

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？

答：自变量是显示的文字条件，有两种水平，一是显示文字与它们的打印颜色匹配（一致文字条件），二是显示文字与它们打印颜色不匹配（不一致文字条件）。因变量是参与者说出同等大小列表中的墨色名称的时间。

2. 此任务的适当假设集是什么？你想执行什么类型的统计测试？为你的选择提供正当理由。

答：2a：

符号定义：零假设为 H_0 ，备择假设为 H_1 ，条件 1 为一致文字条件，总体均值为 μ_1 ，条件 2 为不一致文字条件，总体均值为 μ_2 。

零假设： $H_0: \mu_1 = \mu_2$ 参与者在两种条件下说出墨色名称的时间没有差异

备择假设： $H_1: \mu_1 \neq \mu_2$ 参与者在两种条件下说出墨色名称的时间有差异

此任务为双侧检验，因为我们没有定义两种条件下时间差异的大小。

2b：

执行的统计测试：配对样本 t 检验。

因为有两个条件，我们认为是两个样本的差异，总体方差未知，所以采用 t 检验。然后由于参与者同时接受两种不同的实验条件，因此采用配对样本 t 检验。

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。

答：该数据集的描述性统计如下：

描述统计

	样本	平均值	标准差
Incongruent	24	22.01592	4.797057
Congruent	24	14.05112	3.559358
有效样本数	24		

根据已知数据集，样本 $n=24$

条件 1 样本均值： $\bar{x}_1=14.05$ ，条件 2 样本均值： $\bar{x}_2=22.02$

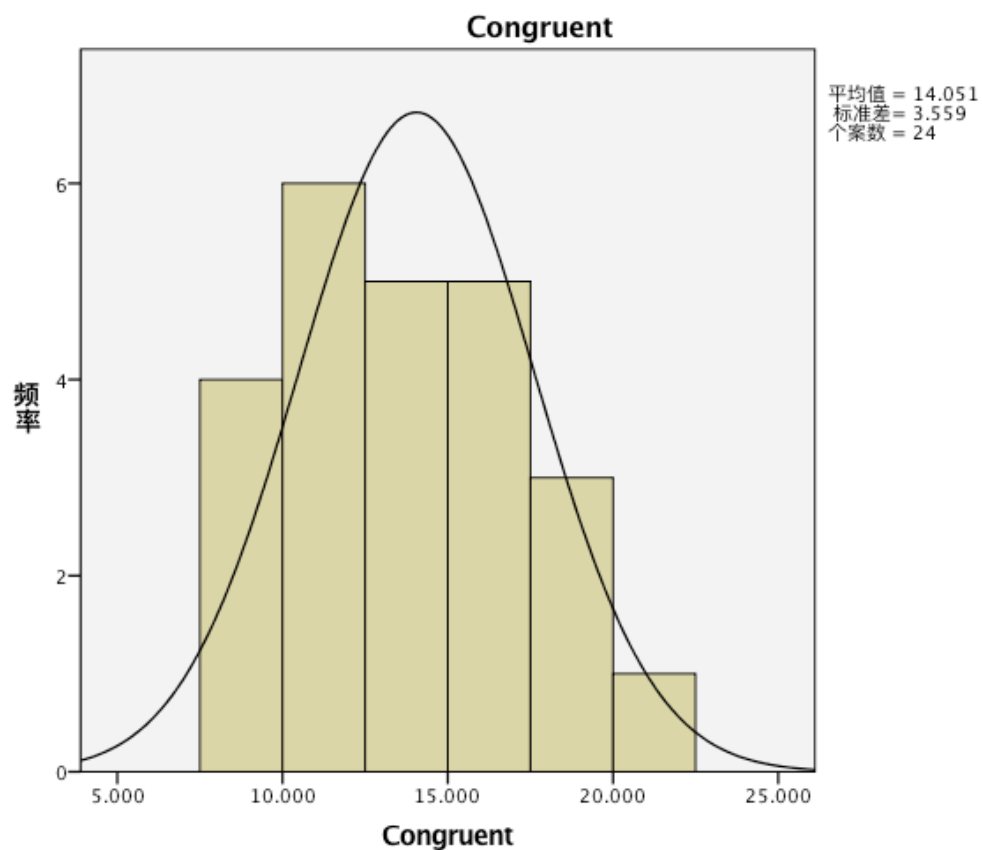
两者的差异为 $D= \bar{x}_2 - \bar{x}_1=22.02-14.05=7.97$

