第八章 假设检验

1. 假设检验的基本概念
2. 单个正态总体均值与方差的假设检验
3. 两个正态总体均值差与方差比的假设检验

§1 假设检验的基本概念

引例 某产品出厂检验规定, 次品率p不超过4%才能出厂. 现从一万件产品中任意抽查12件, 发现3件次品, 问该批产品能否出厂? 若抽查结果发现1件次品, 问能否出厂?

解 假设 $H_0: p \le 0.04$, $H_1: p > 0.04$,

$$P_{12}(3) = C_{12}^3 p^3 (1-p)^9 = 0.0097$$

小概率事件发生

故认为原假设不成立, 即该批产品次品率 $p > 0.04$, 则该批产品不能出厂.

$$P_{12}(1) = C_{12}^1 p^1 (1-p)^{11} = 0.306$$

这不是小概率事件,没理由拒绝原假设,从而接受原假设, 即该批产品可以出厂.

注 本检验方法是概率意义下的反证法.

对立
- 原假设或零假设 $H_0: p \le 0.04,$
- 备择假设或对立假设 $H_1: p > 0.04,$

假设检验的任务　必须在原假设与备择假设之间作一选择.

在原假设成立下, 利用样本观测值对实际问题进行分析, 如果发生小概率事件, 则拒绝接受原假设(接受备择假设). 反之,则接受原假设. 这种统计推断的问题称为假设检验问题.

只对总体中的未知参数提出假设,然后进行检验的问题称为参数检验.

理论基础　小概率事件在一次试验中几乎不可能发生.

假设检验中的小概率值称为显著性水平,记为$\alpha(0<\alpha<1)$. α 通常取为0.01,0.05或0.1.

例 某车间用包装机包装糖果,每袋重量(单位g)服从正态分布.由已往经验知每袋重量的标准差 $\sigma = 0.5g$ 保持不变. 每隔一定时间需要检查包装机的工作情况, 现抽取 9 袋,测得它们的重量为

$$99.0,\ 100.2,\ 99.3,\ 99.1,\ 99.6,\ 99.2,\ 99.9,\ 100.1,\ 99.3,$$

根据质量要求每袋重量为100g,试问这段时间内包装机的工作是否正常(取显著性水平 $\alpha = 0.05$)?

解 设每袋重量 $X \sim N(\mu, 0.5^2)$,

包装机的工作是否正常,相当于判断 $\mu = \mu_0 = 100$是否正确.

原假设H_0: $\mu = \mu_0 = 100$

备择假设H_1: $\mu \neq 100$,

因为样本均值$\overline{X}$是总体均值μ的无偏估计,所以$|\overline{X}-\mu|$应集中在0点附近,考虑$\left|\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}\right|$,在$H_0$正确条件下,$\mu=\mu_0$,适当选取一个常数$k$,若$\left|\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k$

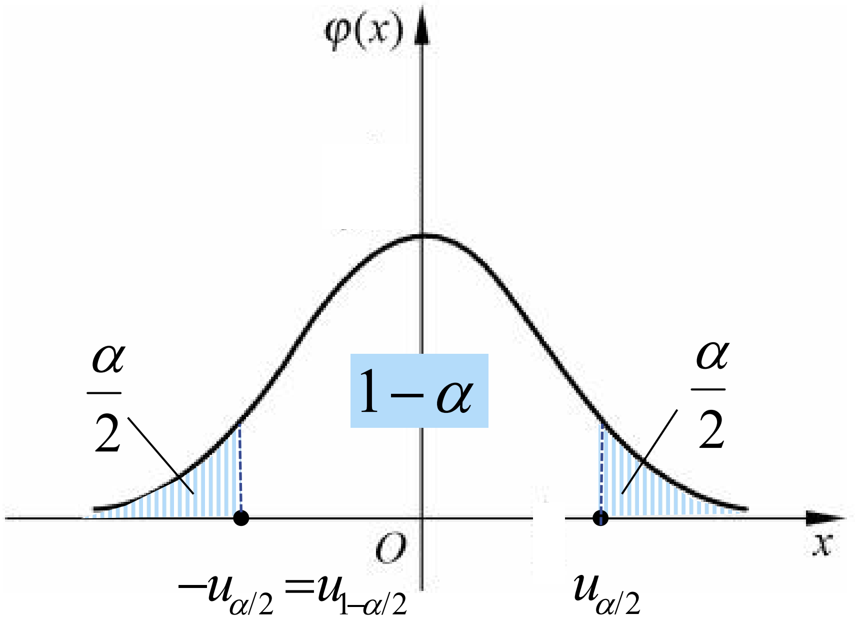
就有理由怀疑原假设H_0的正确性,应该拒绝H_0.

