第五章：假设检验

苏勤亮

中山大学计算机学院

超算中心503M

suqliang@mail.sysu.edu.cn

主要内容

- 假设检验介绍
- 正态总体参数的假设检验
- 似然比检验

假设检验实例

- 什么是假设检验?

根据观察到的样本对所提出的假设作出接受或拒绝的决定

例：某车间用一台包装机包装葡萄糖，包得的袋装糖重是一个随机变量，它服从正态分布。当机器正常时，其均值为0.5千克，标准差为0.015千克。某日开工后为检验包装机是否正常，随机地抽取它所包装的糖9袋，称得净重为(千克)：0.497　0.506　0.518　0.524　0.498　0.511　0.520　0.515　0.512．问机器是否正常？（长期实践可知，标准差是稳定的且$\sigma = 0.015$）

分析：由题意可知，每袋糖的重量X是一个服从正态分布$N(\mu, 0.015^2)$的随机变量，正常工作时$\mu = 0.5$

- 由题意可知

$$\text{机器是否正常} \Leftrightarrow \mu\text{是否等于}0.5$$

- 判断机器是否正常等价于决定接受如下两个对立假设中的哪一个

$$H_0:\ \mu = \mu_0 \ \text{和}\ H_1:\ \mu \neq \mu_0$$

其中，$\mu_0 \triangleq 0.5$

- 因此，判断机器是否正常工作问题转化为判断假设H_0是否成立

- 由于$\bar{X}$是总体均值μ的无偏估计量，因此可以根据$\bar{X}$与μ_0的距离$|\bar{X}-\mu_0|$来判断H_0成立与否的一个依据

- 衡量$|\bar{X}-\mu_0|$的大小等价于衡量$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}}$的大小。正常情况下（$H_0$成立），$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}}$的值应该很小。事实上，由由前面学习可知，若$H_0$成立，有下式成立

$$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

因此，若H_0成立，$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$服从标准正态分布，$\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|$取较大值是小概率事件

- 为判断差距是否够大，选定一个门限c使其满足 （α通常取0.01，0.05等）

$$P\left(\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} > c\right) = \alpha$$

1) 如果 $\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq c$, 则称$\bar{x}$与$\mu_0$的差异是显著的,则拒绝 H_0

2) 如果 $\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| < c$, 则称$\bar{x}$与$\mu_0$的差异是不显著的, 则接受 H_0

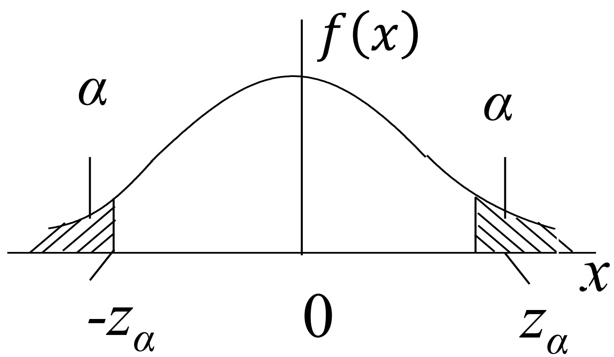
其中，α 称为显著性水平，$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$称为检验统计量

注意：上述$\bar{x}$与μ_0有无显著差异的判断是在显著性水平α之下作出的

α值与c的关系

设$X \sim N(0,1)$，对给定的α（$0<\alpha<1$）

如果满足条件$P\{X>z_\alpha\} = \alpha$，即$\int_{z_\alpha}^{+\infty} \varphi(x)\mathrm{d}x = \alpha$，那么$z_\alpha$表示$N(0,1)$分布的上$\alpha$分位点



由$P\left(\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} > c\right) = \alpha$可知，$c = z_{\alpha/2}$

- 因此，之前的结论可具体表示为

1) 如果$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$时，拒绝$H_0$

2) 如果$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2}$时，接受$H_0$

- 若取定$\alpha = 0.05$, 则 $c = z_{\alpha/2} = z_{0.025} = 1.96$，又已知$n = 9$，$\sigma = 0.015$

由给定样本可计算得$\bar{x} = 0.511$，可知得到

$$\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} = 2.2 > 1.96,$$

于是拒绝假设H_0，认为包装机工作不正常

上述检验方法的合理性

由于α取值很小，当H_0为真时，$\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$是一个小概率事件。因此，我们可以合理推断：如果$H_0$为真，不等式$\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$几乎是不会成立的