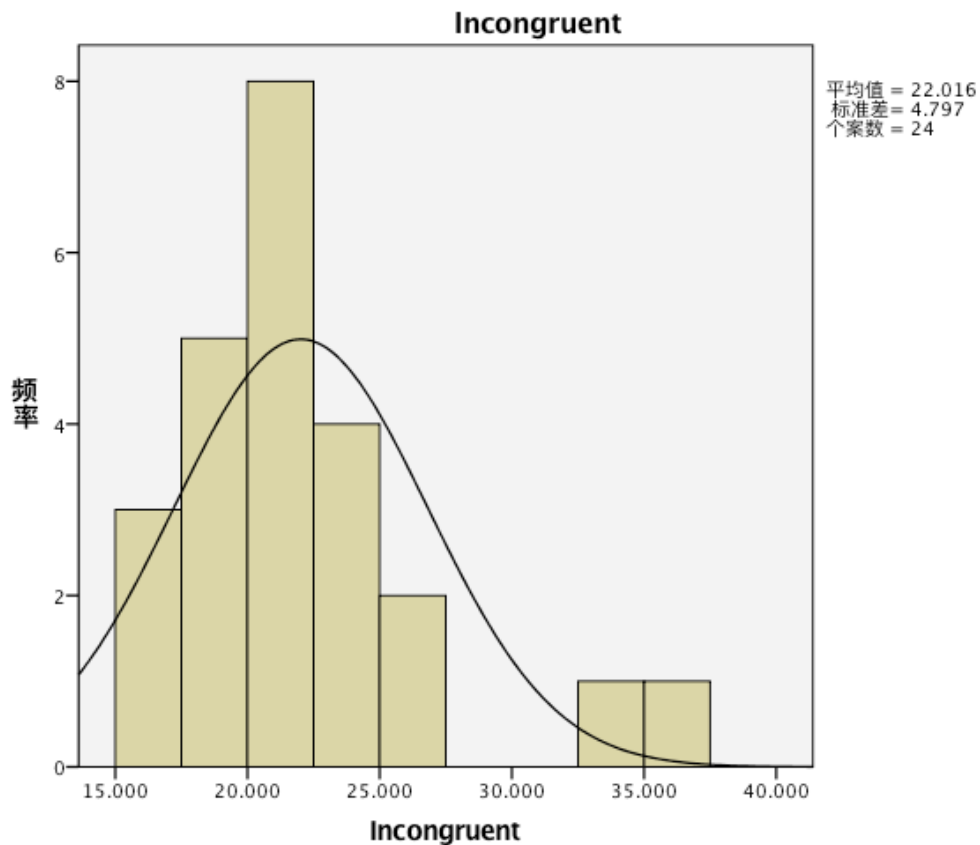
两者差异的标准差： $s=4.86$

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。

答：

可视化的直方图如下：





由以上两幅图可以看出，一致条件的反应时频率最高集中在 10.0~12.5，而不一致条件反应时频率最高集中在 20.0~20.5；不一致条件下反应时均值大于一致条件下的反应时；一致条件下的反应时分布更为集中，不一致条件下的反应时分布更为零散。

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？

		t	自由度	显著性（双尾）
配对 1	Congruent - Incongruent	-8.021	23	.000

答：置信水平 $\alpha = 0.01$

关键统计值为：

样本统计量 $t = -8.021$

则 $p = 0.0001$

该统计结果成功拒绝了零假设，因此我们可以认为参与者在一致条件下比不一致条件下说出墨色名称的时间更短，说明实验效果有效。

该结果与我期望的一致，因为当阅读不一致的墨色名称时，需要辨别反应时和选择反应时，增加了反应时间。

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：导致所观察到的效应的原因有以下：

1. **stroop** 范式中，参与者需要进行文字识别和颜色识别，需要的反应时间为，识别文字的时间，识别颜色的时间，读出颜色的时间，当条件一致时，不需要进行选择，因此直接能读出颜色，而当条件不一致时，参与者会受到语义的干扰，需要花更多的时间来选择读出颜色的名称，所以增加了反应时间。
2. 类似效应的任务有，内隐联想测验是基于反应时的对相关的两概念间联系紧密程度的一种测量。在实验任务中，被试需要对代表目标概念，诸如种族（白人、黑人），的词语或图片，和具有积极或消极性情的属性概念做出归类（通常是电脑屏幕左右方所提示的两个类别）。每个目标概念的类别将分别与积极或消极的属性概念相配对。对一种配对的更快的归类则表明该配对具有更强的联系或更容易被归在一起（例如，更快地对狗的概念与积极的概念的配对做出归类），因此就对此概念（狗）持相似的积极或消极的态度。

参考附录：

Greenwald, A.G., McGhee, D.E., & Schwartz, M. (2005). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.

[内隐联想测验 wiki](#)

[内隐联想测验程序](#)

Hawli

2017 年 7 月