第八章 假设检验

- 假设检验
- 正态总体均值的假设检验
- 正态总体方差的假设检验
- 置信区间和假设检验的关系
- 样本容量的选取
- 分布拟合检验
- 假设检验问题的p值检验法

§8.1 假设检验

统计推断的另一类重要问题是假设检验问题。它包括

（1）已知总体分布的形式，需对其中的未知参数给出假设检验. —参数检验

（2）总体的分布形式完全未知的情况下，对总体的分布或数字特征进行假设检验.—非参数检验

§ 8.1.1 问题的提出

例1 设某种清漆的9个样品，其干燥时间（以小时计）分别为：

6.0 5.7 5.5 6.5 7.0 5.8 5.2 6.1 5.0

根据以往经验，干燥时间的总体服从正态分布N(6.0, 0.36),现根据样本检验均值是否与以往有显著差异？

例2 一种摄影药品被其制造商声称其贮藏寿命是均值180天、标准差不多于10天的正态分布。某位使用者担心标准差可能超过10天。他随机选取12个样品并测试，得到样本标准差为14天。根据样本有充分证据证明标准差大于10天吗？

假设

在假设检验中，常把一个被检验的假设称为原假设，用H_0表示，通常将不应轻易加以否定的假设作为原假设.

当原假设H_0被拒绝时而接受的假设称为备择假设，用H_1表示，它们通常成对出现.

$$H_0:\ \theta = \theta_0,\quad H_1:\ \theta \neq \theta_0\text{（双边检验）}$$

$$H_0:\ \theta \leq \theta_0,\quad H_1:\ \theta > \theta_0\text{（右边检验）}$$

$$H_0:\ \theta \geq \theta_0,\quad H_1:\ \theta < \theta_0\text{(左边检验）}$$

§ 8.1.2 检验统计量和拒绝域

如果统计量$T = T(X_1, \cdots, X_n)$的取值大小和原假设H_0是否成立有密切联系，可将之称为对应假设问题的检验统计量，对应于拒绝原假设H_0时，样本值的范围称为拒绝域，记为W,其补集$\overline{W}$称为接受域.

■ 对例1的统计分析

设清漆的干燥时间为X，由已知$X \sim N(\mu，\sigma^2)$，

其中σ^2=0. 36，考虑有关参数μ的假设：

$$H_0: \mu = 6.0, H_1: \mu \neq 6.0 \text{（双边检验）}$$

因样本均值$\bar{X}$是μ的无偏估计，$\bar{X}$的取值大小反映了μ的取值大小，当原假设成立时，$|\bar{X}-6.0|$取值应偏小。

检验规则：其中C是待定的常数.

当$|\bar{X}-6.0| \geq C$时，拒绝原假设H_0;

当$|\bar{X}-6.0| < C$时，接受原假设H_0，

C?

上述例子中，可取检验统计量为$\bar{X}$(或$\bar{X}-6.0$),拒绝域为

$$W=\{(X_1,\cdots,X_n):|\bar{X}-6.0| \geq C\}$$

§ 8.1.3 两类错误

由于作出决策的依据是一个样本，因此，可能出现“实际上原假设成立，但根据样本作出拒绝原假设”或者“实际上原假设不成立，但根据样本作出接受原假设”的决策。

在假设检验中有四种可能结果：

决策	原假设H_0	
	真的	假的
接受H_0	正确决策	第二类错误
拒绝H_0	第一类错误	正确决策

第一类错误：原假设H_0成立时，作出拒绝原假设的决策；
第二类错误：备择假设H_1成立时，作出接受原假设的决策。

犯第I类错误的概率

$$\alpha = P\{\text{第I类错误}\} = P\{\text{拒绝}H_0 \mid H_0\text{是真的}\}$$

犯第II类错误的概率

$$\beta = P\{\text{第II类错误}\} = P\{\text{接受}H_0 \mid H_0\text{是假的}\}$$

实际中常常将犯第一类错误的概率控制在一定限度内，即事先给定较小的数α（0< α <1）（称为显著性水平），使得

$$P_{H_0}(\left|\overline{X} - 6.0\right| \geq c) \leq \alpha$$

例1中，给定显著性水平$\alpha = 0.05$，当原假设成立时，

总体$X \sim N(6.0, 0.6^2)$,因此，$\bar{X} \sim N(6.0, \frac{0.6^2}{9})$

$$P(\left|\bar{X} - 6.0\right| \geq c) = P(\frac{\left|\bar{X} - 6.0\right|}{0.6/3} \geq \frac{c}{0.6/3}) = \alpha$$

拒绝域为：$\dfrac{\left|\bar{X} - 6.0\right|}{0.2} \geq z_{\alpha/2} = z_{0.025} = 1.96$

根据样本得$\bar{x} = 5.87, \dfrac{\left|\bar{x} - 6\right|}{0.2} = 0.67 < 1.96 = z_{0.025}$.

即$\bar{x}$不落在拒绝域内，$\bar{x}$与$\mu = 6$的差异不显著，因此接受原假设，认为干燥时间的均值与以往无显著差异。

例1中，犯第I类错误的概率

$$\alpha(C)=P\{拒绝H_0 \mid H_0是真的\}$$

$$=P\{|\bar{X}-6.0|\geq C \mid \mu=6.0\}$$

$$=P\{\frac{|\bar{X}-6.0|}{\sigma/\sqrt{n}}\geq\frac{C}{\sigma/\sqrt{n}} \mid \mu=6.0\}$$

由于当H_0成立时，即μ=6. 0时，$\frac{\bar{X}-6.0}{\sigma/\sqrt{n}}\sim N(0,1)$,因此

$$\alpha(C)=2-2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right)$$

犯第II类错误的概率

$$
\begin{aligned}
\beta(C) &= P\{接受H_0 \mid H_0是假的\} \\
&= P\{|\bar{X}-6.0|<C \mid \mu \neq 6.0\} \\
&= P\{6.0-C<\bar{X}<6.0+C \mid \mu \neq 6.0\} \\
&= P\left\{\frac{6.0-C-\mu}{\sigma/\sqrt{n}}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<\frac{6.0+C-\mu}{\sigma/\sqrt{n}} \,\middle|\, \mu \neq 6.0\right\} \\
&= \Phi\left\{\frac{6.0+C-\mu}{\sigma/\sqrt{n}}\right\}-\Phi\left\{\frac{6.0-C-\mu}{\sigma/\sqrt{n}}\right\}, \quad \mu \neq 6.0
\end{aligned}
$$

显然，犯第I类错误的概率$\alpha(C)$关于C是单调减函数，而犯第II类错误的概率$\beta(C)$关于C是单调增函数.

在给定的样本量下n，不可能找界值C，使得$\alpha(C)$和$\beta(C)$都尽能小.

犯两类错误的概率相互制约！

Neyman-Pearson原则：

首先控制犯第I类错误的概率不超过某个常数 $\alpha\in(0,1)$，再寻找检验，使得犯第II类错误的概率尽可能小.

其中的常数α称为显著水平.

常取α=0.01, 0.05, 0.1等.

在例1中，若取显著水平$\alpha=0.05$，则有

$$\alpha(C)=2-2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right)\leq 0.05$$

计算得

$$\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right) \geqslant 0.975 \qquad \frac{C}{\sigma/\sqrt{n}} \geqslant z_{0.025}$$

$$C\geq z_{0.025}\sigma/\sqrt{n}=1.96\times 0.6/3=0.392$$

由于犯第II类错误的概率$\beta(C)$关于C单调增函数，根据$\mathrm{N}eyman-Pearson$原则，应取

$$C=0.392$$

因此拒绝域为

$$W=\left\{(X_1,\cdots,X_9):|\bar{X}-6.0|\geq 0.392\right\}$$

根据实际样本资料，$\bar{x}=5.87$，$|\bar{x}-6.0|=0.13<0.392$，样本落入接受域.

根据上述检验规则，犯第I类错误的概率

α（0. 392）=0. 05=α

犯第II类错误的概率

$$\beta(0.392) = \Phi\left\{\frac{6.0+0.392-\mu}{0.6/\sqrt{9}}\right\} - \Phi\left\{\frac{6.0-0.392-\mu}{0.6/\sqrt{9}}\right\}$$

$$= \Phi\left\{\frac{6.392-\mu}{0.2}\right\} - \Phi\left\{\frac{5.608-\mu}{0.2}\right\}, \qquad \mu \neq 6.0$$

例如，当μ=5. 4时，犯第II类错误的概率

$$\beta = \Phi\left\{\frac{6.392-5.4}{0.2}\right\} - \Phi\left\{\frac{5.608-5.4}{0.2}\right\} = \Phi(4.96) - \Phi(1.04)$$

$$\approx 1.00-0.85$$

$$=0.15$$

处理假设检验问题的基本步骤

1. 根据实际问题的要求，提出原假设H_0和备择假设H_1；

2. 给定显著性水平α以及样本容量n；

3. 确定检验统计量以及拒绝域的形式；

4. 按照$P\{当H_0为真拒绝H_0\} \le \alpha$求出拒绝域；

5. 根据样本观测值作出决策，接受原假设H_0还是拒绝原假设H_0 .