$$u=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$ —— 检验统计量

为了给出恰当的k值,对于给定的显著性水平α,令

$$P\left\{\frac{|\overline{X}-\mu_0|}{\sigma/\sqrt{n}}\geq k\right\}=\alpha.$$

根据标准正态分布上α分位点的定义知$k=u_{\alpha/2}$.



$$P\left\{\frac{|\overline{X}-\mu_0|}{\sigma/\sqrt{n}} \ge u_{\alpha/2}\right\} = \alpha .$$

如果统计量 u 满足 $|u| = \left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| \ge u_{\alpha/2}$,则意味着小概率事件 $\alpha = 0.05$ 发生,

根据实际推断原理,应拒绝原假设H_0.

如果统计量 u 满足 $|u| = \left|\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\right| < u_{\alpha/2}$ ，则应接受原假设H_0.

$W = \{ |u| \ge u_{\alpha/2} \}$ ———— 原假设H_0拒绝域

例 某车间用包装机包装糖果,每袋重量(单位g)服从正态分布.由已往经验知每袋重量的标准差 $\sigma=0.5g$ 保持不变. 每隔一定时间需要检查包装机的工作情况, 现抽取 9 袋,测得它们的重量为

$$99.0,\ 100.2,\ 99.3,\ 99.1,\ 99.6,\ 99.2,\ 99.9,\ 100.1,\ 99.3,$$

根据质量要求每袋重量为100g,试问这段时间内包装机的工作是否正常(取显著性水平 $\alpha=0.05$)?

解 设每袋重量 $X\sim N(\mu,0.5^2)$, H_0: $\mu=\mu_0=100$, H_1: $\mu\neq100$,

现在 $\mu_0=100, \sigma=0.5, n=9, \bar{x}=\dfrac{1}{9}\sum_{i=1}^{9}x_i=99.52, u_{\alpha/2}=1.96,\ W=\{|u|\geq u_{\alpha/2}\}$,

$$|u|=\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}=\frac{|99.52-100|}{0.5/\sqrt{9}}=2.88>1.96$$

所以拒绝H_0,即认为这段时间内包装机的工作不正常.

假设检验的结果可能犯两类错误

第一类错误(弃真) 当原假设H_0为真时, 作出的决定却是拒绝H_0, 犯这类错误的概率不超过显著性水平α,记为

$$P\{拒绝H_0 \mid H_0正确\} = \alpha.$$

第二类错误(取伪) 当原假设H_0不正确时,作出的决定却是接受H_0, 犯这类错误的概率不妨记为β,

$$P\{接受H_0 \mid H_0不正确\} = \beta.$$

	H_0为真	H_0不真
接受H_0	正确	第二类错误
拒绝H_0	第一类错误	正确

在确定检验法则时,应尽可能使犯两类错误的概率都较小.

当样本容量给定以后,若减少犯某一类错误的概率,则犯另一类错误的概率往往会增大,要使犯两类错误的概率都减小,只好增大样本容量.

在给定样本容量的情况下,我们总是控制犯第一类错误的概率,让它小于或等于α,而不考虑犯第二类错误的概率β,这类检验问题称为显著性检验.

只对总体中的未知参数提出假设,然后进行检验的问题称为**参数检验.**

对总体中的未知参数θ进行检验,

原假设H_0: $\theta=\theta_0$, 备择假设H_1: $\theta\neq\theta_0$. ——— 双边检验.

