

# 变量

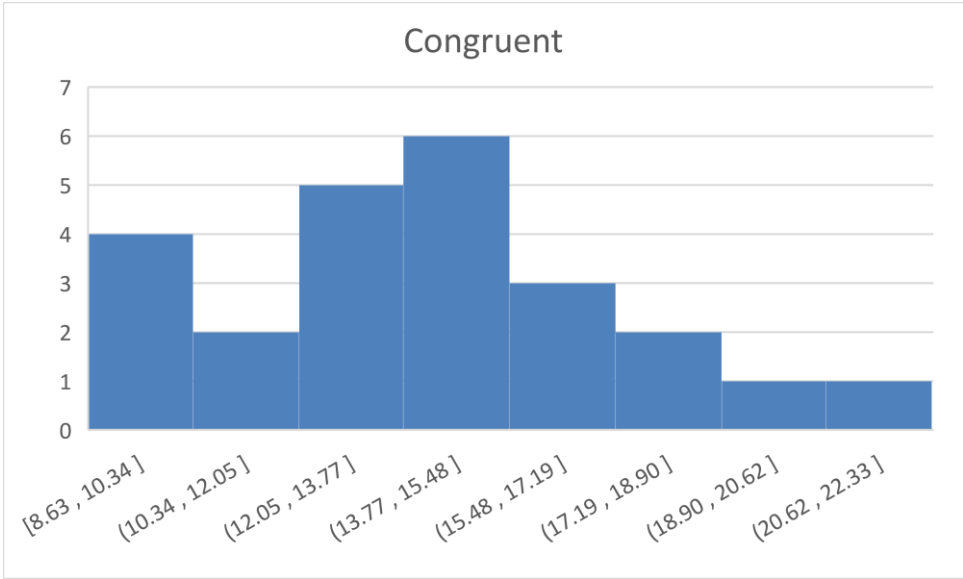
自变量是文字与颜色是否匹配  
因变量是参与者全部完成所使用的时间

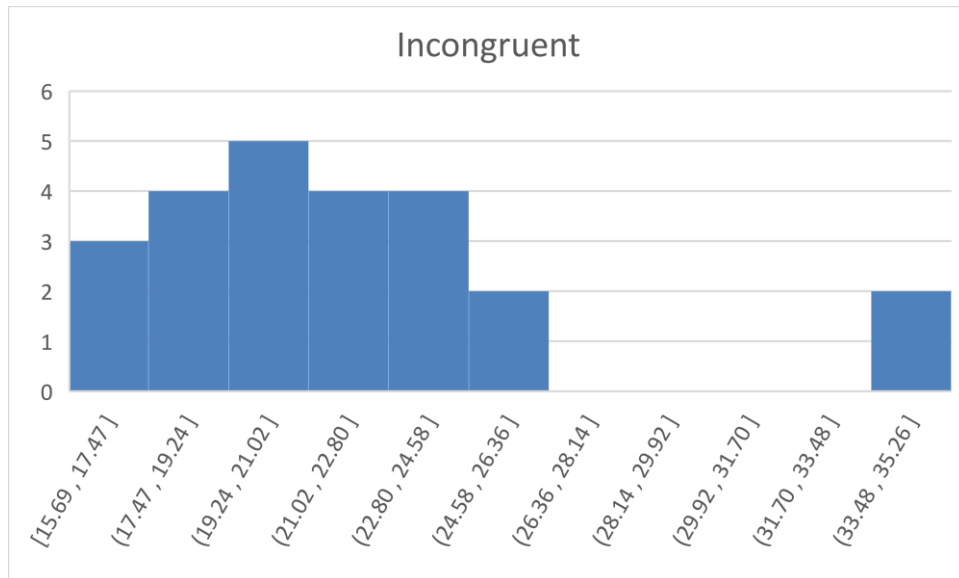
# 假设

适当假设：当文字与颜色匹配时，参与者全部完成所使用的时间更少。  
零假设：文字与颜色匹配不匹配，参与者全部完成所使用的时间无明显差异。  
对立假设：文字与颜色匹配时，参与者全部完成所使用的时间更少。  
这里用 $H_0$ 表示零假设，用 $H_1$ 表示对立假设，用 $\mu_c$ 表示文字与颜色匹配时参与者使用时间的总体均值，用 $\mu_i$ 表示文字与颜色不匹配时参与者使用时间的总体均值。