- 若某次试验观察到不等式 $\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$成立，我们有充分理由怀疑假设$H_0$的正确性，因而拒绝$H_0$，因为正常情况下，$\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$发生的可能性是非常低的
- 若出现观察值$\bar{x}$ 满足不等式 $\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\alpha/2}$, 则没有充分理由拒绝假设$H_0$, 因而只能接受$H_0$

假设检验的相关概念

- 假设检验问题通常表述为（以总体均值假设检验为例）

在显著性水平α下，检验如下两个假设哪个成立

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

1) 原假设与备择假设

H_0称为原假设或零假设, H_1 称为备择假设

2) 否定域（拒绝域）

当检验统计量在某个区域W中时, 拒绝原假设H_0, 则称区域W为否定域(拒绝域)

3) 两类错误

假设检验所作出的结论有可能是错误的，这种错误有两类

a) 当原假设H_0为真，观察值却落入否定域，作出了拒绝H_0的判断，称作第一类错误，又叫弃真错误，即：$P(拒绝H_0|H_0真)$

犯第一类错误的概率就等于显著性水平α，

$$P(拒绝H_0|H_0真) = \alpha$$

b) 当原假设H_0不真，而观察值却落入接受域作出了接受H_0的判断，称作第二类错误，即$P(接受H_0|H_0不真)$，常用β表示

真实情况 \ 所作判断	接受H_0	拒绝H_0
H_0为真	正确	第一类错误
H_0为假	第二类错误	正确

- 当样本容量n一定时, 若减少犯第一类错误的概率, 则犯第二类错误的概率往往增大

- 若要使犯两类错误的概率都减小，除非增加样本容量

4) 显著性风险

只对犯第一类错误的概率加以控制，而不考虑犯第二类错误的概率的检验，称为显著性检验

5) 关于原假设与备择假设的选取

在控制第一类错误概率为α的原则下，H_0得到特别的保护

因而, 通常把有把握的、有经验的结论作为原假设，需要去证明的结论放入备择假设

假设检验的一般步骤

1) 根据实际问题的要求, 提出原假设 H_0 及备择假设 H_1

2) 确定检验统计量以及否定域形式

3) 按$P\{拒绝H_0|H_0为真\}=\alpha$求出否定域

4) 根据样本观察值确定接受还是拒绝H_0

例：某厂生产的螺钉，按标准强度为$68N/mm^2$，而实际生产的强度X服$N(\mu, 3.6^2)$. 若$E(X)=\mu=68$，则认为这批螺钉符合要求，否则认为不符合要求．现从整批螺钉中取容量为36的样本, 其均值为$\bar{x}=68.5$，若取$\alpha=0.05$，问螺钉是否符合要求?

解：提出假设：

$$H_0：\mu=68，H_1：\mu\neq 68$$

若原假设正确，那么$X\sim N(68, 3.6^2)$，$\bar{X}=N(68, 3.6^2/36)$

因此，在原假设正确的前提下，$\frac{\bar{X}-68}{3.6/6}\sim N(0,1)$，观察到$\left|\frac{\bar{X}-68}{3.6/6}\right|$较大的可能性较小。由给的$\alpha$值确定常数$c$，使其满足

$$P\left(\left|\frac{\bar{X}-68}{3.6/6}\right|>c\right)=\alpha$$

由$\alpha = 0.05$，可知$c = z_{\alpha/2} = z_{0.025} = 1.96$

由 $\left|\frac{\bar{X}-68}{3.6/6}\right| > 1.96$ $\Longleftrightarrow$ $\bar{X} > 69.18$或 $\bar{X} < 66.824$