原假设H_0: $\theta=\theta_0(\theta\geq\theta_0)$, 备择假设$H_1$: $\theta<\theta_0$. ——— 左边检验.

原假设H_0: $\theta=\theta_0(\theta\leq\theta_0)$, 备择假设$H_1$: $\theta>\theta_0$. ——— 右边检验.

参数检验的一般步骤

1．根据问题的要求,提出原假设 H_0和备择假设 H_1;

2．在H_0正确下确定检验统计量T 及其分布;

3．对于给定的显著性水平α, 确定拒绝域W;

4．根据样本值计算T 的观察值t, 当 $t\in W$时,拒绝H_0, 否则接受H_0.

§2 单个正态总体的参数假设检验

1. 单个正态总体均值的假设检验
2. 单个正态总体方差的假设检验

假设总体 $X \sim N(\mu, \sigma^2)$, $X_1, X_2, \ldots, X_n$是来自总体X的样本,样本均值和样本方差分别为$\overline{X}, S^2$.

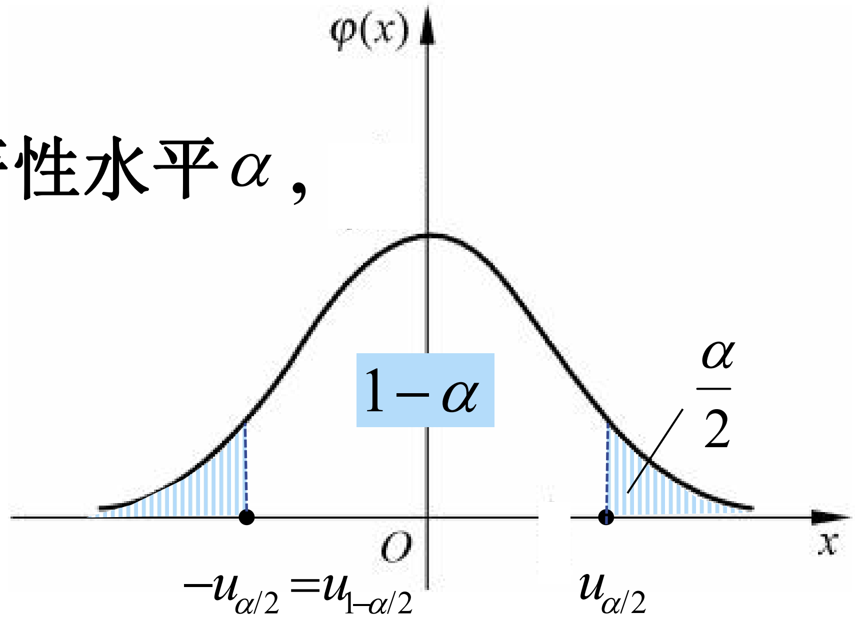
2.1 单个正态总体均值的假设检验

1. σ^2已知时,关于μ的假设检验 ——u检验

双边检验 H_0: $\mu = \mu_0$,H_1: $\mu \neq \mu_0$.

取检验统计量$u = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0, 1)$, 对于给定的显著性水平$\alpha$,

$$P\left\{\frac{|\overline{X} - \mu_0|}{\sigma / \sqrt{n}} \geq u_{\alpha/2}\right\} = \alpha .$$

