**检验心理学现象**

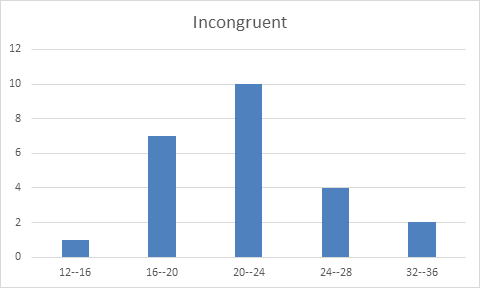
作业地址：https://classroom.udacity.com/nanodegrees/nd009-cn-basic/parts/3151ab46-a95f-420a-a0e3-f2dd677b2f3e/modules/56375cce-b181-4b39-aa2f-df05e8293e66/lessons/4582204201239847/concepts/45861894150923

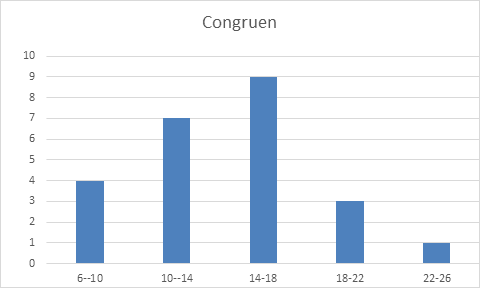
**1. 我们的自变量是什么？因变量是什么？**

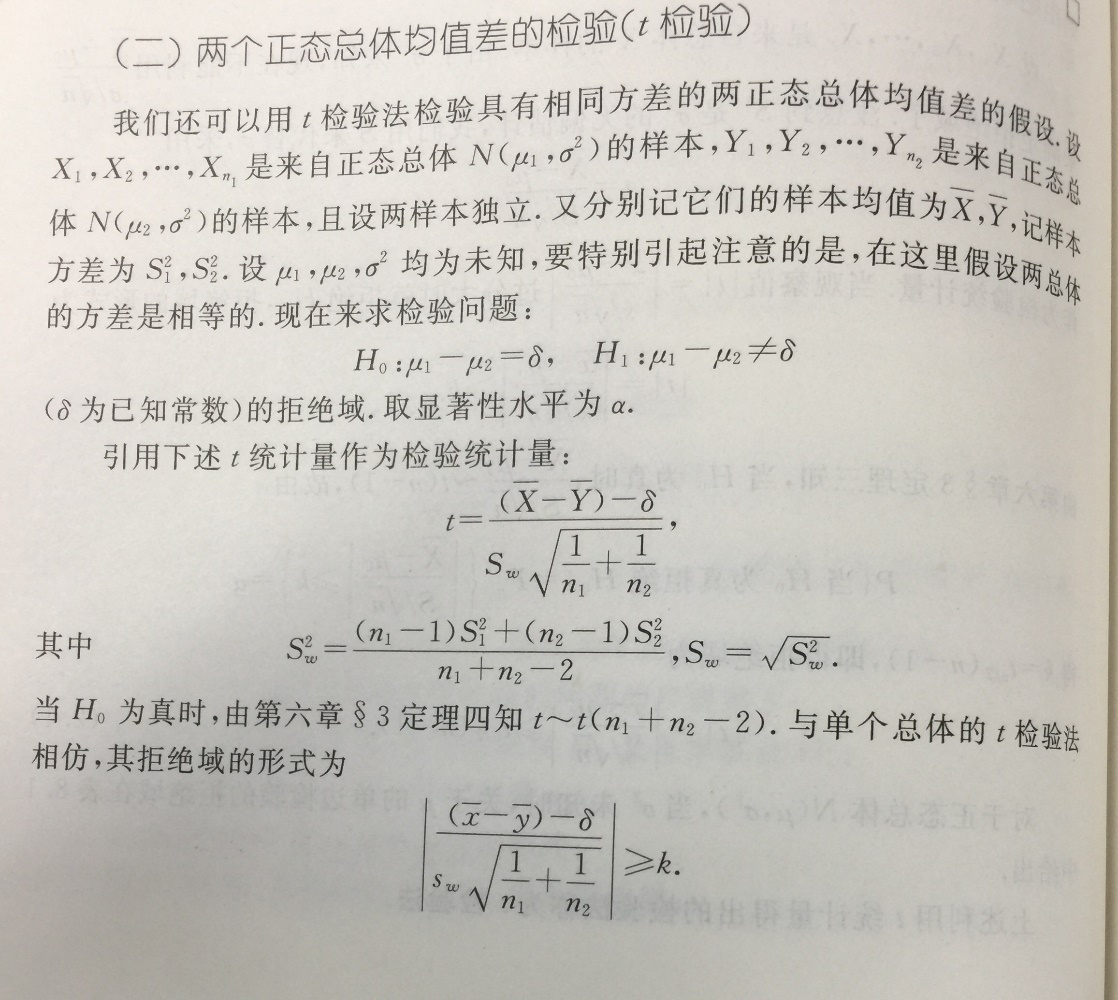
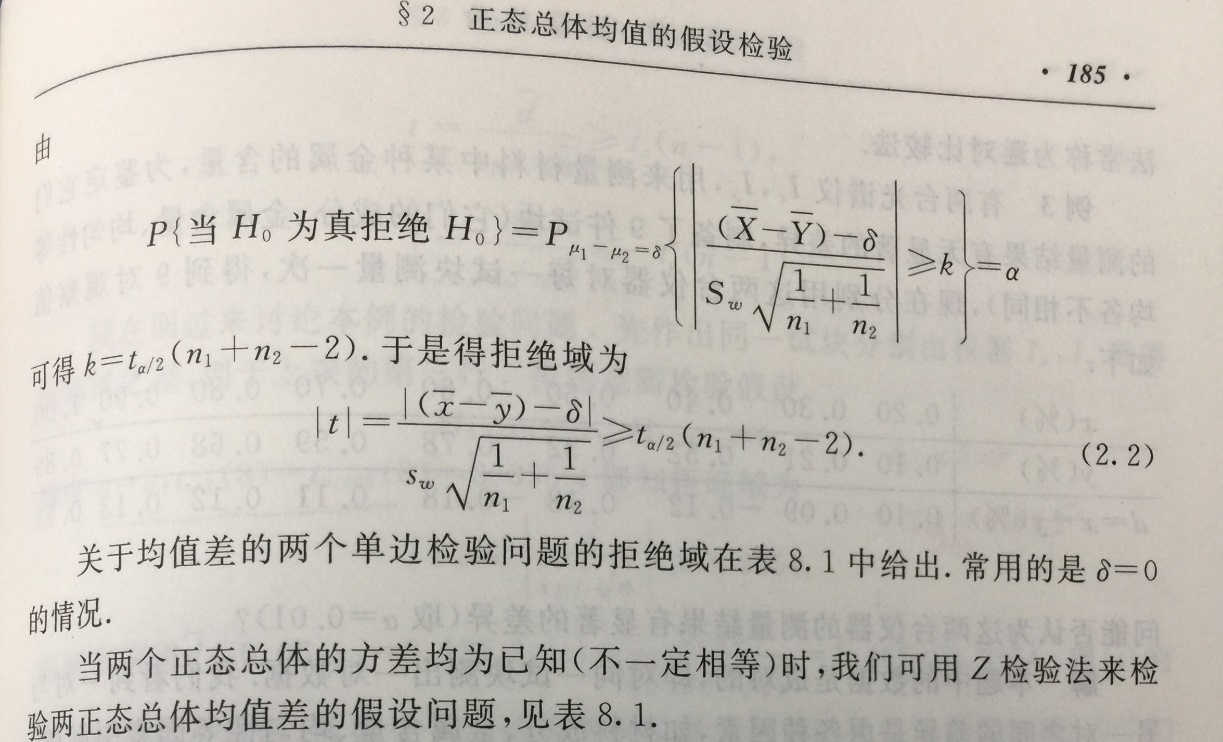
* 自变量：一致文字条件，和不一致文字条件。这两种可能。
* 因变量：说出同等大小的列表中的墨色名称的时间。

