第 7 章　　假设检验

在科研、生产和日常生活中，我们常常要对很多问题做出一定的论断或猜测，这就是假设，而假设需要做出是或非的回答。为此，我们需要安排一些试验，试验的结果与我们感兴趣的问题有着某种关系，我们可以根据试验的结果对所作出的论断制或猜测制定判断规则并做出是或非的回答。以上的过程我们称之为假设检验。假设检验为统计推断的重要内容之一。本章在建立假设检验的有关概念之后，重点介绍正态总体的参数的检验方法。

7．1 假设检验的基本概念

1．统计假设

设总体 X 的分布函数为 $F(x)$，$F(x)$ 一般完全或部分未知，对未知的总体分布所作假设称为一个统计假设。当总体分布的类型已知，对分布的一个或几个未知参数的值作出假设，或者对总体分布函数的类型或某些特征提出某种假设，这种假设称为待检假设或零假设，通常用 H_0 表示。事实上，当我们提出了零假设时，也同时给出了另外一个假设，即提供给我们选择的备择假设，记为 H_1。H_0 与 H_1 是互不相容的。

在参数模型下，如果总体的分布类型已知，仅是某个参数未知，只要对未知参数作出假设就可确定总体的分布。这种仅涉及到总体分布的参数的统计假设称为参数假设。若是对总体的分布类型或某些特征提出假设，这称为非参数假设。

例 1　手表厂生产的女表表壳，在正常情况下，其直径（单位：mm）服从正态分布 $N(20，1)$。为了检查该厂某天生产是否正常，对生产过程中的手表表壳随机的抽查了 5 只，测的表面直径分别为 19，19.5,19,20,20.5。问这天生产是否正常?

由问题的提出可知，我们实际上是要检查这天生产的手表表壳的直径 μ 是否为 20？即提出假设 H_0：$\mu_0 = 20$ 及备择假设 H_1：$\mu_0 \neq 20$。这样，问题就转化为如何利用抽查得到的样本去检验零假设 $\mu_0 = 20$ 的真伪。因此，这就需要设置一种检验的规则以及如何根据规则进行检验作进一步的讨论。

2．检验法则

在确定了待检假设以后，我们必须在 H_0 与 H_1 之间作出抉择，而对一个假设的确定只有接受和拒绝两种，例如在例 1 中，如果我们接受 H_0，即表示该厂这天的生产是正常的，如果拒绝 H_0，亦即接受 H_1，则表示该天生产不正常。为此，必须设计一种合理的法则，根据这一法则，就可利用已得到的样本作出判断。在例 1 中，由于要检验的假设涉及总体均值 μ，故容易想到可否借助样本均值 $\bar{x}$ 这一统计量来进行判断。这是因

为$\bar{x}$是 μ 的无偏估计，样本均值$\bar{x}$的大小在一定的程度上反映了μ 的大小。因此，当假设H_0为真时，则观测值$\bar{x}$与$\mu_0=20$的偏差$|\bar{x}-20|$一般不应太大。如果$|\bar{x}-20|$过分大，我们就应怀疑假设 H_0的正确性并拒绝 H_0。而衡量$|\bar{x}-\mu_0|$的大小归结为衡量统计量

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}$$

的大小，在H_0为真时统计量 $\dfrac{\bar{x}-\mu_0}{\sigma\sqrt{n}}\sim N(0,1)$

基于上面的设想，我们可适当限定一正数k，使得当$\bar{x}$满足不等式

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}\geq k$$

时就拒绝H_0。反之，若 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}<k$

时则接受H_0。正数k的每一个选择都对应着一个不同的检验法则。

在给定的一个检验法则中，以$\aleph_0$表示在此检验法中引起拒绝H_0的所有可能的样本观察值的集合，并称$\aleph_0$为此检验法的**拒绝域**，而它的余集称为**接受域**。显然，检验法与拒绝域是一一对应的。