因此，可以得到

- 区间$(-\infty, 66.824)$与$(69.18, +\infty)$为检验的拒绝域
- 区间$(66.824, 69.18)$为检验的接受域

现$\bar{x} = 68.5$落入接受域, 则接受原假设$H_0: \mu = 68$

- 犯第一类错误的概率

$$P(\text{拒绝}H_0|H_0\text{为真}) = P(\bar{X} < 66.824 \cup \bar{X} > 69.18|\mu = 68)$$

$$= \alpha = 0.05$$

- 犯第二类错误的概率β

$$\beta = P(\text{接受}H_0|H_0\text{不真})$$

H_0不真，即$\mu \neq 68$，μ可能小于68，也可能大于68，β的大小取决于μ的真值的大小

如果$\mu = 66$，那么$\bar{X} \sim N(66,\ 3.6^2/36)$

$$\beta_{\mu=66} = P(66.82 \leq \bar{X} \leq 69.18\ |\mu = 66)$$

$$= \Phi\left(\frac{69.18-66}{0.6}\right) - \Phi\left(\frac{66.82-66}{0.6}\right)$$

$$= \Phi\ (5.3) - \Phi\ (1.37) = 1 - 0.9147 = 0.0853$$

若$\mu = 69$，那么$\bar{X} \sim N(69,\ 3.6^2/36)$

$$\beta_{\mu=69} = P(66.82 \leq \bar{X} \leq 69.18\ |\mu = 69)$$

$$= \Phi\left(\frac{69.18-69}{0.6}\right) - \Phi\left(\frac{66.82-69}{0.6}\right)$$

$$= \Phi\ (0.3) - \Phi\ (-3.63) = 0.6177$$

现增大样本容量，取$n = 64$

由于$\alpha = 0.05$, 则$c = z_{\alpha/2} = z_{0.025} = 1.96$

由$\left|\frac{\bar{X}-68}{3.6/\sqrt{64}}\right| > 1.96$，可以确定当$n = 64$时的拒绝域为

$$(-\infty, 67.118) \cup (68.882, +\infty)$$

接受域为

$$(67.118, 68.882)$$

当样本容量$n = 64$且$\mu = 66$时，$\bar{X} \sim N(66,\ 3.6^2/64)$。第二类错误可计算如下

$$\beta_{\mu=66} = P(\ 67.118 \le \bar{X} \le 68.882\ |\mu = 66)$$

$$\approx \Phi\left(\frac{68.88-66}{0.45}\right) - \Phi\left(\frac{67.12-66}{0.45}\right)$$

$$= \Phi\ (6.4) - \Phi\ (2.49)$$

$$\approx 1 - 0.9936 = 0.0064 < 0.0853$$

$$\beta_{\mu=69} = P(\ 67.12 \le \bar{X} \le 68.88\ |\mu = 69)$$

$$= 0.3936 < 0.6177$$

例：设$X_1, X_2, \cdots, X_n$是来自正态总体$N(\mu, 100)$的一个样本, 要检验 $H_0: \mu = 0$ vs $H_1: \mu \neq 0$，在下列两种情况下, 分别确定常数d，使得以W_1为否定域的检验犯第一类错误的概率为0.05

1） $n = 1$, $W_1 = \{x_1 \,||x_1| > d\}$

2） $n = 25$, $W_1 = \{(x_1, \cdots, x_{25}) \,||\bar{x}| > d\}$, 其中 $\bar{x} = \frac{1}{25}\sum_{i=1}^{25} x_i$

解：1） 当$n = 1$时，若H_0成立，那么$\frac{X_1}{10} \sim N(0,1)$

$$P(\text{犯第一类错误}) = P(|X_1| > d) = P\left(\left|\frac{X_1}{10}\right| > \frac{d}{10}\right) = 0.05$$

当H_0成立时，$\frac{X_1}{10} \sim N(0,1)$. 因此，由$P\left(\left|\frac{X_1}{10}\right| > \frac{d}{10}\right) = 0.05$可知

$$\Longrightarrow \frac{d}{10} = z_{0.025}，\text{因此，} \frac{d}{10} = 1.96，\text{即：} d = 19.6$$

2）$n=25$时，若H_0成立，则$\frac{\bar{X}}{10/\sqrt{25}} \sim N(0,1)$

$$P(\text{犯第一类错误}) = P(|\bar{X}| > d)$$

$$= P\left(\left|\frac{\bar{X}}{2}\right| > \frac{d}{2}\right) = 0.05$$

由$P\left(\left|\frac{\bar{X}}{2}\right| > \frac{d}{2}\right) = 0.05$可知 $\frac{d}{2} = z_{0.025}$

