10.2　用 pwr 包做功效分析

Stéphane Champely开发的pwr包可以实现Cohen（1988）描述的功效分析。表10-1列出了一些非常重要的函数。对于每个函数，用户可以设定四个量（样本大小、显著性水平、功效和效应值）中的三个量，第四个量将由软件计算出来。

表10-1　**pwr**包中的函数

函　　数	功效计算的对象
pwr.2p.test()	两比例（n相等）
pwr.2p2n.test()	两比例（n不相等）
pwr.anova.test()	平衡的单因素ANOVA
pwr.chisq.test()	卡方检验
pwr.f2.test()	广义线性模型
pwr.p.test()	比例（单样本）
pwr.r.test()	相关系数
pwr.t.test()	t检验（单样本、两样本、配对）
pwr.t2n.test()	t检验（n不相等的两样本）

四个量中，效应值是最难规定的。计算效应值通常需要一些相关估计的经验和对过去研究知识的理解。但是如果在一个特定的研究中，你对需要的效应值一无所知，该怎么做呢？10.2.7节将会讨论这个难题。本节接下来介绍pwr包在常见统计检验中的应用。在调用以上函数时，请确定已经安装并载入pwr包。

10.2.1　t 检验

对于t检验，pwr.t.test()函数提供了许多有用的功效分析选项，格式为：

```
pwr.t.test(n=, d=, sig.level=, power=, alternative=)
```

其中元素解释如下。

- n为样本大小。
- d为效应值，即标准化的均值之差。

$$d = \frac{\mu_1 - \mu_2}{\sigma} \qquad \text{其中} \quad \begin{aligned} &\mu_1 = \text{组1均值} \\ &\mu_2 = \text{组2均值} \\ &\sigma^2 = \text{误差方差} \end{aligned}$$

- sig.level表示显著性水平（默认为0.05）。
- power为功效水平。
- type指检验类型：双样本t检验（two.sample）、单样本t检验（one.sample）或相依样本t检验（paired）。默认为双样本t检验。

- alternative指统计检验是双侧检验（two.sided）还是单侧检验（less或greater）。默认为双侧检验。

让我们举例说明函数的用法。仍继续10.1节使用手机与驾驶反应时间的实验，假定将使用双尾独立样本t检验来比较两种情况下驾驶员的反应时间均值。

如果你根据过去的经验知道反应时间有1.25 s的标准偏差，并认定反应时间1 s的差值是巨大的差异，那么在这个研究中，可设定要检测的效应值为d=1/1.25=0.8或者更大。另外，如果差异存在，你希望有90%的把握检测到它，由于随机变异性的存在，你也希望有95%的把握不会误报差异显著。这时，对于该研究需要多少受试者呢？

将这些信息输入到pwr.t.test()函数中，形式如下：

```
> library(pwr)
> pwr.t.test(d=.8, sig.level=.05, power=.9, type="two.sample",
   alternative="two.sided")

     Two-sample t test power calculation

              n = 34
              d = 0.8
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

 NOTE: n is number in *each* group
```

结果表明，每组中你需要34个受试者（总共68人），这样才能保证有90%的把握检测到0.8的效应值，并且最多5%的可能性会误报差异存在。

现在变化一下这个问题。假定在比较这两种情况时，你想检测到总体均值0.5个标准偏差的差异，并且将误报差异的几率限制在1%内。此外，你能获得的受试者只有40人。那么在该研究中，你能检测到这么大总体均值差异的概率是多少呢？

假定每种情况下受试者数目相同，可以如下操作：

```
> pwr.t.test(n=20, d=.5, sig.level=.01, type="two.sample",
  alternative="two.sided")

     Two-sample t test power calculation

              n = 20
              d = 0.5
      sig.level = 0.01
          power = 0.14
   alternative = two.sided

NOTE: n is number in *each* group
```

结果表明，在0.01的先验显著性水平下，每组20个受试者，因变量的标准差为1.25 s，有低于14%的可能性断言差值为0.625 s或者不显著（d=0.5=0.625/1.25）。换句话说，你将有86%的可能性错过你要寻找的效应值。因此，可能需要慎重考虑要投入到该研究中的时间和精力。

上面的例子都是假定两组中样本大小相等，如果两组中样本大小不同，可用函数：

```
pwr.t2n.test(n1=, n2=, d=, sig.level=, power=, alternative=)
```