$$H_0 : \mu_c = \mu_i$$
$$H_1 : \mu_c < \mu_i$$

# 统计检验





Z 检验适用于知道总体参数(如 $\mu, \sigma$ )的情况，而这里并不知道总体参数，只有样本，需要比较两个样本之间的差异，来推断总体的情况，故而考虑 t 检验。

t 检验的前提条件是总体服从正态分布和方差齐性。

对两个样本进行分析，其均表现出类似正态分布（如上图所示）。且实验参与者是从人群中随机挑选的，并没有条件限制，故实验样本是随机样本。因此可以推测其总体满足正态分布。

此实验是统计同一组参与者对文字与颜色是否匹配所使用的时间，因此这两个样本是相关样本。故而可以推断其总体的方差近似相等。

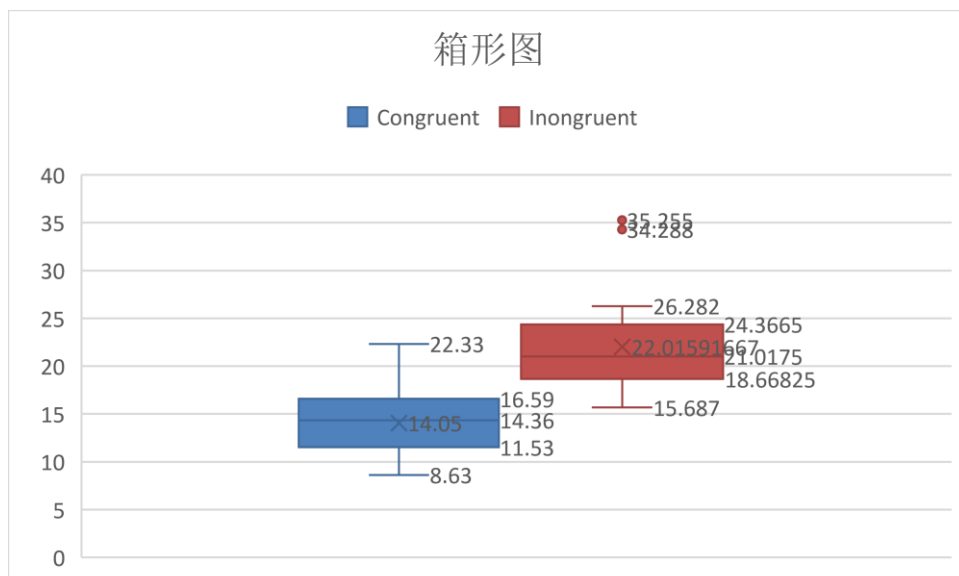
综上，选择用相关样本均值差异的配对 t 检验。

由于 $H_1$ 包含方向，所以采用单尾检验。

选取显著性水平  $\alpha = 0.05$

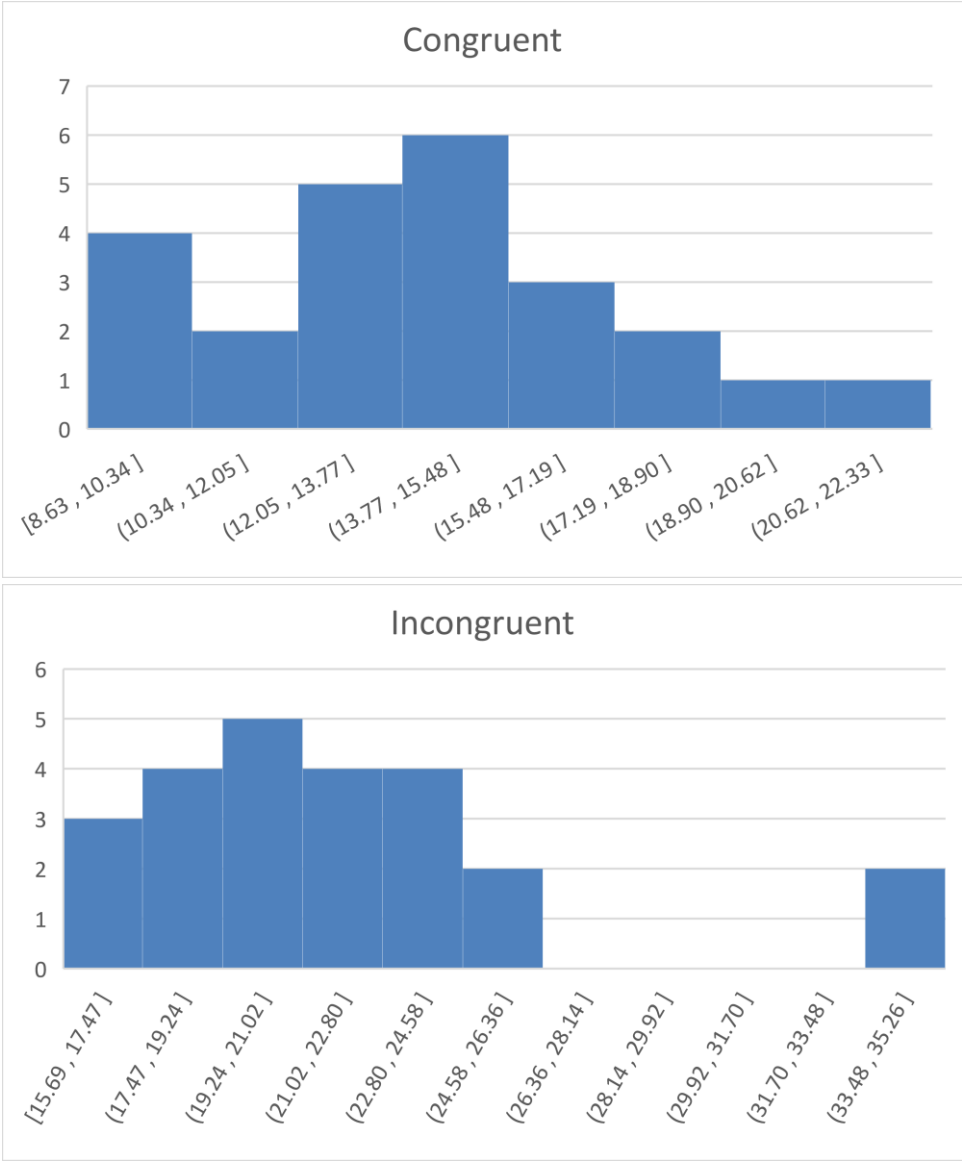
## 样本数据可视化

箱形图：



由箱形图可知，文字和颜色匹配的样本中，四分位差较小，但最大值和最小值差较大；文字和颜色不匹配的样本中，四分位差较大，但最大值和最小值差较小，且有两个异常点存在。

直方图：



从两个样本数据的直方图上我们可以看出，文字与颜色匹配所使用的时间，大多集中在 10 到 18 之间。文字与颜色不匹配所使用的时间，大多集中在 15 到 25 之间。

## 数据集的统计描述

匹配的平均值 $\bar{x}_c=14.05$

不匹配的平均值 $\bar{x}_i=22.02$

