

到，人口数量和高中毕业率的相关系数（-0.10）并不显著地不为0（ $p=0.5$ ）。

#### 其他显著性检验

在7.4.1节中，我们关注了偏相关系数。在多元正态性的假设下，`psych`包中的`pcor.test()`函数<sup>①</sup>可以用来检验在控制一个或多个额外变量时两个变量之间的条件独立性。使用格式为：

```
pcor.test(r, q, n)
```

其中的`r`是由`pcor()`函数计算得到的偏相关系数，`q`为要控制的变量数（以数值表示位置），`n`为样本大小。

在结束这个话题之前应当指出的是，`psych`包中的`r.test()`函数提供了多种实用的显著性检验方法。此函数可用来检验：

- ❑ 某种相关系数的显著性；
- ❑ 两个独立相关系数的差异是否显著；
- ❑ 两个基于一个共享变量得到的非独立相关系数的差异是否显著；
- ❑ 两个基于完全不同的变量得到的非独立相关系数的差异是否显著。

参阅`help(r.test)`以了解详情。

### 7.3.3 相关关系的可视化

以相关系数表示的二元关系可以通过散点图和散点图矩阵进行可视化，而相关图（*correlogram*）则为以一种有意义的方式比较大量的相关系数提供了一种独特而强大的方法。这些图形将在第11章中详述。

## 7.4 t 检验

在研究中最常见的行为就是对两个组进行比较。接受某种新药治疗的患者是否较使用某种现有药物的患者表现出了更大程度的改善？某种制造工艺是否较另外一种工艺制造出的不合格品更少？两种教学方法中哪一种更有效？如果你的结果变量是类别型的，那么可以直接使用7.3节中阐述的方法。这里我们将关注结果变量为连续型的组间比较，并假设其呈正态分布。

为了阐明方法，我们将使用`MASS`包中的`UScrime`数据集。它包含了1960年美国47个州的刑罚制度对犯罪率影响的信息。我们感兴趣的结果变量为`Prob`（监禁的概率）、`U1`（14~24岁年龄段城市男性失业率）和`U2`（35~39岁年龄段城市男性失业率）。类别型变量`so`（指示该州是否位于南方的指示变量）将作为分组变量使用。数据的尺度已被原始作者缩放过。（注意，我原本打算将本节命名为“旧南方的罪与罚”，但是最后理智还是战胜了情感。）

### 7.4.1 独立样本的t检验

如果你在美国的南方犯罪，是否更有可能被判监禁？我们比较的对象是南方和非南方各州，

① 这里可能是作者的笔误，函数`pcor.test`事实上包含于`ggm`包中。——译者注

因变量为监禁的概率。一个针对两组的独立样本t检验可以用于检验两个总体的均值相等的假设。这里假设两组数据是独立的，并且是从正态总体中抽得。检验的调用格式为：

```
t.test(y ~ x, data)
```

其中的 $y$ 是一个数值型变量， $x$ 是一个二分变量。调用格式或为：

```
t.test(y1, y2)
```

其中的 $y_1$ 和 $y_2$ 为数值型向量（即各组的结果变量）。可选参数 $data$ 的取值为一个包含了这些变量的矩阵或数据框。与其他多数统计软件不同的是，这里的t检验默认假定方差不相等，并使用Welsh的修正自由度。你可以添加一个参数`var.equal=TRUE`以假定方差相等，并使用合并方差估计。默认的备择假设是双侧的（即均值不相等，但大小的方向不确定）。你可以添加一个参数`alternative="less"`或`alternative="greater"`来进行有方向的检验。

在下列代码中，我们使用了一个假设方差不等的双侧检验，比较了南方（group 1）和非南方（group 0）各州的监禁概率：

```
> library(MASS)
> t.test(Prob ~ So, data=UScrime)

Welch Two Sample t-test

data:  Prob by So
t = -3.8954, df = 24.925, p-value = 0.0006506
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.03852569 -0.01187439
sample estimates:
mean in group 0 mean in group 1
 0.03851265      0.06371269
```

你可以拒绝南方各州和非南方各州拥有相同监禁概率的假设（ $p < .001$ ）。

**注意** 由于结果变量是一个比例值，你可以在执行t检验之前尝试对其进行正态化变换。在本例中，所有对结果变量合适的变换（ $Y/1-Y$ 、 $\log(Y/1-Y)$ 、 $\arcsin(Y)$ 、 $\arcsin(\sqrt{Y})$ ）都会将检验引向相同的结论。数据变换详述于第8章。

## 7.4.2 非独立样本的t检验

再举个例子，你可能会问：较年轻（14~24岁）男性的失业率是否比年长（35~39岁）男性的失业率更高？在这种情况下，这两组数据并不独立。你不能说亚拉巴马州的年轻男性和年长男性的失业率之间没有关系。在两组观测之间相关时，你获得的是一个非独立组设计（dependent groups design）。前-后测设计（pre-post design）或重复测量设计（repeated measures design）同样也会产生非独立的组。

非独立样本的t检验假定组间的差异呈正态分布。对于本例，检验的调用格式为：

```
t.test(y1, y2, paired=TRUE)
```

其中的 $y_1$ 和 $y_2$ 为两个非独立组的数值向量。结果如下：

```
> library(MASS)
> sapply(UScrime[c("U1", "U2")], function(x) (c(mean=mean(x), sd=sd(x))))
      U1      U2
mean 95.5 33.98
sd   18.0  8.45

> with(UScrime, t.test(U1, U2, paired=TRUE))

      Paired t-test

data:  U1 and U2
t = 32.4066, df = 46, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 57.67003 65.30870
sample estimates:
mean of the differences
      61.48936
```

差异的均值（61.5）足够大，可以保证拒绝年长和年轻男性的平均失业率相同的假设。年轻男性的失业率更高。事实上，若总体均值相等，获取一个差异如此大的样本的概率小于0.000 000 000 000 000 22（即 $2.2e-16$ ）。

### 7.4.3 多于两组的情况

如果想在多于两个的组之间进行比较，应该怎么做？如果能够假设数据是从正态总体中独立抽样而得的，那么你可以使用方差分析（ANOVA）。ANOVA是一套覆盖了许多实验设计和准实验设计的综合方法。就这一点来说，它的内容值得单列一章。你可以随时离开本节转而阅读第9章。

## 7.5 组间差异的非参数检验

如果数据无法满足t检验或ANOVA的参数假设，可以转而使用非参数方法。举例来说，若结果变量在本质上就严重偏倚或呈现有序关系，那么你可能会希望使用本节中的方法。

### 7.5.1 两组的比较

若两组数据独立，可以使用Wilcoxon秩和检验（更广为人知的名字是Mann - Whitney U检验）来评估观测是否是从相同的概率分布中抽得的（即，在一个总体中获得更高得分的概率是否比另一个总体要大）。调用格式为：

```
wilcox.test(y ~ x, data)
```

其中的 $y$ 是数值型变量，而 $x$ 是一个二分变量。调用格式或为：

```
wilcox.test(y1, y2)
```