§ 8.2 正态总体均值的假设检验

设样本$X_1, X_2, \cdots, X_n$来自正态总体$N\left(\mu, \sigma^2\right)$,

$\overline{X}$和S^2分别为样本均值和方差,显著性水平为 α

§ 8.2.1 单个总体的情况

(1) σ^2已知时---Z检验

双边假设问题

$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$， 其中$\mu_0$是已知的常数

检验拒绝域形式为 $\left|\bar{X} - \mu_0\right| \geq c$

在H_0为真时， $\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

根据犯第一类错误概率不大于α， 即$P\left\{\dfrac{\left|\bar{X} - \mu_0\right|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\} = \alpha$

拒绝域为: $\boxed{Z = \dfrac{\left|\bar{X} - \mu_0\right|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}.}$ Z检验法

右边假设问题

$H_0: \mu \le \mu_0, H_1: \mu > \mu_0$， 其中$\mu_0$是已知的常数

检验统计量仍取为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}.$$

检验的拒绝域为 $W = \left\{ Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \ge z_\alpha \right\}$

左边假设问题

$H_0:\mu\geq\mu_0, H_1:\mu<\mu_0$， 其中$\mu_0$是已知的常数

检验统计量仍取为

$$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}.$$

拒绝域形式为

$$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\leq C.$$

检验的拒绝域为 $W=\left\{Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\leq -z_\alpha\right\}$

例1 公司从生产商购买牛奶.公司怀疑生产商在牛奶中掺水以谋利.通过测定牛奶的冰点，可以检验出牛奶是否掺水.天然牛奶的冰点温度近似服从正态分布，均值$\boldsymbol{\mu_0} = -\mathbf{0.545}$℃，标准差$\boldsymbol{\sigma} = \mathbf{0.008}$℃.牛奶掺水可使冰点温度升高而接近水的冰点温度(0 ℃).测得生产商提交的5批牛奶的冰点温度，其均值为$\bar{x} = -\mathbf{0.535}$℃，问是否可以认为生产商在牛奶中掺了水？取$\boldsymbol{\alpha} = \mathbf{0.05}$. 教材（P182例2）

$$H_0: \mu \le \mu_0 = -0.545$$

$$H_1: \quad \mu > \mu_0$$

(2) σ^2未知时---t检验

双边假设问题

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$$

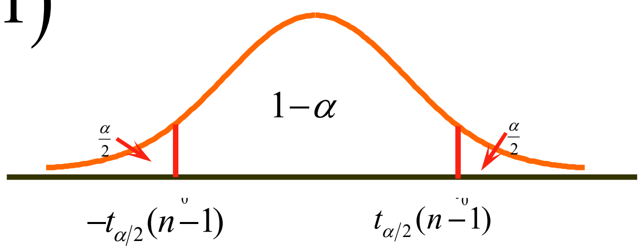
由于σ^2未知，故不能用$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$来确定拒绝域.

用σ的估计量S代替σ,

采用$T = \dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$作检验统计量。

即检验拒绝域的形式为 $\dfrac{\left|\bar{X}-\mu_0\right|}{S/\sqrt{n}} \geq k.$

当原假设成立时， $\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$



检验的拒绝域为 $T=\dfrac{\left|\bar{X}-\mu_0\right|}{S/\sqrt{n}} \geq t_{\alpha/2}(n-1)$

左边假设问题

$H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$， 其中$\mu_0$是已知的常数

拒绝域为

$$W = \left\{ T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_\alpha(n-1) \right\}$$

右边假设问题

$H_0: \mu \le \mu_0, H_1: \mu > \mu_0$， 其中$\mu_0$是已知的常数

拒绝域为

$$W = \left\{ T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \ge t_\alpha(n-1) \right\}$$

26

例2　某种元件的寿命X（以小时记）服从正态分布$N(\mu,\sigma^2)$，μ,σ^2均未知。现测得16只元件的寿命如下：

159	280	101	212	224	379	179	264
222	362	168	250	149	260	485	170

问是否有理由认为元件的平均寿命大于225（小时）？

（取显著性水平为0.05）　教材P184例1

27

解：按题意需检验

$$H_0 : \mu \le \mu_0 = 225, \qquad H_1 : \mu > 225.$$

拒绝域为： $T = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \ge t_\alpha(n-1).$

$n = 16, t_{0.05}(15) = 1.7531. \quad \bar{x} = 241.5, s = 98.7259$

计算得：$t_0 = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = 0.6685 < 1.7531 = t_{0.05}(15).$

没有落在拒绝域内，故不能拒绝原假设，

认为元件的平均寿命不大于225小时。

问：若将原假设和备择假设互换，即考虑左边检验，检验结果怎么样？

$$H_0: \mu \geq \mu_0 = 225, \qquad H_1: \mu < 225.$$

练习 设某次考试的考生成绩服从正态分布，从中随机地取36位考生的成绩，算得平均成绩为66.5分，标准差为15分，问在显著水平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？

§ 8.2.2 两个总体的情况

$X_1, X_2, \cdots, X_{n_1}$来自$N\left(\mu_1, \sigma^2\right), Y_1, Y_2, \cdots, Y_{n_2}$来自$N\left(\mu_2, \sigma^2\right), \mu_1, \mu_2, \sigma^2$未知

$\bar{X}, \bar{Y}, S_1^2, S_2^2$分别为第一,二个总体的样本均值和方差,显著性水平为 α.

双边检验

$$H_0:\mu_1-\mu_2=\delta, H_1:\mu_1-\mu_2\neq\delta. \quad (\delta\text{为已知常数})$$

检验拒绝域的形式为：$\left|\bar{X}-\bar{Y}-\delta\right|\geq k$

即等价于 $\dfrac{\left|\bar{X}-\bar{Y}-\delta\right|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\geq c$ 　其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}, S_w=\sqrt{S_w^2}$

在原假设成立时，$t=\dfrac{\left(\bar{X}-\bar{Y}-\delta\right)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$

检验拒绝域为：$\boxed{t=\dfrac{\left|\bar{X}-\bar{Y}-\delta\right|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\geq t_{\alpha/2}(n_1+n_2-2)}$ 　t检验法

右边检验

$$H_0:\mu_1\le\mu_2,H_1:\mu_1>\mu_2 \qquad \sigma_1^2=\sigma_2^2=\sigma^2,\sigma^2\text{未知}$$

检验拒绝域形式为：$$\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\ge c$$

其中 $$S_w^2=\frac{\left(n_1-1\right)S_1^2+\left(n_2-1\right)S_2^2}{n_1+n_2-2},S_w=\sqrt{S_w^2}$$

当$\mu_1=\mu_2$时， $$\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t\left(n_1+n_2-2\right)$$

从而，拒绝域为：$$\boxed{t=\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\ge t_\alpha(n_1+n_2-2)}$$

左边检验

$$H_0:\mu_1\ge\mu_2,\ H_1:\mu_1<\mu_2\qquad \sigma_1^2=\sigma_2^2=\sigma^2,\sigma^2\text{未知}$$

检验拒绝域形式为：$$\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\le -c$$

其中 $$S_w^2=\frac{\left(n_1-1\right)S_1^2+\left(n_2-1\right)S_2^2}{n_1+n_2-2},\ S_w=\sqrt{S_w^2}$$

当$\mu_1=\mu_2$时，$$\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t\left(n_1+n_2-2\right)$$

从而，拒绝域为：$$\boxed{t=\frac{\left(\bar{X}-\bar{Y}\right)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\le -t_\alpha(n_1+n_2-2)}$$

例3 在相同条件下对甲、乙两种品牌的洗涤剂分别进行去污试验，测得去污率（%）结果如下：

品牌甲：79.4，80.5,76.2,82.7,77.8,75.6;

品牌乙：73.4,77.5,79.3,75.1,74.7.

假定两品牌的去污率服从正态分布且方差相同，试问两品牌的去污率有无明显差异？（显著性水平为0.05）

解：设X、Y分别为甲、乙两种品牌的去污率，由题设可知$X\sim N(\mu_1,\sigma^2)$, $Y\sim N(\mu_2,\sigma^2)$，建立假设：

$$H_0: \mu_1-\mu_2=0, \quad H_1: \mu_1-\mu_2 \neq 0$$

由题意$n_1=6,\bar{x}=78.4,s_1^2=7.28$, $n_2=5,\bar{y}=76,s_2^2=5.6$

查表得$t_{0.025}(9)=2.262$，计算可得

$$t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}=1.034<t_{0.025}(9)$$

故接受H_0，即认为两品牌的去污率无明显差异.