两者的平均值差 $\bar{x}_D = \bar{x}_c - \bar{x}_i = -7.96$

标准差 $s_d = \sqrt{\frac{\sum (x_i - \bar{x})^2}{(n-1)}} = 4.86$

标准误差  $SEM = s_d / \sqrt{n} = 0.99$

t 统计量  $t_{STAT} = \frac{\overline{x_D}}{SEM} = -8.02$

自由度  $df = 23$

t 临界值  $t_{critical} = 1.714$

效应量  $r^2 = \frac{t^2}{t^2 + df} = 0.74$ ，即有 74% 的差异是由文字和颜色匹不匹配所造成的的

自由度 23，95% 置信区间的 t 临界值是 2.069，误差界限为  $t_{0.025} * SEM = 2.05$

置信区间  $CI = \overline{x_D} \pm 2.05 = -7.96 \pm 2.05 = (-10.01, -5.91)$

## 决策

由于  $t_{STAT} = -8.02$  落入 -1.714 以左的拒绝域， $P < 0.0005$ ，零假设被拒绝。所以有理由相信文字与颜色匹配时，参与者全部完成所使用的时间更少。

## 参考文献

《商务统计学第 5 版》 作者：美)戴维·M. 莱文 出版社：中国人民大学出版社  
[https://en.wikipedia.org/wiki/Stroop\\_effect](https://en.wikipedia.org/wiki/Stroop_effect)