第八章　　假 设 检 验

假设检验是统计推断中的另一类重要问题，是凭借样本信息对总体的某种假设进行判断，从而决定是接受或拒绝这一假设的一个决策过程。.

§8.1　假设检验的基本思想和概念

（一）　假设检验问题

先看一个具体实例　.

设某厂生产一种灯泡，其寿命 X 服从$N(\mu,\sigma^2)$．若已知σ^2=4000，按该厂的经验，灯泡的平均寿命为μ=1500 小时，现采用新工艺，为探求新工艺是否改变了灯泡的寿命，随机抽取 25 只作试验，测得平均寿命为 1675 小时，问采用新工艺后，灯泡的寿命是否有明显的变化？

我们的问题就是要判别新产品的寿命是服从 $\mu>1500$ 的正态分布呢？还是与老产品一样仍然服从$\mu=1500$的正态分布？如是前者，我们就说新产品的寿命有显著提高；如是后者，就说没有显著提高.

在上面的例子中，我们可以把所涉及到的两种情况用统计假设的形式表示出来. 第一个统计假设 $\mu=1500$ 表示采用新工艺后产品平均寿命没　有显著增加　，第二个统计假设 $\mu>1500$ 表示采用新工艺后产品平均寿命有显著增加. 我们称第一个假设$\mu=1500$为原假设，用符号$H_0:\mu=1500$表示；第二个假设$\mu>1500$为备择假设，用符号$H_1:\mu>1500$表示．至于在两个假设中用哪个作为原假设，哪个作为备择假设，要看具体的目的和要求而定．

假设检验问题的一般提法是：在给定的备择假设H_1下对原假设H_0做出判断，若拒绝原假设H_0就意味着接受备择假设H_1，否则就接受原假设H_0．简单地说，假设检验问题就是在原假设H_0和备择假设H_1中做出表态：即接受原假设H_0或者接受备择假设H_1，这类假设检验问题常常简称为H_0对H_1的检验问题.

在H_0对H_1的检验问题中要做出某种判断，必须从样本$X_1,X_2,\cdots,X_n$出发，制定一个法则，一旦样本的观察值$(x_1,x_2,\cdots,x_n)$确定后，利用我们所构造的法则作出判断：拒绝

H_0 还是拒绝 H_1．这种法则就成为 H_0 对 H_1 的一个检验法则，有时也简称一个检验.

我们所构造的检验法则，是以定义在样本空间上的一个函数为依据所构成的一个准则．这个法则本质上是将样本空间划分成两个互不相交的子集 C 和 C^*，使得样本观察值 $(x_1,x_2,\cdots,x_n)\in C$ 时，将拒绝原假设 H_0（同时接受备择假设 H_1）；当样本观察值 $(x_1,x_2,\cdots,x_n)\in C^*$ 时，将接受原假设 H_0（同时拒绝备择假设 H_1），这样划分构成的一个准则，我们称样本空间的子集 C 为检验的拒绝域（或叫临界域）.

（二）假设检验的基本思想和方法

假设检验的基本思想是依据实际推断原理，即“概率很小的事件在一次试验中实际上几乎是不会发生的”．如果概率很小的事件在一次试验中竟然发生了，这就使我们怀疑假定的正确性，从而拒绝所提出的假设，否则接受假设 .

回到本文的例子．对假设 $H_0:\mu=1500$ 和假设 $H_1:\mu>1500$，若 H_0 为真（成立），则在新工艺下生产的灯泡，其寿命均应当服从正态分布 $N(1500,4000)$．由于样本均值 $\overline{X}$ 是总体均值 μ 的无偏估计，因此，当 H_0 为真时，$\left|\overline{X}-1500\right|$ 不应该过大，应确定个临界值 k，使得当 $\left|\overline{X}-1500\right|>k$，则否定假设“$H_0$ 为真”，即拒绝 H_0．那么如何来确定临界值 k 呢？

若给定一个 $\alpha(0<\alpha<1)$，α 一般比较小，那么我们可以通过

$$P\left\{\left|\overline{X}-1500\right|>k\right\}=\alpha$$

来确定出 k．因为对于抽取的样本，如果 $\left|\overline{X}-1500\right|>k$ 成立，而 $\left|\overline{X}-1500\right|>k$ 发生的概率比较小（为 α），则意味着小概率事件发生了，按照实际推断原理，有理由认为原来的假设（即假设 H_0 为真）是错误的，由此拒绝原假设.

