

1. 自变量是什么，因变量是什么？

自变量：文字颜色与表达意思是否一致；

因变量：各个测试者分别说出不同条件下文字列表的时间。

2. 任务的适当假设集是什么？

原假设为两种文字条件下，测试者说出同等大小列表内容中的墨色名称所需时间的总体均值无明显区别；备择假设为两种文字条件下，测试者所需时间的总体均值有明显区别。

设：一致条件和不一致条件的测试时间样本分别来自不同的正态分布总体：

测试样本种类	样本大小	样本均值	样本标准差
一致条件	$n_1 = 24$	$\bar{x}_1 = 14.051$	$s_1 = 3.559$
不一致条件	$n_2 = 24$	$\bar{x}_2 = 22.016$	$s_2 = 4.797$

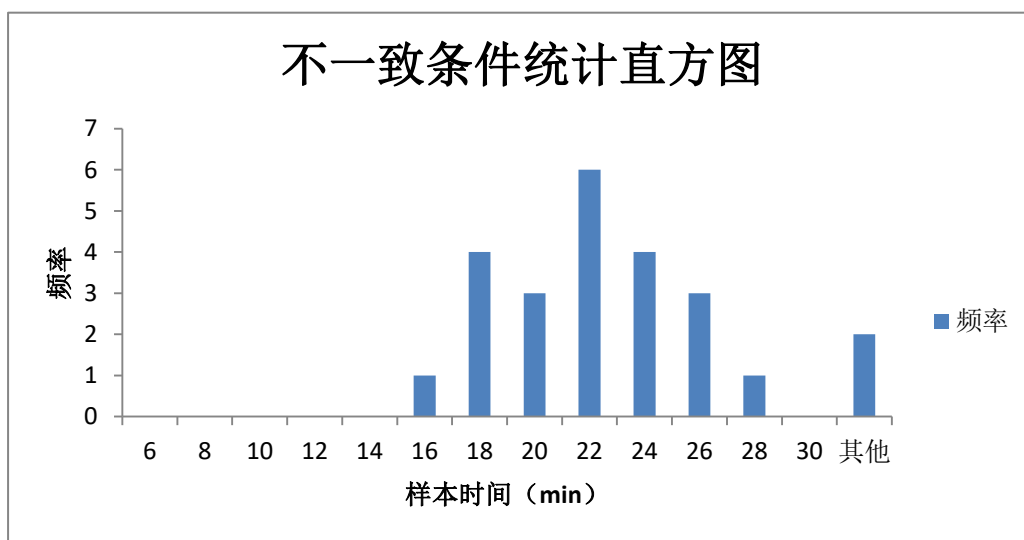
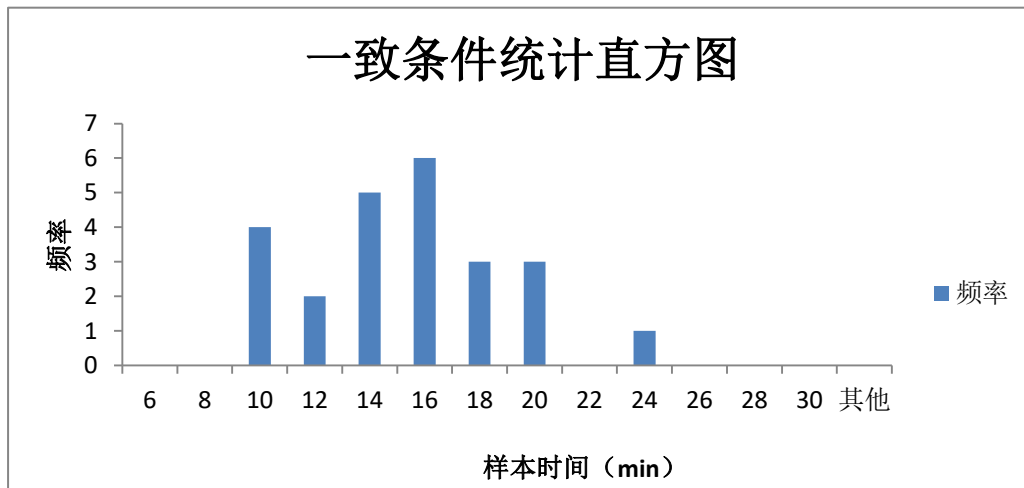
$H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$, 设显著性水平 $\alpha=0.05$, 其中 μ_1 为一致文字条件下测试时间的总体均值, μ_2 为不一致文字条件下的测试时间的总体均值。这里选择了双尾检验方法。

这里采用 t 检验。样本是对相同测试对象进行两组不同的测量, 属于“相关样本”, 并且假设两组样本各自大致满足正态分布条件, 方差相同, 只是均值可能不同; 数据不是全部测试者的结果, 属于“样本数据”, 其统计量属于“样本统计量”, 而 z 检验需要满足已知总体标准差, 且是大样本两个条件。

3. 数据集的一些描述性统计

参考上一小节的表格。

4. 样本数据的一两个可视化



从图中看出，一致性测试和不一致性测试相比，不一致性均值明显高于一致性测试结果。

5. 统计测试并报告结果，置信水平、关键统计值、是否拒绝零假设、结论？

$\mu_1 - \mu_2$ 的点估计值为 $\bar{x}_1 - \bar{x}_2 = -7.965$

两个样本逐项做差之后的差值，它的样本标准差 $s = 4.865$

t 统计值 = $\frac{\mu_1 - \mu_2}{s/\sqrt{n}} = -8.021$

t 临界值 = $\pm t_{0.025}(24-1) = \pm 2.069$

显然统计值落在了拒绝域，故拒绝 H_0 。测试结果显示，不一致测试和一致测试平均花费时间不能认为一致。

6. （可选）导致所观察到的原因是什么？

可以猜测，不一致性测试内容需要测试者花更多时间确定字体颜色与内容是否一致，故总体需要花费更多时间才能完成。