第 4 章　假设检验模拟验证

统计功效（statistical power ）是指，在假设检验中，拒绝原假设后，接受正确的替换假设的概率。我们知道，在假设检验中有α错误和β 错误。α错误是弃真错误，β错误是取伪错误。取伪错误是指，原假设为假，样本观测值没有落在拒绝域中，从而接受原假设的概率，即在原假设为假的情况下接受原假设的概率。由此可知，统计功效等于$1-\beta$。

4.1　参数检验模拟验证

$\bar{X} \sim N(\mu, \frac{1}{n})$

$\sim N(\mu, \frac{1}{n})$

? $\bar{X} \to \mu$　$n \to \infty$

例4.1.1　设 $X_1, X_2, \cdots, X_n$ 独立同分布来自 $N(\mu, 1)$ 分布，检验原假设和备择假设分别为：$H_0 : \mu = 0$ 和 $H_1 : \mu \neq 0$。

对于上述假设检验问题，设样本均值 $\bar{X} = \sum_{i=1}^{n} X_i/n$，统计量为 $T_z = \sqrt{n}\bar{X}$。在原假设成立的情况下，T_z 服从 $N(0, 1)$ 分布。在备择假设成立的情况下，T_z 服从 $N(\sqrt{n}\mu, 1)$ 分布。取 $u_{1-\frac{\alpha}{2}}$ 为标准正态分布的 $1 - \alpha/2$ 分位数，当 $|T_z| > u_{1-\frac{\alpha}{2}}$ 时，我们拒绝 H_0。如果 $\mu = n^r C$，当 $r = -1/2$ 时，T_z 服从 $N(C, 1)$ 分布；当 $r > -1/2$ 时，T_z 趋于无穷；当 $r < -1/2$ 时，T_z 趋于 $N(0, 1)$。

$\sqrt{n}\bar{X}$ $N(\mu, 1)$

$P(拒 \mid H_0) = \alpha$　Size

检验的功效为：$p = P(拒绝H_0 \mid H_1成立)$。对于上述假设检验，注意到

$T_z \Leftrightarrow T_z^2$?

μ　$\{\hat{\mu}_1, \hat{\mu}_2\}$?

$$
\begin{aligned}
& P\{|\sqrt{n}\bar{X}| \leq u_{1-\frac{\alpha}{2}} \mid H_1成立\} \\
= \; & P\{-u_{1-\frac{\alpha}{2}} \leq \sqrt{n}(\bar{X} - \mu) + \sqrt{n}\mu \leq u_{1-\frac{\alpha}{2}} \mid H_1成立\} \\
= \; & P\{-u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu \leq \sqrt{n}(\bar{X} - \mu) \leq u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu \mid H_1成立\} \\
= \; & \Phi(u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu) - \Phi(-u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu) \\
= \; & \Phi(u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu) + \Phi(u_{1-\frac{\alpha}{2}} + \sqrt{n}\mu) - 1,
\end{aligned}
$$

模

因此，在本例中功效为

$$p = P\{|\sqrt{n}\bar{X}| > u_{1-\frac{\alpha}{2}} \mid H_1\text{成立}\} = 2 - \Phi(u_{1-\frac{\alpha}{2}} - \sqrt{n}\mu) - \Phi(u_{1-\frac{\alpha}{2}} + \sqrt{n}\mu).$$

但是，对于大部分的假设检验问题，统计量的功效没有显示表达式。 此时，我们需要估计统计量的功效：$\hat{p} = \hat{P}(\text{拒绝}H_0 \mid H_1\text{成立})$.

问题：对于上述例子，如果 p 没有显示表达式，如何估计 $p = P\{|\sqrt{n}\bar{X}| > u_{1-\frac{\alpha}{2}}\}$？

可以用备择假设成立时，拒绝原假设的发生频率来估计功效 p。在本例中，我们从 $N(\mu,1)(\mu \neq 0)$ 中产生 N 组数据，其观测值为 $\{x_{k1},\cdots,x_{kn}\}, k = 1,\cdots,N$，对于第 k 组数据，设样本均值为 $\bar{x}_k = \sum_{i=1}^{n} x_{ki}$。 则统计量 T_z 的观测值为$\{\sqrt{n}\bar{x}_k, k = 1,\cdots,N\}$，取 $I(\cdot)$ 为示性函数，功效 p 的估计为

$$\hat{p} = \frac{1}{N}\sum_{k=1}^{N} I\left(|\sqrt{n}\bar{x}_k| > u_{1-\frac{\alpha}{2}}\right).$$

$\frac{p(1-p)}{N}$?

问题：估计 $\hat{p}$ 的性质，均值和方差？

由于 $N\hat{p}$ 服从二项分布 $b(N,p)$，可得

$$Var(N\hat{p}) = Np(1-p), \quad Var(\hat{p}) = \frac{p(1-p)}{N}.$$

若 $N = 1000$，$\alpha = 0.05$，则

$$p = 2 - \Phi(1.96 - \sqrt{n}\mu) - \Phi(1.96 + \sqrt{n}\mu), \quad Var(\hat{p}) = \frac{p(1-p)}{1000}.$$

在本例中，取样本量 $n = 100$，$\mu = 0.3$，重复次数 $N = 1000$。估计功效的程序如下：

```
n=100
res<-c()
mu=0.3;
  for (i in 1:1000) {
    data<-rnorm(n)
    E_data<-mean(data)+mu
    stat=E_data*sqrt(n)
```

```
    res[i]<-as.numeric(abs(stat)>=qnorm(0.975,0,1))
  }
  result=mean(res)
  criti=qnorm(0.975,0,1);
  power=2-pnorm(criti-sqrt(n)*mu,0,1)-pnorm(criti+sqrt(n)*mu,0,1)
  c(result, power)
```

例4.1.2 设 $X_1, X_2, \cdots, X_n$ 和 $Y_1, Y_2, \cdots, Y_m$ 是来自 $N(0, \sigma_1^2)$ 和 $N(0, \sigma_2^2)$ 的两个独立样本，检验原假设和备择假设分别为：$H_0: \sigma_1 = \sigma_2$ 和 $H_1: \sigma_1 \neq \sigma_2$。

$\hat\sigma_1^2 - \hat\sigma_2^2 = 0$?