3．两类错误

我们做出判断的依据是一个样本，由于样本的随机性，我们进行假设检验时不可避免地出现误判而犯错误，当H_0为真时，仍可能做出拒绝H_0的判断，这类错误称为犯第Ⅰ类错误，也称为“**弃真**”或“**拒真”；**也可能在H_0不真时，却接受H_0，称为犯第Ⅱ类错误，也称为“**取伪**”或“**受伪”。**犯第一类错误的概率为

$$P\{拒绝H_0|为真\}$$

由于在实际中无法排除犯这类错误的可能性，因此，我们自然希望犯第Ⅰ类错误的概率控制在一定的限度之内。例如可给定一个较小的正数α（$0<\alpha<1$），并使

$$P\{拒绝H_0|H_0为真\}\leq\alpha$$

α 一般称为检验水平。下面我们将作进一步的讨论。

4．水平为α的检验

犯两类错误的大小自然就决定着相应的检验法则的优劣，但在样本容量固定的条件

下可以证明犯两类错误的概率不可能同时达到很小。通常 H_0 是比较重要的假设，因此如何犯第 I 类错误的概率控制在小概率的范围内就显得非常重要。其做法如下：给定 α（$0<\alpha<1$），构造一个检验的拒绝域，使其犯第 I 类错误的概率不超过 α，即 $P\{$拒绝 $H_0 | H_0$ 为真$\} \le \alpha$。例如在例 1 中，为了检验 H_0：$\mu_0 = 20$，我们构造了一个统计量

$$\frac{\left|\bar{x}-\mu_0\right|}{\sigma/\sqrt{n}}=\frac{\left|\bar{x}-20\right|}{1/\sqrt{5}} \sim N(0,1)$$

如果给定 $\alpha=0.05$，并使犯第 I 类错误的概率最大为 α，由此可构造一个拒绝域为:

$$\frac{\left|\bar{x}-\mu_0\right|}{1/\sqrt{5}} \ge z_{\frac{\alpha}{2}}$$

并使 $$P\left\{\frac{\left|\bar{x}-\mu_0\right|}{1/\sqrt{5}} \ge z_{\frac{\alpha}{2}}\right\}=\alpha$$

这里 $z_{\alpha/2}$ 可由标准正态分布表查得 $z_{\frac{\alpha}{2}}=z_{0.025}=1.96$

这时可得拒绝域为 $\aleph_0:\dfrac{\left|\bar{x}-20\right|}{1/\sqrt{5}} \ge 1.96$,即$\left|\bar{x}-20\right| \ge 1.96\sqrt{5}$

这种把犯第 I 类错误的概率控制在不超过给定的 α 的检验法称为显著性水平为 α 的检验，并称 α 为**显著性水平，**或简称为水平。

通过上面的例子，我们给出构造检验的一般方法。

例 2 设 X_1，X_2，…，X_n 为来自总体 X 的一个样本，$X \sim N(\mu, \sigma_0^2)$，其中 σ_0^2 已知，μ 未知。给定显著性水平为 α（$0<\alpha<1$），试构造检验假设为

$$H_0:\mu=\mu_0\,,\ H_1:\mu\ne\mu_0(\mu_0\text{为已知})$$

的水平为 α 的检验

解 考虑 μ 的无偏估计 $\overline{X}$，且知道 $\overline{X} \sim N(\mu,\dfrac{\sigma_0^2}{n})$.

当 H_0 为真时, $$\overline{X} \sim N(\mu_0,\frac{\sigma_0^2}{n}).$$

于是 $\overline{X}$ 的值应落在 μ_0 的附近。所以当 H_0 为真时，$|\overline{X}-\mu_0|$ 取较大值应为小概率事件。由此选择 H_0 的拒绝域为 $$\aleph_0=\{(x_1,x_2,\ldots,x_n):|\bar{x}-\mu_0| \ge k\}$$