§ 8.2.3 基于成对数据的检验

逐对比较法：

为了比较两种产品、两种仪器、两种方法等的差异，常在相同的条件下做对比试验，得到一批成对的观察值，然后分析观察数据作出推断.

一般，设有n对独立的观测结果：$(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$，

令$D_1=X_1-Y_1,D_2=X_2-Y_2,\ldots,D_n=X_n-Y_n$，

则$D_1,D_2,\ldots,D_n$相互独立，服从同一分布，设为$N(\mu_D,\sigma_D^2)$.

分别将$D_1,D_2,\ldots,D_n$的样本均值和样本方差

的观测值记为$\bar{d},S_D^2$，

转化为单个正态总体的均值的假设检验。

双边假设 $H_0: \mu_d = 0 \leftrightarrow \mu_d \neq 0$

$$拒绝域为: \frac{|\bar{d}|}{S_D/\sqrt{n}} \geq t_{\alpha/2}(n-1),$$

左边假设 $H_0: \mu_d \geq 0,\ H_1: \mu_d < 0$

$$拒绝域为: \frac{\bar{d}}{S_D/\sqrt{n}} \leq -t_\alpha(n-1)$$

右边假设 $H_0: \mu_d \leq 0,\ H_1: \mu_d > 0$

$$拒绝域为: \frac{\bar{d}}{S_D/\sqrt{n}} \geq t_\alpha(n-1)$$

例4 为了试验两种不同谷物种子的优劣，选取了十块土质不同的土地，并将每块土地分为面积相同的两部分，分别种植这两种种子。设在每块土地的两部分人工管理等条件完全一样。下面给出各块土地上的产量。

土地	1	2	3	4	5	6	7	8	9	10
种子A(xi)	23	35	29	42	39	29	37	34	35	28
种子B(yi)	26	39	35	40	38	24	36	27	41	27
di=xi-yi	-3	-4	-6	2	1	5	1	7	-6	1

问：以这两种种子种植的谷物产量是否有显著的差异（取显著性水平为0.05)？

40

解：检验假设 $H_0:\mu_D=0, H_1:\mu_D\neq 0$

分别将$D_1,D_2,...,D_n$的样本均值和样本方差记为$\bar{D}, S_D^2$,

拒绝域为：$\dfrac{\left|\bar{D}\right|}{S_D/\sqrt{n}}\geq t_{\alpha/2}\left(n-1\right)$,

$n=10$, 查表得：$t_{0.025}\left(9\right)=2.2622, \bar{d}=-0.2, s_d=4.442$,

计算得：$\dfrac{\left|\bar{d}\right|}{s_d/\sqrt{n}}=0.142<2.2622$

接受原假设H_0，认为两种种子的产量没有显著差异。

§ 8.3 正态总体方差的假设检验

§ 8.3.1 单个总体的情况

双边检验 $H_0:\sigma^2=\sigma_0^2, H_1:\sigma^2\neq\sigma_0^2$

其中σ_0^2是已知常量。此时σ^2的无偏估计量为样本方差S^2,

且有 $$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$$

因此可取检验统计量为 $$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}$$

检验拒绝域形式为:

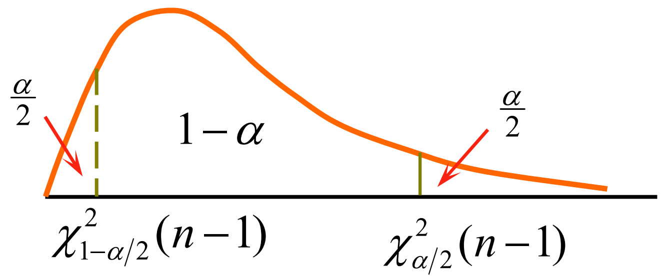
$$\frac{(n-1)S^2}{\sigma_0^2}\leq k_1, \text{或} \frac{(n-1)S^2}{\sigma_0^2}\geq k_2.$$

在原假设成立时, $$\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$$

$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2} \le k_1, \text{或} \frac{(n-1)S^2}{\sigma_0^2} \ge k_2\right\} = \alpha$$

为计算方便，习惯上取

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



于是有 $k_1 = \chi^2_{1-\alpha/2}(n-1)$，$k_2 = \chi^2_{\alpha/2}(n-1)$。

拒绝域为：$\boxed{\frac{(n-1)S^2}{\sigma_0^2} \le \chi^2_{1-\alpha/2}(n-1), \text{或} \frac{(n-1)S^2}{\sigma_0^2} \ge \chi^2_{\alpha/2}(n-1)}$ χ^2检验法

43

类似地，对于左边检验

$$H_0:\sigma^2 \geq \sigma_0^2, H_1:\sigma^2 < \sigma_0^2$$

拒绝域为：
$$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi_{1-\alpha}^2(n-1);$$

类似地，对于右边检验

$$H_0:\sigma^2 \leq \sigma_0^2, H_1:\sigma^2 > \sigma_0^2$$

拒绝域为：
$$\frac{(n-1)S^2}{\sigma_0^2} \geq \chi_{\alpha}^2(n-1);$$

44

例1 一个园艺科学家正在培养一个新品种的苹果,这种苹果除了口感好和颜色鲜艳以外,另一个重要特征是单个重量差异不大（对照品种的方差$\sigma^2=7$)。为了评估新苹果,她随机挑选了25个测试重量(单位：克),其样本方差为$S^2=4.25$.

在α=0.05下检验新品种是否比对照品种方差小？

解：$H_0: \sigma^2 \geq 7, \quad H_1: \sigma^2 < 7$

拒绝域：$\dfrac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\alpha}(n-1)$

查表得：$\chi^2_{0.95}\left(24\right) = 13.848,$

计算得：$\dfrac{\left(25-1\right)\times 4.25}{7} = 14.57 > 13.848$

不拒绝原假设，即认为新品种的方差并不比对照组的小。

例2 某厂生产的某种型号的电池，其寿命（以h计）长期以来服从方差$\boldsymbol{\sigma^2 = 5000}$的正态分布，现有一批这种电池，从它的生产情况来看，寿命的波动性有所改变.现随机取**26**只电池，测出其寿命的样本方差$\boldsymbol{s^2 = 9200}$.问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著变化（取$\boldsymbol{\alpha = 0.02}$）

练习 某炼铁厂铁水的含碳量X在正常情况下服从正态分布，现对操作工艺进行了某些改变，从中抽取7炉铁水的试样，测得含碳量数据如下：4.421, 4.052, 4.357, 4.394, 4.326, 4.287, 4.683. 问是否可以认为新工艺炼出的铁水含碳量的方差仍为0.112^2？（显著性水平为0.05）

§ 8.3.2 两个总体的情况

$$\mathrm{H}_0: \sigma_1^2 = \sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2$$ 设μ_1, μ_2未知

检验拒绝域的形式为： $\dfrac{S_1^2}{S_2^2} \le k_1$, 或 $\dfrac{S_1^2}{S_2^2} \ge k_2$

在原假设成立时， $\dfrac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$

$$P\{\text{当}H_0\text{为真拒绝}H_0\} = P_{\sigma_1^2=\sigma_2^2}\left\{\frac{S_1^2}{S_2^2} \le k_1, \text{或} \frac{S_1^2}{S_2^2} \ge k_2\right\} = \alpha$$

$$\text{取}P_{\sigma_1^2=\sigma_2^2}\left\{\frac{S_1^2}{S_2^2} \le k_1\right\} = \frac{\alpha}{2}, \quad P_{\sigma_1^2=\sigma_2^2}\left\{\frac{S_1^2}{S_2^2} \ge k_2\right\} = \frac{\alpha}{2}$$

因此，检验拒绝域为：

$$\boxed{\frac{S_1^2}{S_2^2} \le F_{1-\alpha/2}(n_1 - 1, n_2 - 1), \text{或} \frac{S_1^2}{S_2^2} \ge F_{\alpha/2}(n_1 - 1, n_2 - 1)}$$

左边检验 $H_0:\sigma_1^2 \geq \sigma_2^2, H_1:\sigma_1^2 < \sigma_2^2$

的检验拒绝域为： $F \leq F_{1-\alpha}(n_1-1, n_2-1)$

右边检验 $H_0:\sigma_1^2 \leq \sigma_2^2, H_1:\sigma_1^2 > \sigma_2^2$

的检验拒绝域为： $F \geq F_{\alpha}(n_1-1, n_2-1)$

例3 两台机床生产同一个型号的滚珠，从甲机床生产的滚珠中抽取8个，从乙机床生产的滚珠中抽取9个，测得这些滚珠的直径(毫米)如下：

甲机床	15.0	14.8	15.2	15.4	14.9	15.1	15.2	14.8	
乙机床	15.2	15.0	14.8	15.1	14.6	14.8	15.1	14.5	15.0

设两机床生产的滚珠直径分别为X,Y,且$X\sim N\left(\mu_1,\sigma_1^2\right),Y\sim N\left(\mu_2,\sigma_2^2\right)$