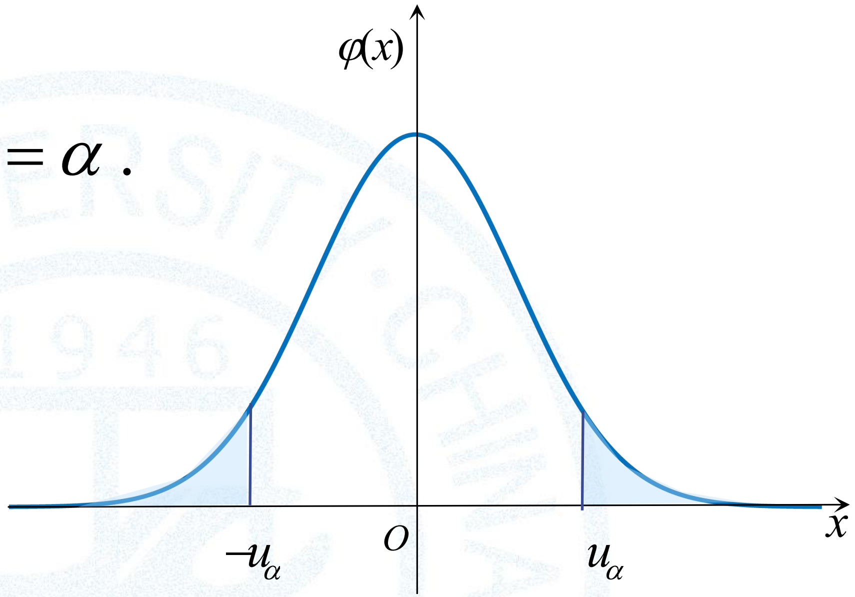


拒绝域为 $W = \{|u| \geq u_{\alpha/2}\}$

左边检验 $H_0: \mu = \mu_0 \left(\mu \geq \mu_0\right)$，$H_1: \mu < \mu_0$.

对于给定的显著性水平α, 考虑 $P\left\{\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \leq -u_\alpha\right\} = \alpha$.

拒绝域为 $W = \{u \leq -u_\alpha\}$



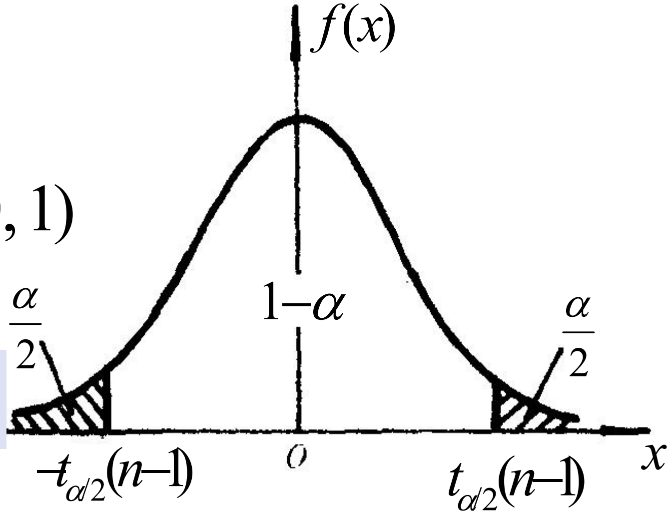
右边检验 $H_0: \mu = \mu_0 \left(\mu \leq \mu_0\right), H_1: \mu > \mu_0$.

对于给定的显著性水平α ,考虑 $P\left\{\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geq u_\alpha\right\} = \alpha$.

拒绝域为 $W = \{u \geq u_\alpha\}$

2. σ^2 未知时,关于μ的假设检验——t 检验

假设检验 $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$.



取检验统计量 $t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$. $u = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

对于给定的显著性水平α, 拒绝域为 $W = \{|t| \geq t_{\alpha/2}(n-1)\}$

左边检验 $H_0: \mu = \mu_0 \left(\mu \geq \mu_0\right)$, $H_1: \mu < \mu_0$.

对于给定的显著性水平α, 拒绝域为 $W = \{t \leq -t_\alpha(n-1)\}$

右边检验 $H_0: \mu = \mu_0 \left(\mu \leq \mu_0\right)$, $H_1: \mu > \mu_0$.

对于给定的显著性水平α, 拒绝域为 $W = \{t \geq t_\alpha(n-1)\}$

例 某车间加工一种零件,要求长度为150(单位:mm),今从一批加工后的这种零件中抽取 9 个,测得长度如下:

147, 150, 149, 154, 152, 153, 148, 151, 155

假设零件长度X服从正态分布,问这批零件是否合格(显著性水平$\alpha=0.05$)?