这里 k 为某待定正常数。

当检验水平为 α 时，k 应满足 $\alpha=P_{H_0}\{(X_1,X_2,\ldots.X_n)\in X_0\}$

$$= P_{\mu=\mu_0}\left\{|\overline{X}-\mu_0| \geq k\right\} = P_{\mu=\mu_0}\left\{\frac{|\overline{X}-\mu_0|}{\sigma_0/\sqrt{n}} \geq \frac{k}{\sigma_0/\sqrt{n}}\right\},$$

故 $\quad \frac{k}{\sigma_0/\sqrt{n}} = z_{\frac{\alpha}{2}}$ 即 $\quad k = z_{\frac{\alpha}{2}}\frac{\sigma_0}{\sqrt{n}}$

因此所取得水平为 α 的拒绝域为 $\aleph_0 = \{(x_1,x_2,...,x_n):\frac{|\overline{x}-\mu_0|}{\sigma_0/\sqrt{n}} \geq z_{\frac{\alpha}{2}}\}$

由此例可见，统计量 $\overline{X}-\mu_0$ 或 $\frac{|\overline{X}-\mu_0|}{\sigma_0/\sqrt{n}}$

在检验的构造过程中起着关键作用，一般称其为检验统计量。我们要求在 H_0 为真时，检验统计量的分布应是确定的(已知)的，且不含任何未知参数。例如在例 1 中，其检验统计量为

$$\frac{|\overline{X}-\mu_0|}{\sigma_0/\sqrt{n}} \sim N(0,1)$$

它满足上述要求。故在 α =0.05 时，拒绝域为

$$\aleph_0 = \{(x_1,x_2,...,x_n):\frac{|\overline{x}-\mu_0|}{\sigma_0/\sqrt{n}} = \frac{|\overline{x}-20|}{1/\sqrt{5}} \geq z_{\frac{\alpha}{2}} = 1.96\}$$

由样本算得 $\quad \overline{X} = 1.96$

代入检验统计量中可得 $\quad \frac{|\overline{X}-20|}{1/\sqrt{5}} = |19.6-20|\sqrt{5} = 0.8944 < z_{\frac{\alpha}{2}} = 1.96$

这表明样本值落在接受域内，故应接受 H_0，从而认为该天生产的女表表壳可得直径的均值是 20，亦即认为该天的生产是正常的。

例 3 安装一台新仪器要求元件尺寸的均值保持在原有仪器的水平。已知原有仪器的元件尺寸均值为 3.278cm，均方差为 0.002cm。现测量 10 个新元件，得尺寸数据(单位:cm)为

3.277 3.281 3.278 3.278 3.286 3.279 3.278 3.281 3.279 3.280.

设元件尺寸服从正态分布。且新、旧元件尺寸分布的方差相同，问新装仪器的元件尺寸的均值与原有仪器的元件尺寸均值有误显著差别？（取 α =0.05）

解 设元件尺寸 X 服从正态分布 $X \sim N(\mu,\sigma^2)$，因新旧元件尺寸的方差相同，故 σ^2=0.02。由题意知待检假设为

$$H_0:\mu = 3.278, H_1:\mu \neq 3.278$$

由例 2 可知，水平为 α 的拒绝域为 $\aleph_0 = \{(x_1,x_2,...,x_n):\frac{|\overline{x}-\mu_0|}{\sigma_0/\sqrt{n}} \geq z_{\frac{\alpha}{2}}\}$

现在α=0.05，由例 1 知 $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$.

又由样本算得均值为 $\overline{X} = 3.2795$

且 $\mu_0 = 3.278, \sigma^2 = 0.002^2, n = 10,$

从而可算得 $\dfrac{|\overline{X} - \mu_0|}{\sigma_0/\sqrt{n}} = \left|\dfrac{3.2795 - 3.278}{0.02} \cdot \sqrt{10}\right|. = 2.37 > 1.96.$