因此，$\frac{d}{2} = 1.96$，即：$d = 3.92$

可以看出，当$n=1$时，需要$\bar{x} > 19.6$才可以拒绝H_0，但当$n=25$时，只要$\bar{x} > 3.92$就可以拒绝H_0

主要内容

- 假设检验介绍
- 正态总体参数的假设检验
  1) 均值的假设检验
  2) 方差的假设检验
- 似然比检验

均值μ的检验——σ^2已知的情况

1） 假设$H_0: \mu = \mu_0$； $H_1: \mu \neq \mu_0$ （双边检验）

- 如果H_0为真时，可知统计量$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$，即：

$$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

- 由$P\left\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}\right\} = \alpha$可知，否定域为

$$|Z| \geq z_{\alpha/2}$$

如果统计量的观测值$|Z| = \left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$，则拒绝原假设；否则接受原假设；否则，接受原假设

Z检验法

2） 假设H_0: $\mu \le \mu_0$； H_1: $\mu > \mu_0$ （右边检测）

或 $H_0: \mu = \mu_0$； $H_1: \mu > \mu_0$

- 定义统计量

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

- 若H_0若成立，那么

  a） 真实均值等于μ，因此有$\tilde{Z} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

  b） 由于$\mu \le \mu_0$，有$Z \le \tilde{Z}$恒成立

- 因为$\tilde{Z} \sim N(0,1)$，有$P\{\tilde{Z} > z_\alpha\} = \alpha$。由于$Z \le \tilde{Z}$，可以知道

$$P\{Z > z_\alpha\} \le \alpha$$

即：如果H_0成立，事件$\{Z > z_\alpha\}$发生概率一定小于α

- 因此，对于右边假设检验H_0: $\mu \le \mu_0$；H_1: $\mu > \mu_0$，

检验统计量：$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ 拒绝域：$Z > z_\alpha$

> 这种方法可以保证第一类错误的概率一定小于α

3） 假设H_0: $\mu \ge \mu_0$； H_1: $\mu < \mu_0$ （左边检测）

或 H_0:$\mu = \mu_0$； H_1:$\mu < \mu_0$

类似的，在假设H_0: $\mu \ge \mu_0$成立的条件下，有$Z > \tilde{Z}$。因此，

$$P\{Z < -z_\alpha\} \le P\{\tilde{Z} < -z_\alpha\} = \alpha$$

因此，对于右边假设检验，检验统计量为

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

拒绝域为

$$Z < -z_\alpha$$

可以看出，无论是双边还是单边检测，检验统计量都是 $Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$，

唯一不同的是拒绝域的形式不同

均值μ的检验——σ^2未知的情况

1） 假设$H_0: \mu = \mu_0$； $H_1: \mu \neq \mu_0$ （双边检验）

由于σ^2未知，不能使用检验统计量$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ 来确定否定域了

由前面的学习可知，统计量$\frac{\bar{X}-\mu_0}{S/\sqrt{n}}$服从$t(n-1)$分布

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

- 因此，σ^2未知情况下的双边检测的检验统计量为

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

由t分布的分位数可知，

$$P\left\{\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)\right\} = \alpha$$

- 因此，可确定否定域为

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

如果统计量的观测值$|t| \geq t_{\alpha/2}\,(n-1)$，则拒绝原假设；否则，接受原假设

2) 假设 $H_0: \mu \le \mu_0$； $H_1: \mu > \mu_0$ 右边检测

或 $H_0: \mu = \mu_0$； $H_1: \mu > \mu_0$

当 $\mu = \mu_0$时，

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

由$P\{t \ge t_\alpha(n-1)\} = \alpha$，得显著性水平为$\alpha$的拒绝域为

$$t \ge t_\alpha(n-1)$$

3） 假设 $H_0: \mu \geq \mu_0$； $H_1: \mu < \mu_0$ 左边检测

或 $H_0: \mu = \mu_0$； $H_1: \mu < \mu_0$

当$\mu = \mu_0$时，

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

由$P\{t \leq -t_\alpha(n-1)\} = \alpha$，得显著性水平为$\alpha$的拒绝域为

$$t \leq -t_\alpha(n-1)$$

可以看出，对于σ^2未知的情形，无论是双边还是单边检测，检验统计量都是$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$，唯一不同的是拒绝域的形式不同