进一步我们来探讨如何确定 k.

由于 $X\sim N(\mu,{\sigma_0}^2)$，在 σ_0 已知时，统计量 $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$，假设 H_0 为真，则对于给定的 α，有

$$P\left\{\left|\frac{\overline{X}-1500}{\sigma_0/\sqrt{n}}\right|>Z_{\alpha/2}\right\}=\alpha$$

其中$Z_{\alpha/2}$是正态分布的双侧α分位点

上式变形有

$$P\left\{\left|\overline{X}-1500\right|>Z_{\alpha/2}\sigma_0/\sqrt{n}\right\}=\alpha$$

由此确定$k=Z_{\alpha/2}\sigma_0/\sqrt{n}$.

在本文的例子中，$\left|\overline{X}-1500\right|=175$，若$\alpha=0.05$，则$k=1.96\times200/\sqrt{25}=78.4$. 因为$175>78.4$，故拒绝$H_0$，同时接受$H_1$，即认为新工艺使得灯泡寿命发生了显著变化.

我们称这种假设检验为显著性检验，把α称为检验的显著性水平.

（三）关于两类错误

假设检验问题的基本思想是依据实际推断原理，但是，在进行检验过程中，由于样本的随机性，我们可能会犯两类错误.

第一类错误（弃真错误）:即本来H_0为真，但依据样本的信息错误地拒绝了H_0. 设这类错误的概率为α，即

$$P\{拒绝H_0 \mid H_0为真\}=\alpha$$

第二类错误（采伪错误）,即本来H_0不真，但依据样本的信息错误地接受了H_0. 设犯这类错误的概率为β，即

$$P\{接受H_0 \mid H_0不真\}=\beta$$

在样本容量一定的情况下，β与α的数值呈反向变化.

为明确起见，我们将两类错误列于下表

表 8.1　两 类 错 误

样本	总体情况	
$(x_1,x_2,\cdots,x_n)$	H_0为真	H_0不真
$\in C$（拒绝域）	犯第一类错误	正确
$\in C^*$（接受域）	正确	犯第二类错误

对于给定的一对假设H_0和H_1，总可以找出很多拒绝域，我们总是希望寻得这种拒绝

域，使犯两类错误的概率α和β都很小．但是，在样本容量固定时，要使得α和β同时减少是不可能的，否则将会导致样本容量的无限增大，这又是不实际的．基于这种情况，奈曼和皮尔逊（Neyman-Pearson）提出一个原则，即在控制犯第一类错误的概率α的条件下，尽量使犯第二类错误的概率 β小，因为人们常常把拒绝 H_0比错误地接受 H_0看得更重要些．在实际操作中，我们将上述原则简化，只对α加以限制，而不考虑β，这时，寻找拒绝域时就只涉及原假设H_0，而不涉及备择假设H_1．按照这种原则的假设检验即是上述的显著性检验，显著性水平α即犯第一类错误的概率.

（四）显著性检验的具体步骤

综上所述，显著性检验的具体步骤如下

1．依据具体问题，提出假设H_0和备择假设H_1；

2．选取适当的统计量，当H_0为真时，它的分布应该是已知的;

3．给定显著性水平α（一般取比较小的数）,确定拒绝域;

4．依据样本观察值$(x_1, x_2, \cdots, x_n)$，计算统计量的值;

5．作出判断：若统计量的值落入拒绝域，则拒绝H_0，否则就接受H_0.