由于它落在拒绝域$\aleph_0$内，故拒绝 H_0，接受 H_1，即认为新旧元件尺寸的均值之间存在显著差别。

5. 假设检验的程序

上面叙述的检验法则具有普遍意义，可用在各种各样的假设检验问题上。由此我们归结出假设检验的一般步骤：

1. 根据题意合理地建立**零假设** H_0 **和设备假设** H_1；

若零假设为 H_0：$\mu = \mu_0$，则备择假设 H_1 根据实际情况可以有下面三种：

$$H_1: (1)\ \mu \neq \mu_0\ (2)\ \mu < \mu_0\ (3)\ \mu > \mu_0$$

在一般情况下 H_1 常选择（1），这时称为**双侧检验**；若选择（2）或（3）称为**单侧检验**。如所考虑总体的均值越大越好时，H_1 可选择（3）。

2. 选择适当的检验统计量 T;

要求在 H_0 为真时，统计量 T 的分布是确定和已知的；

3. 规定检验水平α，并由 H_0 和 H_1 确定一个合理的拒绝域($\aleph_0$ 含有待定常数）.

4. 样由本观测值，计算出统计量 T_0 的值；

5. 作出判断：若统计量的值 T_0 落在拒绝域内，则拒绝 H_0，否则接受 H_0。

7. 2 正态总体的参数分布

下面我们讨论正态总体参数的假设检验问题，分单个正态总体与两个正态总体的情形来讨论。

1．单个总体均值 μ 的检验

设样本 X_1，X_2，…，X_n来自正态总体 $X \sim N(\mu, \sigma^2)$。

(1) 方差σ^2已知，检验假设 H_0：$\mu = \mu_0$

这时备择假设 H_1 可根据具体问题选择 $H_1: (1)\ \mu \neq \mu_0\ (2)\ \mu < \mu_0\ (3)\ \mu > \mu_0$

中的一种。检验统计量取 $u = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1).$

在给定水平为α时，若备择假设为 H_1：$\mu \neq \mu_0$，则其拒绝域为

$$\aleph_0=\{(x_1,x_2,\ldots,x_n):\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}\geq z_{\frac{\alpha}{2}}\}.$$

若 H_1 为 $\mu<\mu_0$, 则其拒绝域为若 $\aleph_0=\{(x_1,x_2,\ldots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\leq z_\alpha\}$

若 H_1 为 $\mu>\mu_0$，则拒绝域为 $\aleph_0=\{(x_1,x_2,\ldots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\geq z_\alpha\}$

由于在以上检验中，我们取统计量为 $u=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1).$

故称为 U 检验法。

有时, 我们还可能遇上更为完整的假设，或者说由具体问题提出以下的假设更为合理：

$$H_0:\mu\leq\mu_0,\ H_1:\mu>\mu_0,$$

在这里 H_0 包含了多种情形，称为复合假设。可以证明，这时在水平 α 下的拒绝域为

$$\aleph_0=\{(x_1,x_2,\ldots,x_n):\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\geq z_\alpha\}$$

(2) 方差 σ^2 未知，检验假设 H_0：$\mu=\mu_0$

这时由于 σ^2 未知，$\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$

已不能作为检验统计量，由于样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$

是方差 σ^2 的无偏估计，故以 S 代替 σ 可得 T 统计量 $T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$,

在 H_0 为真时，统计量 $T\sim t(n-1)$ 在给定 α 时，若 H_1 为 $\mu\neq\mu_0$，则其拒绝域为

$$\aleph_0=\{(x_1,x_2,\ldots,x_n):|\frac{\bar{x}-\mu_0}{S/\sqrt{n}}|\geq t_{\frac{\alpha}{2}}(n-1)\}$$

这种检验法称为 **T 检验法。**

类似地可以写出 T 检验法的单边假设检验结果见文字教材的附表 8.1。

例 1 用某仪器间接测量温度, 重复五次, 测得结果为

$$1250^\circ\quad 1265^\circ\quad 1245^\circ\quad 1260^\circ\quad 1275^\circ.$$

设测量值 X 服从正态分布 $N(\mu,\sigma^2)$，水平 $\alpha=0.05$。问是否有理由认为该仪器测量值大于 1277^0（真实值)？

解 由题意知，H_0：$\mu \le \mu_0$。因方差σ^2未知，用 T检验法，这时拒绝域为

$$\aleph_0 = \{(x_1, x_2, ..., x_n): \frac{\bar{x}-\mu_0}{S/\sqrt{n}} \ge t_\alpha(n-1)\}$$