(1) 检验假设$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$($\alpha$=0.1)；

(2) 检验假设$H_0:\mu_1\le\mu_2, H_1:\mu_1>\mu_2$($\alpha$=0.1)；

(3)检验假设$H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2$($\alpha$=0.1)。

51

解：(1) 当μ_1,μ_2未知时，检验 $H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$

的拒绝域为：$\dfrac{S_1^2}{S_2^2}\le F_{1-\alpha/2}(n_1-1,n_2-1)$, 或 $\dfrac{S_1^2}{S_2^2}\ge F_{\alpha/2}(n_1-1,n_2-1)$

查表得：$F_{0.05}(7,8)=3.50, F_{0.95}(7,8)=\dfrac{1}{F_{0.05}(8,7)}=\dfrac{1}{3.73}=0.268$

本题中 $n_1=8,\ \bar{x}=15.05,\ S_1^2=0.0457$；$n_2=9,\ \bar{y}=14.9,\ S_2^2=0.0575$

计算得：$0.268<\dfrac{S_1^2}{S_2^2}=0.795<3.50$

不拒绝原假设，故认为方差没有显著差异。

52

$n_1 = 8,\ \overline{x} = 15.05,\ S_1^2 = 0.0457;\ n_2 = 9,\ \overline{y} = 14.9,\ S_2^2 = 0.0575$

$(2)\ H_0: \mu_1 \le \mu_2, H_1: \mu_1 > \mu_2$

的拒绝域为: $\dfrac{\overline{X} - \overline{Y}}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \ge t_\alpha(n_1 + n_2 - 2)$

$$t_{0.1}(15) = 1.3406, S_w = 0.228, \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.486$$

计算得: $\dfrac{\overline{X} - \overline{Y}}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = 1.354 > 1.3406$, 从而拒绝原假设。

53

$n_1=8,\ \overline{x}=15.05,\ S_1^2=0.0457;\ n_2=9,\ \overline{y}=14.9,\ S_2^2=0.0575$

(3) $H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2$

的拒绝域为：$\dfrac{\left|\overline{X}-\overline{Y}\right|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$

计算得：$\dfrac{\left|\overline{X}-\overline{Y}\right|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}=1.354<t_{0.05}(15)=1.7531$, 从而接受原假设。

练习 某洗衣粉包装机，在正常工作情况下，每袋标准质量为1000g，标准差不能超过15g. 假设每袋洗衣粉的净重服从正态分布. 某天为检查机器工作是否正常，从已装好的袋中，随机抽查10袋，测其净重（克）为

1020, 1030, 968, 994, 1014, 998, 976, 982, 950, 1048

问这天机器工作是否正常？（显著性水平为0.05）

§ 8.4 置信区间与假设检验

考虑单个正态总体方差已知时有关均值的统计推断.

设$X_1, X_2, \ldots, X_n$是来自总体$N(\mu, \sigma^2)$样本, σ^2已知.

μ的置信水平为$1-\alpha$的置信区间为

$$\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

假设检验问题 $H_0: \mu = \mu_0 \; H_1: \mu \neq \mu_0$,

显 著 性 水 平 为 α 的 检 验 拒 绝 域 为

$$W = \left\{ \frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} \geq z_{\alpha/2} \right\},$$

接受域为

$$\bar{W} = \left\{ \frac{|\bar{X} - \mu_0|}{\sigma / \sqrt{n}} < z_{\alpha/2} \right\}$$

$$= \left\{ \bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu_0 < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right\}$$

将接受域中的μ_0改写成μ时，所得结果正好是参数μ置信水平为$1-\alpha$的置信区间.

一般地，若假设检验问题　$H_0:\theta=\theta_0$ $H_1:\theta\neq\theta_0$的显著水平为α的接受域能等价地写成

$$\hat{\theta}_L<\theta_0<\hat{\theta}_U$$

那么（$\hat{\theta}_L$，$\hat{\theta}_U$）是参数θ的置信水平为$1-\alpha$的置信区间.

反之，若（$\hat{\theta}_L$，$\hat{\theta}_U$）是θ的置信水平为$1-\alpha$的置信区间，则当$\theta_0\in$（$\hat{\theta}_L$，$\hat{\theta}_U$）时，接受双边检验$H_0:\theta=\theta_0$，$H_1:\theta\neq\theta_0$中的原假设H_0，且检验的拒绝域为 $\theta_0\leq\hat{\theta}_L$或$\theta_0\geq\hat{\theta}_U$.

58

单侧置信限与单边假设检验的关系：

（1）若$\hat{\theta}_L$是θ的置信水平为$1-\alpha$的单侧置信下限，

则当$\theta_0 \ge \hat{\theta}_L$时，接受右边检验$H_0:\theta \le \theta_0$，　$H_1:\theta > \theta_0$

中的原假设H_0,反之，拒绝原假设.

（2）若$\hat{\theta}_U$是θ的置信水平为$1-\alpha$的单侧置信上限，

则当$\theta_0 \le \hat{\theta}_U$时，接受左边检验$H_0:\theta \ge \theta_0$，　$H_1:\theta < \theta_0$

中的原假设H_0,反之，拒绝原假设.

正态总体均值、方差的置信区间与假设检验

	待估参数	原假设	枢轴量	检验统计量	分 布	置信区间	拒绝域
一个正态总体	μ （σ^2已知）	$\mu=\mu_0$ （σ^2已知）	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{\sigma/\sqrt{n}}<z_{\alpha/2}$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{\sigma/\sqrt{n}}\ge z_{\alpha/2}$
	μ （σ^2未知）	$\mu=\mu_0$ （σ^2未知）	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t(n-1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{S/\sqrt{n}}<t_{\alpha/2}(n-1)$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{S/\sqrt{n}}\ge t_{\alpha/2}(n-1)$
	σ^2 (μ未知)	$\sigma^2=\sigma_0^2$ (μ未知)	$\dfrac{(n-1)S^2}{\sigma^2}$	$\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi^2_{1-\alpha/2}(n-1)<\dfrac{(n-1)S^2}{\sigma^2}<\chi^2_{\alpha/2}(n-1)$	$\dfrac{(n-1)S^2}{\sigma_0^2}\le\chi^2_{1-\alpha/2}(n-1)$ 或 $\dfrac{(n-1)S^2}{\sigma_0^2}\ge\chi^2_{\alpha/2}(n-1)$
两个正态总体	$\mu_1-\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\mu_1=\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t(n_1+n_2-2)$	$\dfrac{\lvert(\bar{X}-\bar{Y})-(\mu_1-\mu_2)\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}<t_{\alpha/2}(n_1+n_2-2)$	$\dfrac{\lvert\bar{X}-\bar{Y}\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\ge t_{\alpha/2}(n_1+n_2-2)$
	$\dfrac{\sigma_1^2}{\sigma_2^2}$	$\sigma_1^2=\sigma_2^2$	$\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}$	$\dfrac{S_1^2}{S_2^2}$	$F(n_1-1,n_2-1)$	$F_{1-\alpha/2}(n_1-1,n_2-1)<\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}<F_{\alpha/2}(n_1-1,n_2-1)$	$\dfrac{S_1^2}{S_2^2}\le F_{1-\alpha/2}(n_1-1,n_2-1)$ 或 $\dfrac{S_1^2}{S_2^2}\ge F_{\alpha/2}(n_1-1,n_2-1)$

§ 8.5 样本容量的选取

若C是参数θ的某检验问题的一个检验法，

$$\beta(\theta)=P_\theta(\text{接受}H_0)$$

称为检验法C的施行特征函数或OC函数，其图形称为OC曲线.