**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）。**

* 假设集：
  + 零假设*H*​0​​：两种情况下，所使用的时间并没有显著差异（*u*​*con*​​−*u*​*incon*​​=0）；
  + 对立假设H\_1*H*​1​​：两种情况下，所使用的时间有显著差异（*u*​*con*​​−*u*​*incon*​​≠0）。
* 统计检验的选择及理由
  + 两个正太总体均值差的检验--t检验，理由如下：
  + 由题意知，样本为随机样本；
  + 分析所得数据，处理后画出柱状图，可推测随机试验总体服从正态分布；



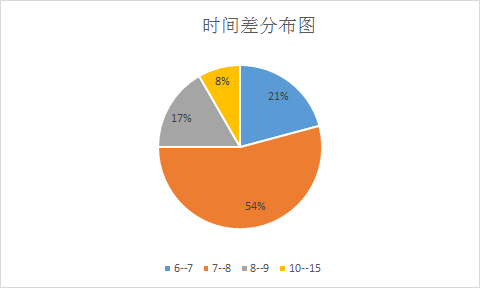


* + 试验是同一受试者参与两组条件不同，其余变量受控而得两个样本数据，可以合理推测其总体方差相近似。
  + 上诉3个条件满足《概率论与数理统计》（第四版）（高等教育出版社）184页。两个正太总体均值差的检验（t检验）条件。内容如下：
* 
* 

**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。**

* 样本量：*n* = 24
* 自由度：*d*​*f*​​ = 23
* “一致条件”的样本均值: *u*​*con*​​ = 14.051125
* “不一致条件”的样本均值*u*​*incon* ​​= 22.01591666666667
* 样本均值的差值*u*​*D*​​ = 7.9647916666666685
* “一致条件”的样本方差：*σ*​*con*​2 ​​= 12.141152859375003
* “不一致条件”的样本方差：*σ*​*incon*​2​​ = 22.05293382638889
* 样本差值的方差*σ*​*D*​2 ​​= 3.360664248