这里 $n=5$，$t_\alpha(n\text{-}1) = t_{0.05}(4) = 2.1348$ 由样本算得 $\overline{X} = 1259$ ，$S^2 = 142.5$

代入可得 $$T_0 = \frac{1259-1277}{\sqrt{\frac{142.5}{5}}} \approx -3.372 < 2.1348.$$

因此接受 H_0即认为测量值不大于 1277^0。

2. 单个总体的方差σ^2的检验

设样本 X_1，X_2，...，X_n 来自正态总体 $X \sim N(\mu, \sigma^2)$，均值μ未知，检验假设 H_0: $\sigma^2=\sigma_0^2$，H_1：$\sigma^2 \ne \sigma_0^2$，（σ_0^2为已知常数）。

由前可知，样本方差 $$S^2 = \frac{1}{n-1}\sum_{n-1}^{n}(X_i - \overline{X})^2$$

是总体方差σ^2的无偏估计，故当 H_0为真时，样本方差 S^2的值应在σ_0^2的附近。这时我们取检验统计量为 $$X^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1).$$

其拒绝域应有以下形式：$\aleph_0 = \{(x_1, x_2, ..., x_n): \frac{(n-1)S^2}{\sigma_0^2} \le k_1 或 \frac{(n-1)S^2}{\sigma_0^2} \ge k_2\}$

此处 k_1，k_2由下式确定：

$$P\{拒绝H_0 \mid H_0真\} = P_{\sigma_0^2}\left\{\left(\frac{(n-1)S^2}{\sigma_0^2} \le k_1\right)\right\} \cup \left\{\left(\frac{(n-1)S^2}{\sigma_0^2} \ge k_2\right)\right\} = \alpha,$$

为了计算方便起见，习惯上取

$$P_{\sigma_0^2} = \left\{\left(\frac{(n-1)S^2}{\sigma_0^2} \le k_1\right)\right\} = \frac{\alpha}{2}, \qquad P_{\sigma_0^2} = \left\{\left(\frac{(n-1)S^2}{\sigma_0^2} \ge k_2\right)\right\} = \frac{\alpha}{2},$$

故得 $$k_1 = \chi^2_{1-\frac{\alpha}{2}}(n-1), \quad k_2 = \chi^2_{\frac{\alpha}{2}}(n-1)$$

因此拒绝域为$\aleph_0 = \{(x_1, x_2, ..., x_n): \frac{(n-1)s^2}{\sigma_0^2} \le \chi^2_{1-\frac{\alpha}{2}}(n-1) 或 \frac{(n-1)s^2}{\sigma_0^2} \ge \chi^2_{\frac{\alpha}{2}}(n-1).$

上述检验法称为χ^2**检验法**。从以上的构造过程可知，k_1，k_2的取法可以不唯一，我们这样取完全在于方便计算，并且对于具有对称分布的检验统计量，这种取法具有优越性。

当μ 已知时，通常取检验统计量为 $\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$,

在H_0为真时，有 $$\chi^2=\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n).$$

可得 $$\aleph_0=\{(x_1,x_2,...,x_n):\chi^2=\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\geq\chi^2_{\frac{\alpha}{2}}(n)\ \text{or}\ \chi^2=\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\leq\chi^2_{1-\frac{\alpha}{2}}(n)\ \}.$$

但在μ已知时，也可用μ 未知时的χ^2检验法，两者比较，后者的拒绝域比前者有较小的犯第Ⅱ类错误的概率，因而更有效一些。

例 2 某厂生产的铜丝，质量一向比较稳定，今从中随机地抽出 10 根检查其折断力，测得数据(单位:千克)如下： 575 576 570 569 582 577 580 571 585 设铜丝的折断力服从正态分布 $N(\mu,\sigma^2)$，检验水平为α=0.05。试问：是否可以相信该厂的铜丝的折断力的方差为 64?