§8.2　正态总体的参数假设检验

正态总体是最常见、也是最重要的总体。正态分布$N(\mu,\sigma^2)$含有两个参数μ和σ^2，因此下面的假设都是针对这两个参数的假设。

（一）单个总体$N(\mu,\sigma^2)$

1° σ^2已知，关于μ的检验（U 检验）

需检验假设$H_0: \mu = \mu_0$； $H_1: \mu \neq \mu_0$。

假设H_0为真，则统计量

$$U = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1) \qquad (8.1)$$

对于给定的显著性水平α，有

$$P\{|U| \geq z_{\alpha/2}\} = \alpha$$

由此，我们便得到拒绝域 $C=\{|U|\geq z_{\alpha/2}\}$。从样本观察值算出 （8.1）式中$U$ 的值，若$|U|\geq z_{\alpha/2}$，则拒绝原假设H_0，即认为总体均值与μ_0有显著差异；若$|U|<z_{\alpha/2}$ ，则接受原假设H_0，即认为总体均值与μ_0无显著差异。称这种检验法为U 检验法。

2° σ^2未知，关于μ的检验（t检验）

需检验假设$H_0:\mu=\mu_0$； $H_1:\mu\neq\mu_0$。

假设H_0为真，则统计量

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1) \tag{8.2}$$

对于给定的显著性水平α，有

$$P\{|t|\geq t_{\alpha/2}(n-1)\}=\alpha$$

由此，我们便得到拒绝域$C=\{|t|\geq t_{\alpha/2}(n-1)\}$。从样本观察值算出（8.2）式中$t$的值，若$|t|\geq t_{\alpha/2}(n-1)$，则拒绝原假设$H_0$，即认为总体均值与$\mu_0$有显著差异；若$|t|<t_{\alpha/2}(n-1)$ ，则接受原假设H_0，即认为总体均值与μ_0无显著差异。称这种检验法为t检验法。

下面举一个例子。

例 1. 设某次考试的考生成绩服从正态分布，从中随机抽取 36 位考生的成绩，算得平均成绩为 66.5 分，标准差为 15 分，问在显著性水平$\alpha=0.05$下，是否可以认为这次考试全体考生的平均成绩为 70 分？

解：设该次考试的学生成绩为X，则$X\sim N(\mu,\sigma^2)$。提出假设：

$$H_0:\mu=70，\ H_1：\ \mu\neq 70。$$

由于σ^2未知，用t检验法，选用统计量

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$

在显著性水平$\alpha=0.05$下，拒绝域为

$$(-\infty,-t_{\alpha/2}(35))\cup(t_{\alpha/2}(35),+\infty),$$

查t分布表得 $t_{\alpha/2}(35)=2.0301$，即拒绝域为

$$(-\infty,-2.0301))\cup(2.0301,+\infty),$$

计算统计量样本观测值，可得$t=\dfrac{66.5-70}{15/\sqrt{36}}=-1.4$。因为$t=-1.4$未落入拒绝域，从而接受$H_0$，即可以认为这次考试全体考生的平均成绩为70分.

3° μ未知，关于σ^2的检验（χ^2检验）

需检验假设$H_0:\sigma^2=\sigma_0{}^2$，$H_1:\sigma^2\neq\sigma_0{}^2$。

假设H_0为真，则统计量

$$\chi^2=\frac{(n-1)S^2}{\sigma_0{}^2}\sim\chi^2(n-1) \qquad (8.3)$$

对于给定的显著性水平α，有

$$P\{\chi^2>\chi^2{}_{\alpha/2}(n-1)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n-1)\}=\alpha$$

由此，我们便得到拒绝域$C=\left\{\chi^2>\chi^2{}_{\alpha/2}(n-1)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n-1)\right\}$。从样本观察值算出（8.3）式中$\chi^2$的值，若$\chi^2>\chi^2{}_{\alpha/2}(n-1)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n-1)$，则拒绝原假设$H_0$，即认为总体方差与$\sigma_0^2$有显著差异；若$\chi^2{}_{\alpha/2}(n-1)\geq\chi^2\geq\chi^2{}_{1-\alpha/2}(n-1)$，则接受原假设$H_0$，即认为总体方差与$\sigma_0^2$无显著差异。称这种检验法为$\chi^2$检验法。

4° μ已知，关于σ^2的检验（χ^2检验）

需检验假设$H_0:\sigma^2=\sigma_0{}^2$，$H_1:\sigma^2\neq\sigma_0{}^2$。

假设H_0为真，则统计量

$$\chi^2=\frac{\sum_{i}^{n}(X_i-\mu)^2}{\sigma_0{}^2}\sim\chi^2(n) \qquad (8.4)$$