此处，n1和n2是两组的样本大小，其他参数含义与pwr.t.test()的相同。可以尝试改变pwr.t2n.test()[1]函数中的参数值，看看输出的效应值如何变化。

10.2.2 方差分析

pwr.anova.test()函数可以对平衡单因素方差分析进行功效分析。格式为：

```
pwr.anova.test(k=, n=, f=, sig.level=, power=)
```

其中，k是组的个数，n是各组中的样本大小。

对于单因素方差分析，效应值可通过f来衡量：

$$f=\sqrt{\frac{\sum_{i-1}^{k} p_i\times(\mu_i-\mu)^2}{\sigma^2}}$$

其中，$p_i = n_i/N$，

n_i = 组i的观测数目

N = 总观测数目

μ_i = 组i均值

μ = 总体均值

σ^2 = 组内误差方差

让我们举例说明函数用法。现对五个组做单因素方差分析，要达到0.8的功效，效应值为0.25，并选择0.05的显著性水平，计算各组需要的样本大小。代码如下：

```
> pwr.anova.test(k=5, f=.25, sig.level=.05, power=.8)

     Balanced one-way analysis of variance power calculation

              k = 5
              n = 39
              f = 0.25
      sig.level = 0.05
          power = 0.8

 NOTE: n is number in each group
```

结果表明，总样本大小为5 × 39，即195。注意，本例中需要估计在同方差时五个组的均值。如果你对上述情况都一无所知，10.2.7节提供的方法可能会有所帮助。

10.2.3 相关性

pwr.r.test()函数可以对相关性分析进行功效分析。格式如下：

```
pwr.r.test(n=, r=, sig.level=, power=, alternative=)
```

其中，n是观测数目，r是效应值（通过线性相关系数衡量），sig.level是显著性水平，power是功效水平，alternative指定显著性检验是双边检验（tow.sided）还是单边检验（less或greater）。

假定正在研究抑郁与孤独的关系。你的零假设和研究假设为：

[1] R中函数名称后面最好加上()。——译者注

H_0: ρ ≤ 0.25 和 H_1: ρ > 0.25
其中，ρ是两个心理变量的总体相关性大小。你设定显著性水平为0.05，而且如果H_0是错误的，你想有90%的信心拒绝H_0，那么研究需要多少观测呢？下面的代码给出了答案：

```
> pwr.r.test(r=.25, sig.level=.05, power=.90, alternative="greater")

     approximate correlation power calculation (arctangh transformation)

              n = 134
              r = 0.25
      sig.level = 0.05
          power = 0.9
    alternative = greater
```

因此，要满足以上要求，你需要134个受试者来评价抑郁与孤独的关系，以便在零假设为假的情况下有90%的信心拒绝它。

10.2.4　线性模型

对于线性模型（比如多元回归），pwr.f2.test()函数可以完成相应的功效分析，格式为：

```
pwr.f2.test(u=, v=, f2=, sig.level=, power=)
```

其中，u和v分别是分子自由度和分母自由度，f2是效应值。

$$f^2 = \frac{R^2}{1-R^2}$$ 其中　R^2 = 多重相关性的总体平方值

$$f^2 = \frac{R_{AB}^2 - R_A^2}{1-R_{AB}^2}$$ 其中　${R_A}^2$ = 集合A中变量对总体方差的解释率[①]
${R_{AB}}^2$ = 集合A和B中变量对总体方差的解释率

当要评价一组预测变量对结果的影响程度时，适宜用第一个公式来计算f2；当要评价一组预测变量对结果的影响超过第二组变量（协变量）多少时，适宜用第二个公式。

现假设你想研究老板的领导风格对员工满意度的影响，是否超过薪水和工作小费对员工满意度的影响。领导风格可用四个变量来评估，薪水和小费与三个变量有关。过去的经验表明，薪水和小费能够解释约30%的员工满意度的方差。而从现实出发，领导风格至少能解释35%的方差。假定显著性水平为0.05，那么在90%的置信度情况下，你需要多少受试者才能得到这样的方差贡献率呢？

此处，sig.level = 0.05，power = 0.90，u = 3（总预测变量数减去集合B中的预测变量数），效应值为f2 = (0.35 − 0.30)/(1 − 0.35) = 0.076 9。将这些信息输入到函数中：

```
> pwr.f2.test(u=3, f2=0.0769, sig.level=0.05, power=0.90)

     Multiple regression power calculation

              u = 3
```

① 也常称作方差贡献率。——译者注