{X}_D + t_{\text{critical}} \cdot SE) = (-\infty, 7.96 + 1.714 \cdot 0.99) = (-\infty, 9.67)$$

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！

以上可推论出油墨颜色与颜色词会导致人的误解，颜色与字体所代表的都是颜色，混淆测试者的判断，会导致说出正确墨色的时间变长。

优达学城

2016 年 9 月

资料：

<http://www.statisticshowto.com/when-to-use-a-t-score-vs-z-score/>

<http://support.minitab.com/zh-cn/minitab/17/topic-library/basic-statistics-and-graphs/hypothesis-tests/tests-of-means/types-of-t-tests/>

<http://discussions.youdaxue.com/t/topic/42566/2>

<http://www.csic.cornell.edu/Elrod/t-test/t-test-assumptions.html>

<http://discussions.youdaxue.com/t/10b-31-32/41966/2>

<https://wenku.baidu.com/view/3e5a2078a26925c52cc5bfee.html>

<https://www.zhihu.com/question/23448081/answer/132023721>