称 $1-\beta(\theta)=P_\theta(\text{拒绝}H_0)$ 为检验法C的**功效函数**。

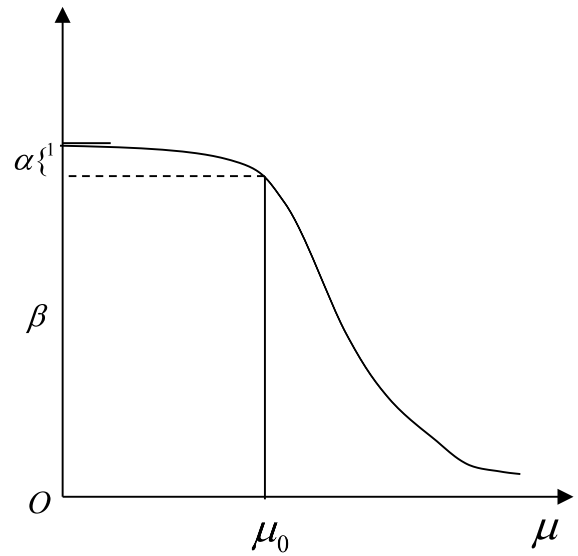
$$\beta(\theta)=\begin{cases}P_\theta(H_0\text{真时接受}H_0) & \theta\in H_0\\ P_\theta(\text{犯第}II\text{类错误}) & \theta\in H_1\end{cases}$$

$$1-\beta(\theta)=\begin{cases}P_\theta(\text{犯第}I\text{类错误}) & \theta\in H_0\\ P_\theta(H_0\text{不真时拒绝}H_0) & \theta\in H_1\end{cases}$$

§ 8.5.1 Z检验法的OC函数

右边检验

$$H_0 : \mu \le \mu_0, H_1 : \mu > \mu_0 \quad \text{的}OC\text{函数}$$



$$\beta(\mu) = P_\mu(\text{接受}H_0) = P_\mu\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} < z_\alpha\right)$$

$$= P_\mu\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right) = \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma/\sqrt{n}}\right)$$

$\beta(\mu)$的性质：

(1)它是$\lambda = \dfrac{\mu - \mu_0}{\sigma/\sqrt{n}}$的单调递减连续函数；

(2) $\lim\limits_{\mu \to \mu_0^+} \beta(\mu) = 1 - \alpha, \quad \lim\limits_{\mu \to \infty} \beta(\mu) = 0.$

给定$\beta(<1-\alpha)$, 无论n多大，无法满足$\beta(\mu)\le\beta$，对一切$\mu>\mu_0$

但可以确定n，使得$\beta(\mu)\le\beta$，对一切$\mu\ge\mu_0+\delta(\delta>0$取定）。

$\beta(\mu)$是μ的减函数，要使得对一切$\mu\ge\mu_0+\delta$，$\beta(\mu)\le\beta$。

只要 $\beta(\mu_0+\delta)=\Phi(z_\alpha-\sqrt{n}\delta/\sigma)\le\beta$

$$z_\alpha-\sqrt{n}\delta/\sigma\le -z_\beta, \Rightarrow \boxed{\sqrt{n}\ge\frac{(z_\alpha+z_\beta)\sigma}{\delta}}$$

左边检验

$H_0:\mu\ge\mu_0, H_1:\mu<\mu_0$ 的OC函数

$$\beta(\mu)=P_\mu(\text{接受}H_0)=P_\mu\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}>-z_\alpha\right)=1-\Phi\left(-z_\alpha-\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right)=\Phi\left(z_\alpha+\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right)$$

$\beta(\mu)$是μ的增函数，要使得对一切$\mu\le\mu_0-\delta$，$\beta(\mu)\le\beta$。

只要 $\beta(\mu_0-\delta)=\Phi(z_\alpha-\sqrt{n}\delta/\sigma)\le\beta$

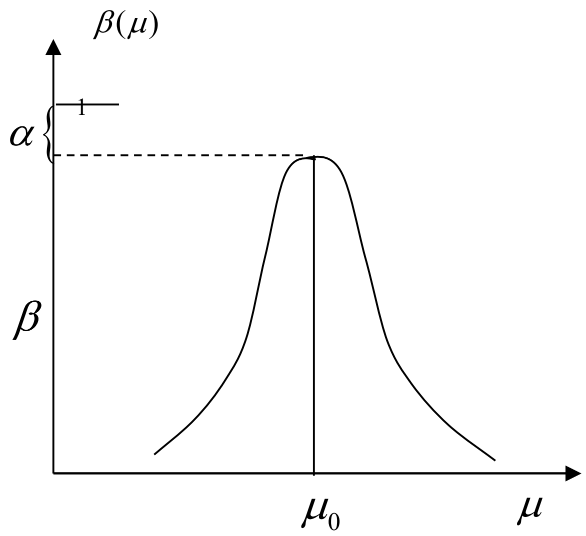
$$z_\alpha-\sqrt{n}\delta/\sigma\le -z_\beta,\Rightarrow \boxed{\sqrt{n}\ge\frac{(z_\alpha+z_\beta)\sigma}{\delta}}$$

双边检验 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$ 的OC函数

$$\beta(\mu) = P_\mu(\text{接受}H_0) = P_\mu\left(-z_{\alpha/2} < \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_{\alpha/2}\right)$$

$$= \Phi\left(z_{\alpha/2} - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right)$$

$$= \Phi\left(z_{\alpha/2} - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right) + \Phi\left(z_{\alpha/2} + \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right) - 1$$



$\beta(\mu)$是$|\mu-\mu_0|$的减函数，要使得对一切$|\mu-\mu_0| \geq \delta$，$\beta(\mu) \leq \beta$。

需解$\beta = \Phi(z_{\alpha/2} - \sqrt{n}\delta/\sigma) + \Phi(z_{\alpha/2} + \sqrt{n}\delta/\sigma) - 1$，确定$n$

因为n很大，可认为$z_{\alpha/2} + \sqrt{n}\delta/\sigma \geq 4$，即$\Phi(z_{\alpha/2} + \sqrt{n}\delta/\sigma) \approx 1$

$$\beta \approx \Phi(z_{\alpha/2} - \sqrt{n}\delta/\sigma) \Rightarrow \boxed{\sqrt{n} \geq \frac{(z_{\alpha/2} + z_\beta)\sigma}{\delta}}$$

例1（工业产品质量抽验方案）设有一大批产品，产品质量指标X服从$N(\mu,\sigma^2)$。以μ小者为佳，厂方要求所确定的验收方案对高质量的产品$(\mu \le \mu_0)$能以高概率1－α为买方所接受. 买方则要求低质产品$(\mu \ge \mu_0+\delta)$能以高概率1－β被拒绝。α，β有厂方与买方协商给出。并采取一次抽样以确定该批产品是否为买方所接受。问应怎样安排抽样方案。已知$\mu_0=120, \delta=20,$ 且由工厂长期经验知$\sigma^2=900.$ 经商定α=β=0. 05。

解：检验问题为 $H_0: \mu \le \mu_0, H_1: \mu > \mu_0$

且要求当 $\mu \ge \mu_0 + \delta$ 时，以概率 $1-\beta = 0.95$ 拒绝H_0。

由Z检验，拒绝域为 $\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \ge z_\alpha$,

$$OC\text{函数}\beta(\mu) = P_\mu\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_\alpha\right) = \Phi\left(z_\alpha - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right)$$

$$\sqrt{n} \ge \frac{(z_\alpha + z_\beta)\sigma}{\delta} \Rightarrow n \ge 24.35$$

取$n = 25.$ 当 $\bar{x}$ 满足 $\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \ge z_\alpha = 1.645$ 时，即

$\bar{x} \ge 129.87$时买方拒绝这批产品，当$\bar{x} < 129.87$时买方接受这批产品。

§ 8.5.2 t检验法的OC函数

右边检验

$$H_0: \mu \le \mu_0, H_1: \mu > \mu_0 \quad \text{的}OC\text{函数}$$

$$\beta(\mu) = P_\mu(\text{接受}H_0) = P_\mu\left\{\frac{\bar{X}-\mu_0}{S/\sqrt{n}} < t_\alpha(n-1)\right\}$$

这里$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} = \left(\dfrac{\bar{X}-\mu}{S/\sqrt{n}} + \dfrac{\mu-\mu_0}{S/\sqrt{n}}\right)\Bigg/\left(\dfrac{S}{\sigma}\right)$，服从自由度为$(n-1)$的非中心$t$分布。在$\mu=\mu_0$时即为通常的$t(n-1)$变量。

给定α, β及$\delta(>0)$，查 附表7 `得 n，

使得当$\dfrac{\mu-\mu_0}{\sigma} \ge \delta$时，犯第$II$类错误的概率不超过$\beta$。

左边检验

$$H_0: \mu \geq \mu_0, H_1: \mu < \mu_0 \quad t\text{检验法}$$

给定α, β及$\delta(>0)$，查 附表7 得 n,

使得当$\dfrac{\mu-\mu_0}{\sigma} \leq -\delta$时，犯第$II$类错误的概率不超过$\beta$。

双边检验

$$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0 \quad t\text{检验法}$$

给定α, β及$\delta(>0)$，查 附表7 得 n,

使得当$\dfrac{|\mu-\mu_0|}{\sigma} \geq \delta$时，犯第$II$类错误的概率不超过$\beta$。

例2 考虑在显著性水平$\alpha=0.05$下进行t检验：

$$H_0:\mu\le 68,\ \ H_1:\mu>68.$$

(1)要求在H_1中$\mu\ge\mu_1=68+\sigma$时犯第II类错误的概率不超过$\beta=0.05$.求样本容量n.