解 设$X\sim N(\mu,\sigma^2)$, 检验假设 H_0: $\mu=\mu_0=150$, H_1: $\mu\neq 150$.

检验统计量 $t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$. 拒绝域为 $W=\{|t|\geq t_{\alpha/2}(n-1)\}$.

这里 $n=9$, $\bar{x}=\dfrac{1}{9}\sum_{i=1}^{9}x_i=151$, $s^2=\dfrac{1}{9-1}\sum_{i=1}^{9}(x_i-\bar{x})^2=7.5$, $s=\sqrt{7.5}=2.739$,

α=0.05, $t_{\alpha/2}(n-1)=t_{0.025}(8)=2.306$.

$$|t|=\frac{|151-150|}{2.739/\sqrt{9}}=1.096<2.306.$$

所以接受H_0,即认为这批零件合格.

例 已知某厂生产的灯泡寿命X(单位h)服从正态分布$N(\mu,200^2)$,根据经验,灯泡的平均寿命不超过1500h,现测试了25只采用新工艺生产的灯泡,测得其平均寿命为 1575 h,试问新工艺是否提高了灯泡的寿命(显著性水平α=0.05).

解 $H_0: \mu = 1500(\mu \le 1500),\ H_1: \mu > 1500.$

检验统计量 $u = \dfrac{\overline{X} - 1500}{\sigma / \sqrt{n}} \sim N(0,1)$

对于给定的显著性水平 α= 0.05, 查表得 $u_\alpha = u_{0.05} = 1.645.$

左边检验,拒绝域为 $W = \{u \ge u_\alpha\} = \{u \ge 1.645\}$

由于 n=25, σ=200, $\bar{x}$= 1575, $u = \dfrac{1575 - 1000}{200/\sqrt{25}} = 1.875 > 1.645$

所以拒绝H_0,而接受H_1,即认为新工艺提高了灯泡的寿命.

2.2 单个正态总体方差的假设检验

1. 均值μ已知时,方差σ^2的假设检验——χ^2检验

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$	$\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n)$ 或 $\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n)$
$\sigma^2\geq\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\leq\chi^2_{1-\alpha}(n)$
$\sigma^2\leq\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\geq\chi^2_{\alpha}(n)$

2．均值μ未知时，方差σ^2的假设检验——χ^2检验

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2\le\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2\ge\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2\ge\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\le\chi^2_{1-\alpha}(n-1)$
$\sigma^2\le\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\ge\chi^2_{\alpha}(n-1)$

例 某厂生产的尼龙纤维的纤度在正常情况下服从正态分布,某日抽取5根纤维,测得它们的纤度为

$$1.32, 1.36, 1.55, 1.44, 1.40$$

试问能否认为这一天尼龙纤维的纤度的标准差σ=0.048(取$\alpha=0.1$).

解 检验假设 H_0: $\sigma=0.048$, H_1: $\sigma\neq0.048$.

检验统计量$\chi^2=\dfrac{(n-1)S^2}{0.048^2}\sim\chi^2(n-1)$. 拒绝域 $W=\{\chi^2\le\chi^2_{1-\alpha/2}(n-1)$ 或 $\chi^2\ge\chi^2_{\alpha/2}(n-1)\}$

这里$\alpha=0.1$, $\chi^2_{1-\alpha/2}(n-1)=\chi^2_{0.95}(4)=0.711$, $\chi^2_{\alpha/2}(n-1)=\chi^2_{0.05}(4)=9.488$,

$\bar{x}=1.414$, $s^2=0.00778$,

$$\chi^2=\frac{(5-1)\times0.00778}{0.048^2}=13.51>9.488.$$

所以拒绝H_0,即不能认为这一天尼龙纤度的标准差$\sigma=0.048$.