解 由题意知要检验假设 H_0：σ^2= 64，H_1：$\sigma^2\neq 64$，因为μ 未知，故检验统计量为 $$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1).$$

这里 $n=10,\alpha=0.05,\chi^2_{\frac{\alpha}{2}}(n-1)=\chi^2_{0.025}(9)=19.02,\chi^2_{1-\frac{\alpha}{2}}(n-1)=\chi^2_{0.975}(9)=2.70,$

并由样本算得 $$\overline{X}=575.7,(n-1)S^2=\sum_{i=1}^{10}(X_i-\overline{X})^2=260.1,$$

由此可算得 $$\chi_0^2=\frac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}=\frac{260.1}{64}\approx 4.06,$$

因为 2.70 ＜ 4.06 ＜ 19,02，根据χ^2检验法应接受 H_0，即认为这批铜丝的折断力的方差为 64。

例 3 在进行工艺改革时，如果方差显著增大，则改革需朝相反方向进行以减少方差；若方差变化不显著，需试行别的改革方案。现在加工 45 个活塞，对某项工艺进行改革，在新工艺下对加工好的 25 个活塞的直径进行测量，并由测量值算得样本方差 S^2=0.00066。已知在工艺改革前活塞直径的方差为 0.00040，问进一步改革的方向又如何?（α=0.05）

解 设测量值 X服从正态分布 $N(\mu,\sigma^2)$。已知工艺改革前方差σ^2=0.00040，现要确

定下一步改革的方向，并由题意可知，需考察改革后的活塞直径的方差σ^2是否不大于改革前的方差？因此待检假设可设为

$$H_0:\sigma^2 \le 0.00040, H_1:\sigma^2 > 0.00040,$$

这是一个复合假设，由前面的讨论可知，拒绝域为

$$\aleph_0 = \{(x_1,x_2,...,x_n):\frac{(n-1)s^2}{\sigma_0^2} > \chi_\alpha^2(n-1)\}$$

这里 n=25，S^2=0.00066，由α=0.05 查χ^2分布表得

$$\chi_\alpha^2(n-1) = \chi_{0.05}^2(24) = 36.415,$$

于是
$$\frac{(n-1)S^2}{\sigma_0^2} = \frac{24\times 0.00066}{0.00040} = 39.60 > 36.415,$$

故应拒绝 H_0，即改革后的方差显著小于改革前的方差，因此，下一步改革应朝相反方向进行。

下面考虑两个正态总体的参数检验问题。

3. 关于均值差 μ_1-μ_2 的假设检验

设样本($X_1, X_2, ..., X_n$)和($Y_1, Y_2, ..., Y_n$)分别取自正态总体 $X\sim N(\mu_1, \sigma_1^2)$和 $Y\sim N(\mu_2, \sigma_2^2)$, 且两个样本相互独立。

考虑假设 H_0：μ_1-μ_2= δ；H_1：μ_1-$\mu_2\ne\delta$，其中δ为已知常数。

当方差σ_1^2，σ_2^2为已知时， μ_1-μ_2的估计量$\overline{X}-\overline{Y}$在 H_0为真时有

$$\overline{X}-\overline{Y} \sim N(\delta, \frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}),$$

这时可取检验统计量为
$$U = \frac{\overline{X}-\overline{Y}-\delta}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \sim N(0,1),$$

易知一个水平为 α 的拒绝域为$|u| = |\frac{\overline{x}-\overline{y}-\delta}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}| \ge z_{\frac{\alpha}{2}}$

当方差σ_1^2，σ_2^2未知，但已知σ_1^2=σ_2^2时，可取检验统计量为
$$T = \frac{\overline{X}-\overline{Y}-\delta}{S_w\sqrt{\frac{1}{n}+\frac{1}{m}}}$$