(2)若$n=30$,问在H_1中$\mu=\mu_1=68+0.75\sigma$时犯第II类错误的概率是多少？

解：（1）此处$\alpha = \beta = 0.05, \mu_0 = 68,$

$$\delta = \frac{\mu_1 - \mu_0}{\sigma} = \frac{68 + \sigma - 68}{\sigma} = 1，\text{查 附表7} \quad n = 13.$$

（2）此处$\alpha = 0.05, n = 30. \mu_0 = 68,$

$$\delta = \frac{\mu_1 - \mu_0}{\sigma} = \frac{68 + 0.75\sigma - 68}{\sigma} = 0.75，\text{查 附表7} \quad \beta = 0.01.$$

例3 考虑显著性水平$\alpha = 0.05$下进行t检验

$$H_0:\ \mu = 14, \qquad H_1:\ \mu \neq 14.$$

要求在H_1中$\frac{|\mu-14|}{\sigma} \geq 0.4$时犯第 II 类错误的概率不超过$\beta = 0.1$，求所需样本容量.

§8.6 分布拟合检验

前面介绍的各种检验都是在总体服从正态分布前提下，对参数进行假设检验的。实际中可能遇到这样的情形，总体服从何种理论分布并不知道，要求我们直接对总体分布提出一个假设 。

例如 要检验在计算机上产生随机数的一个程序。指令该程序产生0到9之间的100个单个数字。观察整数的频数如下表。那么以0.05的显著性水平，有充分的理由相信该批整数不是均匀产生的吗？

整数	0	1	2	3	4	5	6	7	8	9
频数	11	8	7	7	10	10	8	11	14	14

74

例如 从1500到1931年的432年间，每年爆发战争的次数可以看作一个随机变量，据统计，这432年间共爆发了299次战争，具体数据如下:

战争次数*X*	0	1	2	3	4
发生 *X*次战争的年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布假设？

§ 8.6.1 单个分布的χ^2拟合检验法

χ^2拟合检验法：在总体X的分布未知时，根据来自总体的样本，检验关于总体分布的假设的一种检验方法。

设总体X的分布未知，$x_1, x_2 \dots x_n$是来自X的样本值.
检验假设

$$H_0：\text{总体X的分布函数为} F(x)$$
$$H_1：\text{总体X的分布函数不是} F(x)$$

Pearson χ^2检验

1. 将在H_0下，总体X取值的全体分成k个两两不相交的子集$A_1,\ldots,A_k$.

2. 以$f_i(i=1,\ldots,k)$记样本观察值$x_1,\ldots,x_n$中落在A_i的个数（频数）. 则在n次试验中A_i发生的频率为 f_i/n.

3. 当H_0为真时，计算事件A_i发生的概率 $p_i=P_{H_0}(A_i), i=1,\ldots,k$.此时$np_i$称为理论频数.

4. 检验的拒绝域形式为： $\sum_{i=1}^{k} h_i(\frac{f_i}{n}-p_i)^2 \geq c.$

定理. 若n充分大$(n\geq 50)$，则当H_0为真时，统计量

$$\chi^2=\sum_{i=1}^{k}\frac{n}{p_i}\left(\frac{f_i}{n}-p_i\right)^2=\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}=\sum_{i=1}^{k}\frac{f_i^2}{np_i}-n$$

近似服从$\chi^2(k-1)$分布。因此检验拒绝域为

$$\chi^2\geq\chi_\alpha^{\ 2}(k-1).$$

即在显著性水平α下拒绝域为

$$\chi^2=\sum_{i=1}^{k}\frac{f_i^2}{np_i}-n\geq\chi_\alpha^{\ 2}(k-1)$$

注：使用时n不能小于50，另外np_i不能太小，应有$np_i\geq 5$，否则应适当合并A_i，以满足这个要求.

例1 下表列出了某一地区在夏季的一个月中由100个气象站报告的雷暴雨的次数.

i	0	1	2	3	4	5	≥ 6
f_i	22	37	20	13	6	2	0

其中f_i是报告雷暴雨次数i的气象站数.试用χ^2拟合检验法检验雷暴雨的次数X是否服从$\lambda = 1$的泊松分布（取显著性水平$\alpha = 0.05$）.

解：　按题意需检验假设

$$H_0: P\{X = i\} = \frac{\lambda^i e^{-\lambda}}{i!} = \frac{e^{-1}}{i!},\quad i=0,1,\ldots$$

在H_0下X所有可能取的值为Ω={0,1,2 ...}，将Ω分成如表所示的两两不相交的子集$A_0, A_1, A_2, A_3, A_4, A_5, A_6$，则有$P\{X = i\}$为

$$p_i = P\{X = i\} = \frac{e^{-1}}{i!},\quad i=0,1,\ldots 6$$

$$n = 100$$

A_i				${f_i}^2/(np_i)$
A_0: $\{X=0\}$	22	e^{-1}	36.788	13.16
A_1: $\{X=1\}$	37	e^{-1}	36.788	37.21
A_2: $\{X=2\}$	20	$e^{-1}/2$	18.394	21.75
A_3: $\{X=3\}$	13	$e^{-1}/6$	6.131	54.92
A_4: $\{X=4\}$	6	$e^{-1}/24$	1.533	
A_5: $\{X=5\}$	2	$e^{-1}/120$	0.307	
A_6: $\{X=6\}$	0	$1-\sum_{i=0}^{5} p_i$	0.059	

其中某些$np_i < 5$的组予以适当合并，使得每组均有$np_i \geq 5$.

$\chi^2_{0.05}(3)$=7.815，　χ^2=127.04-100>7.815

故在显著性水平$\alpha = 0.05$下拒绝H_0，认为样本不是来自均值$\lambda = 1$的泊松分布.

例2 在研究牛的毛色与牛角的有无，这样两对性状分离现象时，用黑色无角牛与红色有角牛杂交，子二代出现黑色无角牛192头，黑色有角牛78头，红色无角牛72头，红色有角牛18头，共360头，问这两对性状是否符合孟德尔遗传规律中的9:3:3:1的遗传比例？（取显著性水平$\alpha = 0.1$）

解：现将题中的数据列表如下

序号	1	2	3	4
种类	黑色无角	黑色有角	红色无角	红色有角
数量	192	78	72	18
A_i	A_1	A_2	A_3	A_4

以X记各种牛的序号，按题意需检验各类牛的头数符合比例9:3:3:1.需检验假设：

H_0: X的分布律为

X	1	2	3	4
p_k	9/16	3/16	3/16	1/16

$\chi^2_{0.1}(3)$=6.251， χ^2=363.37-360<$\chi^2_{0.1}(3)$

故接受H_0，认为两对性状符合孟德尔遗传规律中的9:3:3:1的遗传比例.

A_i	f_i	p_i	np_i	${f_i}^2/(np_i)$
A_1	192	9/16	202.5	182.04
A_2	78	3/16	67.5	90.13
A_3	72	3/16	67.5	76.8
A_4	18	1/16	22.5	14.4

例3 孟德尔遗传理论断言，当两个品种的豆杂交时，圆的和黄的、起皱的和黄的、圆的和绿的、起皱的和绿的豆的频数将以比例9：3：3：1发生。在检验这个理论时，孟德尔分别得到频数315、101、108、32、这些数据提供充分证据拒绝该理论吗？（取显著性水平$\alpha=0.05$）

解：现将题中的数据列表如下

序号	1	2	3	4
种类	圆&黄	皱&黄	圆&绿	皱&绿
数量	315	101	108	32
A_i	A_1	A_2	A_3	A_4

以X记各种豆的序号，按题意需检验各种豆的数量符合比例9:3:3:1.需检验假设：

H_0: X的分布律为

X	1	2	3	4
p_k	9/16	3/16	3/16	1/16

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
A_1	315	9/16	312.75	317.27
A_2	101	3/16	104.25	97.85
A_3	108	3/16	104.25	111.88
A_4	32	1/16	34.75	29.47

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^2}{np_i} - n = 0.47 < \chi^2_{0.05}(3) = 7.815,$$

因此没有充分的理由否定该理论.

§ 8.6.2 分布族的χ^2拟合检验

当H_0中所假设的X的分布函数$F(x)$含有r 个未知参数时，先用样本求出未知参数的最大似然估计，以估计值为参数值，求出p_i的估计值$\hat{p}_i = \hat{P}_{H_0}(A_i)$， 于是检验统计量为

$$\chi^2 = \sum_{i=1}^{k} \frac{f_i^2}{n\hat{p}_i} - n \quad \overset{\text{近似}}{\sim} \chi^2(k-r-1),$$

因此检验拒绝域为 $\chi^2 \geq \chi_\alpha{}^2(k-r-1)$

例4 从1500到1931年的432年间，每年爆发战争的次数可以看作一个随机变量，据统计，这432年间共爆发了299次战争，具体数据如下:

战争次数X	0	1	2	3	≥4
发生 X次战争的年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布假设？

解：$H_0 : X \sim \pi(\lambda)$，$\lambda$未知，$\hat{\lambda} = \bar{X} = 299/432 = 0.69.$

$$\hat{p}_i = \frac{\hat{\lambda}^i e^{-\hat{\lambda}}}{i!}, i = 0,1,2,3, \quad \hat{p}_4 = \sum_{j=4}^{\infty} \frac{\hat{\lambda}^j e^{-\hat{\lambda}}}{j!}.$$