§3 两个正态总体的参数假设检验

1. 两个正态总体均值差的假设检验
2. 两个正态总体方差比的假设检验

设总体$X\sim N(\mu_1,\sigma_1^2)$与$Y\sim N(\mu_2,\sigma_2^2)$独立, $X_1,X_2,\cdots,X_{n_1}$和$Y_2,Y_2,\cdots,Y_{n_2}$分别为来自总体X与Y的样本.

3.1 两个正态总体均值差的假设检验

1.方差 σ_1^2,σ_2^2 已知时, 均值差 $\mu_1-\mu_2$ 的假设检验——u检验

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2\neq\delta$	$u=\dfrac{\bar X-\bar Y-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\lvert u\rvert\geq u_{\frac{\alpha}{2}}$
$\mu_1-\mu_2\geq\delta$	$\mu_1-\mu_2<\delta$		$u\leq -u_\alpha$
$\mu_1-\mu_2\leq\delta$	$\mu_1-\mu_2>\delta$		$u\geq u_\alpha$

2.方差$\sigma_1^2=\sigma_2^2=\sigma^2$未知时，均值差$\mu_1-\mu_2$的假设检验——t检验

原假设 H_0	备择假设 H_1	检验统计量及其在H_0为真时的分布	拒绝域
$\mu_1-\mu_2=\delta$	$\mu_1-\mu_2\neq\delta$	$t=\dfrac{\bar{X}-\bar{Y}-\delta}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}S_w}\sim t(n_1+n_2-2)$	$\lvert t\rvert\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$
$\mu_1-\mu_2\geq\delta$	$\mu_1-\mu_2<\delta$		$t\leq -t_{\alpha}(n_1+n_2-2)$
$\mu_1-\mu_2\leq\delta$	$\mu_1-\mu_2>\delta$		$t\geq t_{\alpha}(n_1+n_2-2)$

其中$S_w=\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$

例 在甲、乙两个工厂生产的蓄电池中,分别取5个测量电容量,

甲厂：143 141 138 142 140

乙厂：141 143 139 144 141

设甲,乙两厂蓄电池的电容量分别为$X \sim N(\mu_1,\sigma_1^2)$和$Y \sim N(\mu_2,\sigma_2^2)$,且$\sigma_1^2=\sigma_2^2$. 问两厂的电容量有无显著差异(取$\alpha$=0.05)?

解 检验假设H_0：$\mu_1=\mu_2$, H_1：$\mu_1\neq\mu_2$.

检验统计量$t=\dfrac{\bar{X}-\bar{Y}}{s_w\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2)$. 拒绝域为$W=\{\,|t|\geq t_{\alpha/2}(8)=2.306\,\}$.

这里 $\bar{x}=140.8, \bar{y}=141.6, s_1^2=3.699, s_2^2=3.799, s_w^2=\dfrac{4\times s_1^2+4s_2^2}{8}=3.7499$

$$|t|=\frac{|140.8-141.6|}{\sqrt{3.7499}\times\sqrt{1/5+1/5}}=0.6535<2.306,$$

因此接受原假设H_0,即认为两厂蓄电池的电容量无显著差异.

例 在甲、乙两个工厂生产的蓄电池中,分别取5个测量电容量,

甲厂：143 141 138 142 140

乙厂：141 143 139 144 141

设甲,乙两厂蓄电池的电容量分别为$X\sim N(\mu_1,\sigma_1^2)$和$Y\sim N(\mu_2,\sigma_2^2)$,且$\sigma_1^2=\sigma_2^2$. 问两厂的电容量有无显著差异(取$\alpha$=0.05)?

又解 两个样本容量相等,化为单个总体$Z=X-Y$均值是否为零的检验.

设$Z=X-Y$,则$Z\sim N(\mu,\sigma^2)$, 其中$\mu=\mu_1-\mu_2,\sigma^2=\sigma_1^2+\sigma_2^2$. 样本值为2,−2,−1,−2,−1.

检验假设H_0：$\mu=0$, H_1：$\mu\neq0$.