当 H_0为真时，统计量 $T\sim t(m+n-2)$, 其中 $S_w^2 = \frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}$

这时可得水平为 α 的拒绝域为 $|t|=\left|\dfrac{\overline{x}-\overline{y}-\delta}{s_w\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}\right|\geq t_{\frac{\alpha}{2}}(n+m-2)$

例 4 用自动车床采用新旧两种工艺加工同种零件，测得的加工偏差(单位:微米)分别为

旧工艺	2.7	2.4	2.5	3.1	2.7	3.5	2.9	2.7	3.5	3.3
新工艺	2.6	2.1	2.7	2.8	2.3	3.1	2.4	2.4	2.7	2.3

设测量的加工偏差服从正态分布，所得的两个样本相互独立，且总体方差相等。试问自动车床在新旧两种工艺的加工精度有无显著差异？（$\alpha=0.01$）

解 由题意知要检验的假设为 $H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2$.

在 H_0 为真时，检验统计量为

$$T=\frac{|\overline{X}-\overline{Y}|}{S_w\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}\sim t(n+m-2)$$

由此可得水平为 α 的拒绝域为

$$|t|=\left|\frac{\overline{x}-\overline{y}}{s_w\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}\right|\geq t_{\frac{\alpha}{2}}(n+m-2).$$

这里 $m=n$，$\alpha=0.01$，故 $t_{\alpha/2}(m+n-2)=t_{0.005}(18)=2.8784$。并由样本算得

$$\overline{x}=2.93, \overline{y}=2.54, s_{\varpi}^2=0.125$$

于是

$$\frac{|\overline{X}-\overline{Y}|}{S_{\varpi}+\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}}=\frac{|2.93-2.54|}{\sqrt{\dfrac{0.125}{5}}}\approx 2.47<2.8784,$$

故接受 H_0，即认为新旧工艺对零件的加工精度无显著差异。

如果 $\sigma_1{}^2$，$\sigma_2{}^2$ 未知，且也不知两者是否相等，试验配对进行，此时有 $m=n$。这时可作变换 $Z_i=X_i-Y_i$，易知 $Z_i\sim N(\mu_1-\mu_2, \sigma_1{}^2+\sigma_2{}^2)$，$1\leq i\leq n$。这时，样本 $Z_1, Z_2, \ldots, Z_n$ 可视为来自单个正态总体 $N(\mu_1-\mu_2, \sigma_1{}^2+\sigma_2{}^2)$，于是检验假设 H_0：$\mu_1-\mu_2=\delta$ 相应地看作是单个正态主体在方差未知时检验均值是否为 δ 的假设，故这时水平为 α 的拒绝域为

$$\frac{|\overline{z}-\delta|}{s_z/\sqrt{n}}\geq t_{\frac{\alpha}{2}}(n-1).$$

其中

$$\overline{z}=\overline{x}-\overline{y}, s_z^2=\frac{1}{n-1}\sum_{i=1}^{n}[x_i-y_i-(\overline{x}-\overline{y})]^2.$$

例 5 若例 4 中新旧工艺加工同种零件是配对进行的，即第 i 台自动车床按照新旧工艺分别加工两个零件，其加工偏差分别为 x_i 和 y_i，$1\leq i\leq n$，若不知方差是否相同，检验 $H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2, \alpha=0.01.$

解 由样本算得 $\bar{z}=\bar{x}-\bar{y}=2.93-2.54=0.39, s_z^2=0.112$，

故 $$\frac{|\bar{z}|}{s_z/\sqrt{n}}=\frac{0.39}{0.106}=3.68$$

由 $\alpha=0.01$ 查 t 分布表得 $t_{\alpha/2}(n-1)=t_{0.005}(9)=3.2496$，由于

$$\frac{|\bar{z}|}{s_z/\sqrt{n}}=3.68>3.2496$$

故拒绝 H_0，而接受 H_1，即认为新旧工艺下的均值有显著差异。

由于试验方式不同，该例结论与例 4 的结论完全不同。这说明配对数据均值的假设检验比总体均值差的假设检验有较高的"灵敏度"。

4. 方差比 σ_1^2/σ_2^2 的假设检验

在实际问题中，需要考察两个总体的方差是否相等的问题，也就是说要检验假设：

$$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2.$$

由于样本方差 $$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2, S_2^2=\frac{1}{m-1}\sum_{j=1}^{m}(Y_i-\bar{Y})^2$$