战争次数x	0	1	2	3	≥ 4
实测频数 n_i	223	142	48	15	4
概率估计 $\hat{p}_i$	0.502	0.346	0.119	0.027	0.006
理论频数 $n\hat{p}_i$	217	149	51	12	3

$\underbrace{12 \quad 3}_{15}$

检验统计量的观察值为

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^{\ 2}}{n\hat{p}_i} - n = \frac{223^2}{217} + \frac{142^2}{149} + \frac{48^2}{51} + \frac{19^2}{15} - 432 = 1.74$$

即在显著性水平$\alpha = 0.05$下临界值

$$\chi_\alpha^{\ 2}(k - r - 1) = \chi_{0.05}^2(4\text{-}1\text{-}1) = 5.991$$

于是，$1.74 < 5.991$,不能拒绝原假设。

例5 下面列出了84个伊特拉斯坎(Etruscan)人男子的头颅的最大宽度(mm)，试检验这些数据是否来自正态总体（取α=0. 1）

141	148	132	138	154	142	150	146	155	158	150	140
147	148	144	150	149	145	149	158	143	141	144	144
126	140	144	142	141	140	145	135	147	146	141	136
140	146	142	137	148	154	137	139	143	140	131	143
141	149	148	135	148	152	143	144	141	143	147	146
150	132	142	142	143	153	149	146	149	138	142	149
142	137	134	144	146	147	140	142	140	137	152	145

92

解 需检验假设

$$H_0 : X \sim N(\mu, \sigma^2)$$

其中 μ ,σ²未知，其最大似然估计分别为

$\hat{\mu} = 143.8,\ \hat{\sigma}^2 = 6.0^2.$

计算每一事件A_i的概率估计值$\hat{p}_i = \hat{P}(A_i).$

例如

$$\hat{p}_1 = \hat{P}(A_1) = \hat{P}\{X \le 129.5\}$$

$$= \Phi\left(\frac{129.5 - 143.8}{6.0}\right)$$

$$= \Phi(-2.383) = 0.0087,$$

93

A_i		n_i	$\hat{p}_i$	$n\hat{p}_i$	$n_i^2/n\hat{p}_i$
A_1	x≤129.5	1	0.0087	0.73 5.09	4.91
A_2	129.5<x≤134.5	4	0.0519	4.36	
A_3	134.5<x≤139.5	10	0.1752	14.72	6.79
A_4	139.5<x≤144.5	33	0.3120	26.21	41.55
A_5	144.5<x≤149.5	24	0.2811	23.61	24.40
A_6	149.5<x≤154.5	9	0.1336	11.22 14.37	10.02
A_7	154.5<x<∞	3	0.0375	3.15	Σ=87.67

$$\chi^2 = 87.67 - 84 = 3.67$$

$$\chi^2_{0.1}(k-r-1) = \chi^2_{0.1}(2) = 4.605 > 3.67$$

故在水平0.1下接受H_0, 认为数据来自正态总体。

例6 在一实验中，每隔一定时间观察一次由某种铀所放射的到达计数器上α粒子数X，共观察了100次，得结果如下表：

i	0	1	2	3	4	5	6	7	8	9	10	11	≥ 12
f_i	1	5	16	17	26	11	9	9	2	1	2	1	0

其中，f_i是观察到有i个α粒子的次数.从理论上考虑知X应服从泊松分布

$$P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!} \qquad i=0,1,\ldots$$

问上式是否符合实际？(取$\alpha=0.05$)

例7 自1965年1月1日至1971年2月9日共2231天中，全世界记录到里氏震级4级和4级以上地震计162次，统计如下：

相继两次地震间隔天数	0~4	5~9	10~14	15~19	20~24	25~29	30~34	35~39	≥ 40
频数	50	31	26	17	10	8	6	6	8

试检验相继两次地震间隔的天数X服从指数分布（$\alpha = 0.05$）.

§ 8.7 假设检验问题的p值法

假设检验方法
- 临界值法
- p值检验法

定义 假设检验问题的p值(*probability value*)是由检验统计量的样本观察 值得出的原假设可被拒 绝的最小显著性水平.

例1 设总体$X \sim N(\mu,\sigma^2)$, μ未知, $\sigma^2 = 100$,
现有样本 $x_1, x_2, \cdots, x_{52}$，算得$\bar{x} = 62.75$.

现在来检验假设

$$H_0: \mu \le \mu_0 = 60, \quad H_1: \mu > 60.$$

采用Z检验法,检验统计量为

$$z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}.$$

以数据代入,得Z的观察值为 $z_0 = \dfrac{62.75 - 60}{10/\sqrt{52}} = 1.983.$

概率$P\{Z \ge z_0\} = P\{Z \ge 1.983\} = 1 - \Phi(1.983) = 0.0238.$

此概率称为Z检验法的右边检验的p值.

记为p值$=P\{Z \geq z_0\} =$ 0.0238

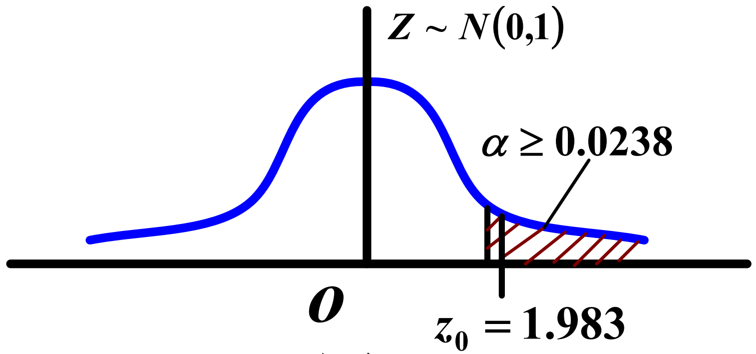


图1

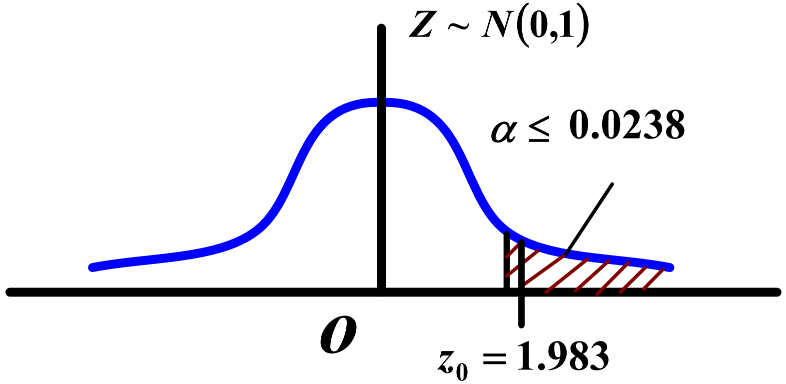


图2

若显著性水平$\alpha \geq p =$0.0238则对应的临界值$z_\alpha \leq 1.983$，这表示观察值$z_0 = 1.983$落在拒绝域内（如图1)，因而拒绝H_0；又显著性水平$\alpha < p = 0.0238$则对应的临界值$z_\alpha > 1.983$，这表示观察值 $z_0 = 1.983$不落在拒绝域内图(2)，因而接受H_0.

任一检验问题的p值可以根据检验统计量 的样本观察值的以及检验 统计量在H_0下一个特定的参数值(一般是 H_0与H_1所规定的参数的分界 点)对应的分布求出.

例如在正态分布$N(\mu,\sigma^2)$均值的检验中, 当σ未知时,可采用检验统计量

$t=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$,在以下三个检验问题中, 当$\mu=\mu_0$时,

$t\sim t(n-1)$.