**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。**

* 时间差值饼状图
* 
* 92%的时间差值在5-10s之间。

**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？**

* 取\alpha*α*水平为0.05
* 自由度为23，拒绝域为： |t|\geq t\_{\alpha/2(46)}=2.0687∣*t*∣≥*t*​*α*/2(46)​​=2.0687
* t统计值：t=6.672745133475093*t*=6.672745133475093
* 效应量：0.65938807423042280.6593880742304228
* 综上，t统计值落在拒绝域中，拒绝零假设H\_0*H*​0​​。说明两种情况下所使用的时间，有统计上的显著差异，并且“不一致”情况所使用的时间会比“一致”情况多6-10秒。有65。94%的差异是由于显示文字与打印颜色不一致造成的。该结果与期望一致。

**6. Python计算代码**

* 本次计算写了数行Python代码辅助计算，代码如下：

*#! /usr/bin/env python3*

*# -\*- coding: utf-8 -\*-*

*# Author:*

*# Date: 2017/8/23*

**import** math

**class** **ArrayAbc**(object):

**def** **\_\_init\_\_**(self, array):

self.array = [float(i) **for** i **in** array]

self.count = self.count()

self.sum = sum(self.array)

self.avg = self.average()

**def** **count**(self):

**return** len(self.array)

**def** **average**(self):

"""求平均数"""

**if** self.count < 1:

**return** 0

**else**:

**return** self.sum / self.count

**def** **median**(self):

"""求中位数"""

**if** self.count < 1:

**return** **None**

**else**:

self.array.sort()

**return** self.array[len(self.array) // 2]

**def** **variance**(self):

"""求方差"""

**if** self.count < 1:

**return** **None**

**else**:

li = [(k-self.avg)\*\*2 **for** k **in** self.array]

s = sum(li)

**return** s / self.count

**class** **TTest**(object):

"""两个正太总体均值差的检验（t检验）。"""

**def** **\_\_init\_\_**(self, n1, n2, x, y, s1, s2):

self.n1 = float(n1)

self.s1 = float(s1)

self.x = float(x)

self.y = float(y)

self.n2 = float(n2)

self.s2 = float(s2)

**def** **get\_sw**(self):

ss = (self.n1 - 1) \* self.s1 + (self.n2 - 1) \* self.s2

ss\_avg = ss / (self.n1 + self.n2 - 2)

**return** math.sqrt(ss\_avg)

**def** **get\_t**(self):

a = self.x - self.y

w = self.get\_sw()

b = w \* math.sqrt(1 / self.n1 + 1 / self.n2)

**return** abs(a / b)

**def** **get\_r**(self):

"""效应量"""

t = self.get\_t()

**return** t\*\*2 / (t\*\*2 + self.n1 - 1)

**if** \_\_name\_\_ == '\_\_main\_\_':

Congruent = [

12.079, 16.791, 9.564, 8.63, 14.669, 12.238, 14.692, 8.987, 9.401, 14.48, 22.328, 15.298,

15.073, 16.929, 18.2, 12.13, 18.495, 10.639, 11.344, 12.369, 12.944, 14.233, 19.71, 16.004,

]

In\_congruent = [

19.278, 18.741, 21.214, 15.687, 22.803, 20.878, 24.572, 17.394, 20.762, 26.282, 24.524, 18.644,

17.51, 20.33, 35.255, 22.158, 25.139, 20.429, 17.425, 34.288, 23.894, 17.96, 22.058, 21.157,

]

obj\_a = ArrayAbc(Congruent)

obj\_b = ArrayAbc(In\_congruent)

n1 = obj\_a.count

n2 = obj\_b.count

x = obj\_a.avg

y = obj\_b.avg

s1 = obj\_a.variance()

s2 = obj\_b.variance()

obj\_t = TTest(n1, n2, x, y, s1, s2)

print("X平均：", x)

print("Y平均：", y)

print("均值差值：", y-x)

print("X方差：", s1)

print("Y方差：", s2)

print("t统计值:", obj\_t.get\_t())

print("效应量r:", obj\_t.get\_r())

"""

X平均： 14.051125

Y平均： 22.01591666666667

均值差值： 7.9647916666666685

X方差： 12.141152859375003

Y方差： 22.05293382638889

t统计值: 6.672745133475093

效应量r: 0.6593880742304228

"""

**7. 参考文献**

* [1] 盛骤等.概率论与数理统计（第四版）[M].浙江大学：高等教育出版社，1979.
* [2] 斯普鲁斯数据：<https://d17h27t6h515a5.cloudfront.net/topher/2016/September/57ce3363_stroopdata/stroopdata.csv>
* [3] Markdown中的数学符号书写：[http://jzqt.github.io/2015/06/30/Markdown中写数学公式/](http://jzqt.github.io/2015/06/30/Markdown%E4%B8%AD%E5%86%99%E6%95%B0%E5%AD%A6%E5%85%AC%E5%BC%8F/)