对于给定的显著性水平α，有

$$P\{\chi^2>\chi^2{}_{\alpha/2}(n)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n)\}=\alpha$$

由此，我们便得到拒绝域$C=\left\{\chi^2>\chi^2{}_{\alpha/2}(n)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n)\right\}$。从样本观察值算出（8.4）式中 χ^2 的值，若$\chi^2>\chi^2{}_{\alpha/2}(n)\text{或}\chi^2<\chi^2{}_{1-\alpha/2}(n)$，则拒绝原假设 H_0，；若

$\chi^2{}_{\alpha/2}(n) \ge \chi^2 \ge \chi^2{}_{1-\alpha/2}(n)$ ，则接受原假设H_0。这种检验法也称为χ^2检验法。

例 2 某电子元件的寿命（单位：小时）$X \sim N(\mu,\sigma^2)$，其中μ,σ^2未知，现测得 16 只元件的寿命如下：

170	264	485	179	260	279	149	224
222	159	250	212	168	101	362	280

问元件寿命的方差是否等于100^2？（$\alpha = 0.05$)

解：这是一个均值未知，检验方差的问题，用χ^2-检验法。

提出统计假设 $H_0:\sigma^2 = 100^2$，$H_1:\sigma^2 \ne 100^2$。选用统计量$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2}$，当$H_0:\sigma^2 = 100^2$成立时，$\chi^2$服从自由度为 15 的$\chi^2$-分布.

在 $\alpha = 0.05$ ， 拒绝域为 $(0,\chi^2{}_{1-\alpha/2}(n-1)] \cup [\chi^2{}_{\alpha/2}(n-1),+\infty)$ ， 查表得$\chi^2{}_{0.975}(15) = 6.262$，$\chi^2{}_{0.025}(15) = 27.488$，即拒绝域为$(0,6.262) \cup (27.488,+\infty)$

计算统计量样本观测值,因为$\bar{x} = 235.25$，$s = 8538.467$，则$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} = 12.81$。

因为χ^2的值未落入拒绝域,从而接受H_0，即可以认为元件寿命的方差与100^2无显著差异.

（二）两个总体$N(\mu_1,\sigma_1^2)$，$N(\mu_1,\sigma_1^2)$

如果要比较两个正态总体的方差是否相等,就要用到下面的F-检验。设$X_1, X_2, \cdots, X_n$是来自总体$N(\mu_1,\sigma_1^2)$的样本，$Y_1,Y_2,\cdots,Y_n$是来自总体$N(\mu_1,\sigma_1^2)$的样本，样本方差分别为S_1^2, S_2^2,且两样本相互独立。若μ_1，μ_2，σ_1^2，σ_2^2均为未知，显著性水平为α。现在需要检验假设

$$H_0:\sigma_1^2 = \sigma_2^2,\quad H_1:\sigma_1^2 \ne \sigma_2^2。$$

在原假设$H_0:\sigma_1^2 = \sigma_2^2$为真时，由第六章定理 6.2 知道，统计量

$$F = \frac{S_n^2}{S_m^2} \sim F(n-1,m-1) \tag{8.5}$$

因此，可由

$$P\left\{F \geq F_{\alpha/2}(n-1,m-1) \text{或} F \leq F_{1-\alpha/2}(n-1,m-1)\right\} = \alpha$$

得到拒绝域 $C = \left\{F \geq F_{\alpha/2}(n-1,m-1) \text{或} F \leq F_{1-\alpha/2}(n-1,m-1)\right\}$。从样本观察值算出（8.5）式中 F 的值，若 $F \geq F_{\alpha/2}(n-1,m-1) \text{或} F \leq F_{1-\alpha/2}(n-1,m-1)$，则拒绝原假设 H_0，；若 $F_{\alpha/2}(n-1,m-1) \geq F \geq F_{1-\alpha/2}(n-1,m-1)$，则接受原假设 H_0。这种检验法也称为 F 检验法。