如果由样本求得统计量 t的观察值为 t_0,那么在检验问题

$$H_0: \mu \le \mu_0, \quad H_1: \mu > \mu_0 \text{中}$$

$$p\text{值} = P_{\mu_0}\{t \ge t_0\} = t_0\text{右侧尾部面积, 如图3;}$$

$$H_0: \mu \ge \mu_0, \quad H_1: \mu < \mu_0 \text{中}$$

$$p\text{值} = P_{\mu_0}\{t \le t_0\} = t_0\text{左侧尾部面积, 如图4;}$$

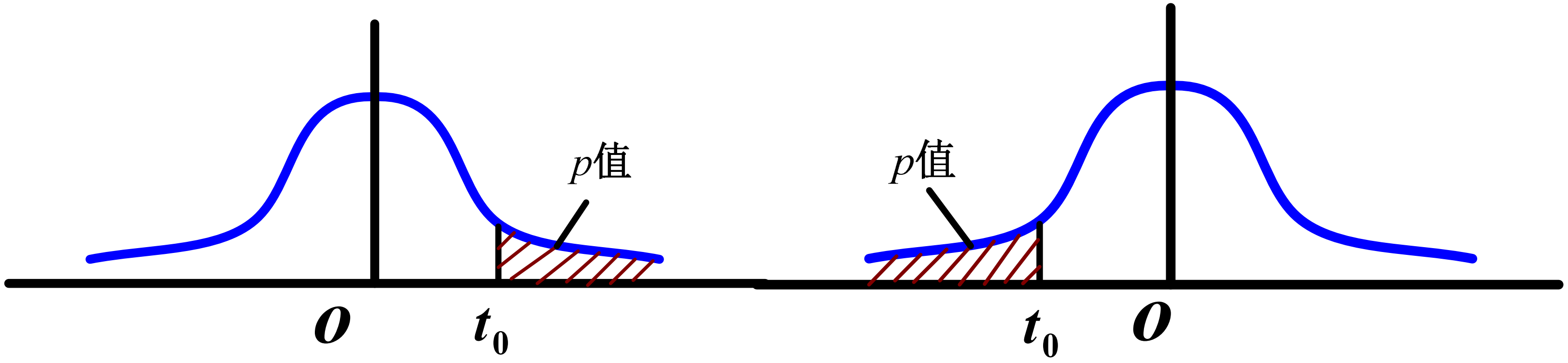


图3　　图4

$H_0:\mu=\mu_0,\quad H_1:\mu\neq\mu_0$中

(i)当$t_0>0$时 $\quad p$值 $=P_{\mu_0}\{|t|\geq t_0\}=P_{\mu_0}\{\{t\leq -t_0\}\cup\{t\geq t_0\}\}$

$=2\times$(t_0右侧尾部面积)如图5;

(ii)当$t_0<0$时 $\quad p$值 $=P_{\mu_0}\{|t|\geq -t_0\}=P_{\mu_0}\{\{t\leq t_0\}\cup\{t\geq -t_0\}\}$

$=2\times$(t_0左侧尾部面积)如图6;

综合(i)(ii),

p值 $=2\times$(由t_0界定的尾部面积)

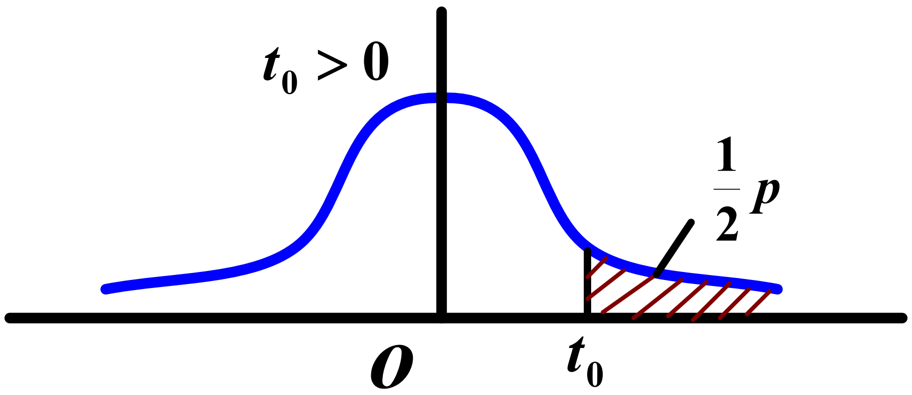


图5

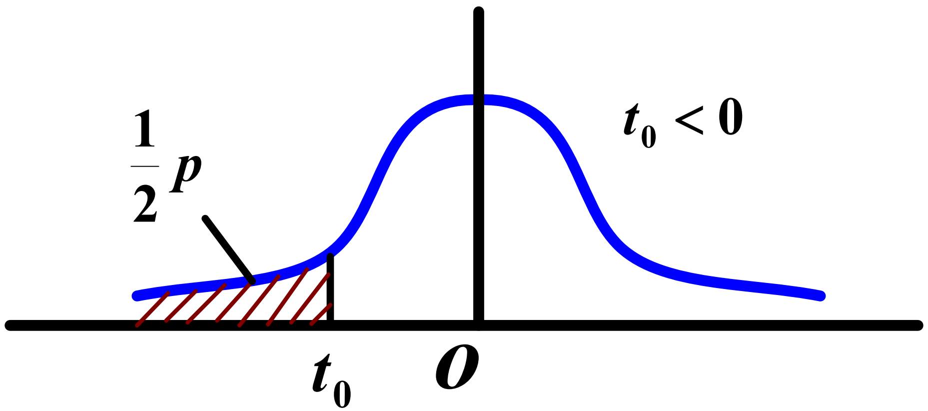


图6

上述各图中的曲线均为$t(n-1)$分布的概率密度曲线.
在现代计算机统计软件中, 一般都给出检验问题的
p值. 按p值的定义, 对于任意指定的显著性水平α,
就有

⑴ 若p值$\le \alpha$, 则在显著性水平α下拒绝H_0;

⑵ 若p值$> \alpha$, 则在显著性水平α下接受H_0.

有了这两条结论就能方便的去定H_0的拒绝域. 这种利用p值来确定检验拒绝域的方法, 称为p值检验法.

用临界值法来确定H_0的拒绝域时，例如当$\alpha = 0.05$时知道要拒绝H_0,再取$\alpha = 0.01$也要拒绝H_0，但不能知道将α再降低一些是否也要拒 绝H_0. 而p值法给出了拒绝 H_0的最小显著性水平 . 因此p值法比临界值法给出了有关拒绝域的更多的信息 .

例2 公司从生产商购买牛奶.公司怀疑生产商在牛奶中掺水以谋利.通过测定牛奶的冰点，可以检验出牛奶是否掺水.天然牛奶的冰点温度近似服从正态分布，均值$\boldsymbol{\mu_0} = -\mathbf{0.545}$℃，标准差 $\boldsymbol{\sigma} = \mathbf{0.008}$℃.牛奶掺水可使冰点温度升高而接近水的冰点温度(0 ℃).测得生产商提交的5批牛奶的冰点温度，其均值为$\bar{\boldsymbol{x}} = -\mathbf{0.535}$℃，问是否可以认为生产商在牛奶中掺了水？取$\boldsymbol{\alpha} = \mathbf{0.05}$.
用p值法检验该问题.

$$H_0: \mu \le \mu_0 = -0.545 \quad H_1: \quad \mu > \mu_0 \quad \alpha = 0.05$$

解 $H_0: \mu \le \mu_0 = -0.545 \quad H_1: \quad \mu > \mu_0 \quad \alpha = 0.05$

用z检验法，现在检验统计量$z = \dfrac{x - \mu_0}{\sigma/\sqrt{n}}$的观察值为

$$z = \frac{-0.535 - (-0.545)}{0.008/\sqrt{5}} = 2.7955.$$

p值＝$P\{Z \ge 2.9775\} = 1 - \Phi$（2.9775）＝0.0026.

p值$< \alpha = 0.05$，故拒绝H_0.

例3 某种元件的寿命X（以小时记）服从正态分布

均未知。现测得16只元件的寿命如下：

159	280	101	212	224	379	179	264
222	362	168	250	149	260	485	170

问是否有理由认为元件的平均寿命大于225（小时）？（取显著性水平为0.05）

用p值法检验该问题.

$$H_0: \mu \le \mu_0 = 225,\ H_1: \mu > 225,\ \alpha = 0.05.$$

解 用t检验法，现在检验统计量$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$的观察值为

$$t=\frac{241.5-225}{9837259/\sqrt{16}}=0.6685.$$

由计算机算得

$$p\text{值}=P\{t\geq 0.6685\}=0.2570.$$

$$p\text{值}>\alpha=0.05,\ \text{故接受}H_0.$$

例4 某厂生产的某种型号的电池，其寿命（以h计）长期以来服从方差$\sigma^2 = 5000$的正态分布，现有一批这种电池，从它的生产情况来看，寿命的波动性有所改变.现随机取26只电池，测出其寿命的样本方差$s^2 = 9200$.问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著变化（取$\alpha = 0.02$）

用p值法检验该问题.

解 用χ^2检验法，现在检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$

的观察值为

$$\chi^2=\frac{25\times 9200}{5000}=46.$$

由计算机算得

$$p\text{值}=2\times P\{\chi^2\geq 46\}=0.0128.$$

$$p\text{值}<\alpha=0.02,\ \text{故拒绝}H_0.$$

p值表示反对原假设H_0的依据的强度，p值越小，反对H_0的依据越强、越充分 (譬如对于某个检验问题的检验统计量的观察值的p值 $= 0.0009$，如此地小，以至于几乎不可能在 H_0为真时出现目前的观察值，这说明拒绝H_0的理由很强，我们就拒绝H_0.

一般, 若p值≤ 0.01,称推断拒绝H_0的依据很强或称检验是高度显著的;

若$0.01 < p$值≤ 0.05,称判断拒绝H_0的依据是强的或称检验是显著的;

若$0.05 < p$值≤ 0.1,称推断拒绝H_0的理由是弱的, 检验是不显著的;

若p值> 0.1, 一般来说没有理由拒绝.

基于p值,研究者可以使用任意希望的显著性水平来作计算.