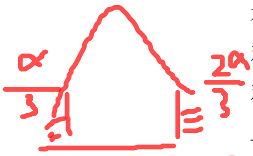
$\dfrac{\hat\sigma_1^2}{\hat\sigma_2^2}$

对于上述假设检验问题，设样本方差

$$S_1^2 = \sum_{i=1}^{n}(X_i - \cancel{\bar X})^2/(n-1), \quad S_2^2 = \sum_{i=1}^{m}(Y_i - \cancel{\bar Y})^2/(m-1),$$

$\sim \chi^2(n-1)$

统计量为 $T_s = S_1^2/S_2^2$。在原假设成立的情况下，T_s 服从 $F(n-1, m-1)$ 分布，备择假设下，$\sigma_2^2 T_s/\sigma_1^2$ 服从 $F(n-1, m-1)$ 分布。取 $F_{1-\frac{\alpha}{2}}(n-1, m-1)$ 和 $F_{\frac{\alpha}{2}}(n-1, m-1)$ 分别是自由度为 $n-1$ 和 $m-1$ 的 F 分布的 $1-\alpha/2$ 和 $\alpha/2$ 分位数，则该假设检验的拒绝域为 $T_s > F_{1-\frac{\alpha}{2}}(n-1, m-1)$ 或 $T_s < F_{\frac{\alpha}{2}}(n-1, m-1)$。本例中，



$$\begin{aligned}
& P(\text{接受}H_0 \mid H_1\text{成立}) \\
= \; & P\{F_{\frac{\alpha}{2}}(n-1, m-1) \le T_s \le F_{1-\frac{\alpha}{2}}(n-1, m-1) \mid H_1\text{成立}\} \\
= \; & P\{\sigma_2^2 F_{\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2 \le \sigma_2^2 T_s/\sigma_1^2 \le \sigma_2^2 F_{1-\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2 \mid H_1\text{成立}\} \\
= \; & F_{(n-1,m-1)}\{\sigma_2^2 F_{1-\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2\} - F_{(n-1,m-1)}\{\sigma_2^2 F_{\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2\},
\end{aligned}$$

其中，自由度为 $n-1$ 和 $m-1$ 的 F 分布的累积分布函数记为 $F_{(n-1,m-1)}$。因此，检验统计量 T_s 的功效为

$$p = 1 - F_{(n-1,m-1)}\{\sigma_2^2 F_{1-\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2\} + F_{(n-1,m-1)}\{\sigma_2^2 F_{\frac{\alpha}{2}}(n-1, m-1)/\sigma_1^2\}.$$

当然我们也可以采用重复试验的方法对其进行估计。例如，我们从 X 和 Y 的总体分布 $N(0, \sigma_1^2)$ 和 $N(0, \sigma_2^2)$ $(\sigma_1 \neq \sigma_2)$ 中产生 N 组数据，对于第 k 组样本，

其观测值为 $\{x_{k1},\cdots,x_{kn};y_{k1},\cdots,y_{km}\},k=1,\cdots,N$，且样本均值为

$$\bar{x}_k=\sum_{i=1}^{n}x_{ki},\quad \bar{y}_k=\sum_{i=1}^{m}y_{ki},$$

样本方差为

$$s_{1(k)}^2=\sum_{i=1}^{n}(x_{ki}-\bar{x}_k)^2/(n-1),\quad s_{2(k)}^2=\sum_{i=1}^{m}(y_{ki}-\bar{y}_k)^2/(m-1).$$

为便于书写，我们将统计量 T_s 的观测值记为 $\{T_s^{(k)},k=1,\cdots,N\}$，其中 $T_s^{(k)}=s_{1(k)}^2/s_{2(k)}^2$，功效 p 的估计为

$$\hat{p}=\frac{1}{N}\sum_{k=1}^{N}I\left\{T_s^{(k)}>F_{1-\frac{\alpha}{2}}(n-1,m-1)\text{或}T_s^{(k)}<F_{\frac{\alpha}{2}}(n-1,m-1)\right\}.$$

同上例一样，在 N 和 α 给定时，我们也可以研究估计 $\hat{p}$ 的均值和方差。取样本量为 $n=100$ 和 $m=50$，总体方差分别为 $\sigma_1^2=1$ 和 $\sigma_2^2=4$，重复次数 $N=1000$。估计功效的程序如下：

```
n=100;m=50
res<-c()
sigma=2
for (i in 1:1000) {
  X_data<-rnorm(n)
  Y_data<-sigma*rnorm(m)
  S_x=sum((X_data-mean(X_data))^2)/(n-1)
  S_y=sum((Y_data-mean(Y_data))^2)/(m-1)
  stat=S_x/S_y
  criti_1=qf(0.975,n-1,m-1)
  criti_2=qf(0.025,n-1,m-1)
  res[i]<-as.numeric(stat>=criti_1)+as.numeric(stat<=criti_2)
}
result=mean(res)
power=1-pf(sigma^2*criti_1,n-1,m-1)+pf(sigma^2*criti_2,n-1,m-1)
c(result, power)
```