分别是 σ_1^2 和 σ_2^2 的无偏估计。容易想到，当 H_0 为真时，S_1^2/S_2^2 应在 1 附近摆动，当此比值很大或很小时，H_0 都不大可能成立。因此可取统计量为 $F=\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\dfrac{S_1^2}{S_2^2}\cdot\dfrac{\sigma_1^2}{\sigma_2^2}$

在 H_0 为真时， $$F=\frac{S_1^2}{S_2^2}\sim F_{\frac{\alpha}{2}}(n-1,m-1)$$

由此可得水平为 α 的拒绝域为 $\dfrac{s_1^2}{s_2^2}\geq F_{\frac{\alpha}{2}}(n-1,m-1)$ 或 $\dfrac{s_1^2}{s_2^2}\leq F_{1-\frac{\alpha}{2}}(n-1,m-1)$

例 6 某一橡胶配方中，原用氧化锌 5 克，现减为 1 克。今分别对两种配方作一批试验，分别测得橡胶伸长率如下：

氧化锌 1 克	565	577	580	575	556	542	560	532	470	461
氧化锌 5 克	540	533	525	520	545	531	541	529	534	

设橡胶伸长率服从正态分布，问两种配方对伸长率总体方差有无显著差异?（$\alpha=0.1$）

解 根据题意需检验 $$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2.$$

这里 $n=10$，$m=9$，$\alpha=0.1$，从而可得

$$F_{\frac{\alpha}{2}}(n-1,m-1)=F_{0.05}(9,8)=3.39, F_{1-\frac{\alpha}{2}}(n-1,m-1)=F_{0.95}(9,8)=\frac{1}{F_{0.95}(8,9)}=\frac{1}{3.23},$$

由样本可算得 $$s_1^2 = 236.8, s_2^2 = 63.86$$

于是 $$\frac{s_1^2}{s_2^2} = \frac{236.8}{63.86} \approx 3.7 > 3.39,$$

故拒绝 H_0，即认为两总体方差有显著性差异.

例 7 现有甲、乙两台车床生产同一型号的滚珠。根据经验认为两台车床生产的滚珠直径都服从正态分布。现从这两台车床生产的产品中分别抽出 8 个和 9 个，测得的之间直径(单位:mm)分别为

甲	15.0	14.5	15.2	14.8	15.1	15.2	14.8
乙	15.2	14.8	15.2	15.0	15.0	14.8	15.1

试问：乙车床生产的滚珠直径的方差是否比甲车床生产的小？（$\alpha = 0.05$）

解 设 X, Y 分别表示甲、乙两台车床的滚珠的直径，即 $X \sim N(\mu_1, \sigma_1{}^2)$，$Y \sim N(\mu_2, \sigma_2{}^2)$，依题意需检验: $$H_0: \sigma_1^2 \le \sigma_2^2, H_1: \sigma_1^2 > \sigma_2^2$$

其水平为 α 的拒绝域为 $$\frac{s_1^2}{s_2^2} \ge F_\alpha(n-1, m-1).$$

这里 n=8，m = 9，$\alpha = 0.05$，从而 $F_\alpha(n-1, m-1) = F_{0.05}(7, 8) = 3.50$，由样本算得

$$s_1^2 = \frac{0.67}{7}, s_2^2 = \frac{0.21}{8}$$

于是 $$F = \frac{S_1^2}{S_2^2} = \frac{0.67}{0.21} \times \frac{8}{7} = 3.65 > 3.50.$$

故拒绝 H_0，即表示乙车床生产的滚珠直径的方差比甲车床生产的小。