检验统计量 $t=\dfrac{\bar{Z}}{S/\sqrt{n}}\sim t(n-1)$. 拒绝域$W=\{|t|\geq t_{0.025}(4)=2.7764\}$.

$$|t|=\frac{|\bar{z}|}{s/\sqrt{n}}=\frac{0.8}{\sqrt{2.699}/\sqrt{5}}=0.2177<2.7764.$$

因此接受原假设H_0,即认为两厂蓄电池的电容量无显著差异

3.2 两个正态总体方差比的假设检验

1．均值μ_1, μ_2已知时,方差比 σ_1^2/σ_2^2 的假设检验——F检验

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$F=\dfrac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}\cdot\dfrac{n_2}{n_1}\sim F(n_1,n_2)$.	$F\le F_{1-\frac{\alpha}{2}}(n_1,n_2)$ 或 $F\ge F_{\frac{\alpha}{2}}(n_1,n_2)$
$\sigma_1^2\ge\sigma_2^2$	$\sigma_1^2<\sigma_2^2$		$F\le F_{1-\alpha}(n_1,n_2)$
$\sigma_1^2\le\sigma_2^2$	$\sigma_1^2>\sigma_2^2$		$F\ge F_{\alpha}(n_1,n_2)$

2．均值μ_1, μ_2未知时,方差比 σ_1^2/σ_2^2 的假设检验——F检验

原假设 H_0	备择假设 H_1	检验统计量及其在H_0为真时的分布	拒绝域
$\sigma_1^2=\sigma_2^2$	$\sigma_1^2\neq\sigma_2^2$	$F=\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$	$F\le F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)$ 或 $F\ge F_{\frac{\alpha}{2}}(n_1-1,n_2-1)$
$\sigma_1^2\ge\sigma_2^2$	$\sigma_1^2<\sigma_2^2$		$F\le F_{1-\alpha}(n_1-1,n_2-1)$
$\sigma_1^2\le\sigma_2^2$	$\sigma_1^2>\sigma_2^2$		$F\ge F_{\alpha}(n_1-1,n_2-1)$

例 从两个正态总体分别独立抽取样本观察值如下：

甲：**4.4 4.0 2.0 4.8**　　　乙：**6.0 1.0 3.2 0.4**

能否认为两个样本观察值来自同一总体(取α=0.05).

解 设两个正态总体分别为$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$,

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

检验统计量$F = \dfrac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$, 拒绝域为$W = \{F \le F_{1-\frac{\alpha}{2}}(3,3)$或$F \ge F_{\frac{\alpha}{2}}(3,3)\}$,

查表得$F_{0.025}(3,3)$=15.44, $F_{0.925}(3,3) = \dfrac{1}{F_{0.025}(3,3)} = \dfrac{1}{15.44} = 0.065$,

这里 $\bar{x} = 3.8$, $\bar{y} = 2.6$, $s_1^2 = 1.55$, $s_2^2 = 6.44$,

$$0.065 < \frac{s_1^2}{s_2^2} = \frac{1.55}{6.44} = 0.24 < 15.44,$$

因此接受原假设H_0,即认为两个正态总体的方差相同.

例 从两个正态总体分别独立抽取样本观察值如下：

甲：4.4 4.0 2.0 4.8　　　乙：6.0 1.0 3.2 0.4

能否认为两个样本观察值来自同一总体(取α=0.05).

解 再检验假设 H_0^*：$\mu_1=\mu_2$，H_1^*：$\mu_1\neq\mu_2$

由于$\sigma_1^2=\sigma_2^2$但未知,取检验统计量为 $t=\dfrac{\bar{X}-\bar{Y}}{s_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)$.

拒绝域为$W=\{|t|\geq t_{\alpha/2}(6)\}$，查表得$t_{0.025}(6)=2.4469$,

$$|t|=\frac{|3.80-2.65|}{\sqrt{\dfrac{3\times1.55+3\times6.44}{6}}\cdot\sqrt{\dfrac{1}{4}+\dfrac{1}{4}}}=0.82<2.4469$$

因此接受H_0^*,即认为两个正态总体的均值相同.

综上在显著性水平α=0.05下,认为两个样本值来自同一总体.