例：某切割机在正常工作时，切割每段金属棒的平均长度为10.5cm，标准差是0.15cm，今从一批产品中随机的抽取15段进行测量，其结果如下：

10.4	10.6	10.1	10.4	10.5	10.3	10.3	10.2
10.9	10.6	10.8	10.5	10.7	10.2	10.7	

假定切割的长度服从正态分布，且标准差没有变化，试问该机工作是否正常？（$\alpha = 0.1$）

解：由题意可知，$X \sim N(\mu, \sigma^2)$ 且 $\sigma = 0.15$

机器是否工作正常等价于检验假设

$$H_0: \mu = 10.5, \quad H_1: \mu \neq 10.5$$

由题$n=15$，$\bar{x}=10.48$，$\alpha=0.1$，可以得到

$$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{10.48-10.5}{0.15/\sqrt{15}}=-0.516$$

查表得$z_{0.05}=1.645$

由此可见，

$$\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|=0.516<z_{0.05}$$

不在拒绝域内，故接受 H_0，认为该机工作正常

例：如果在上例中只假定切割的长度服从正态分布、标准差σ未知，问该机切割的金属棒的平均长度有无显著变化？（$\alpha = 0.05$）

解：依题意$X \sim N(\mu, \sigma^2)$，且μ, σ^2均为未知

问题等价于要检验假设 $H_0: \mu = 10.5$，$H_1: \mu \neq 10.5$

由题目给定信息可知 $n = 15, \bar{x} = 10.48, s = 0.237, \alpha = 0.05$

由此可知，检验统计量

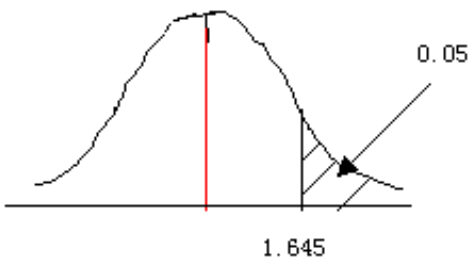
$$|t| = \left|\frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right| = \left|\frac{10.48 - 10.5}{0.237/\sqrt{15}}\right| = 0.327$$

因此， $|t| < t_{\alpha/2}(n-1) = t_{0.025}(14) = 2.1448$

不在拒绝域内，故接受H_0，认为金属棒的平均长度无显著变化

例：根据以往的资料，某厂生产的产品的使用寿命服从正态分布$N(1020, 100^2)$，并设方差稳定。现从最近生产的一批产品中随机抽取16件，测得样本平均寿命为1080小时。问这批产品的使用寿命是否有显著提高（显著性水平：0.05）？

解：提出假设：H_0：$\mu = 1020$，H_1：$\mu > 1020$



检验统计量：

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1080 - 1020}{100/\sqrt{16}} = 2.4$$

由于是右边检测且$\alpha = 0.05$，查表得临界值：$z_\alpha = z_{0.05} = 1.645$

由于$z > z_\alpha$，拒绝H_0，接受H_1，即这批产品的寿命确有提高

例：某厂生产镍合金线，其抗拉强度的均值为10620（kg/mm2）今改进工艺后生产一批镍合金线，抽取10根，测得抗拉强度(kg/mm2)为:

10512, 10623, 10668, 10554, 10776, 10707, 10557, 10581, 10666, 10670.

认为抗拉强度服从正态分布，取$\alpha = 0.05$，问新生产的镍合金线的抗拉强度是否比过去生产的合金线抗拉强度要高?

解： H_0：$\mu = 10620$；H_1：$\mu > 10620$

当$\mu = \mu_0$时：
$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$$

由前面结论可知，该假设检验的拒绝域为$t \geq t_{0.05}(9) = 1.8331$

由于$t = \frac{10631.4-10620}{81/\sqrt{10}} = 0.45 < 1.8331$，不在拒绝域内，所以接受$H_0$

例：设正品镍合金线的抗拉强度服从均值不低于10620（kg/mm2)的正态分布，今从某厂生产的镍合金线中抽取10根，测得平均抗拉强度10600(kg/mm2)，样本标准差为80．问该厂的镍合金线的抗拉强度是否不合格？（$\alpha = 0.1$）

解：H_0：$\mu \geq 10620$；H_1：$\mu < 10620$

当$\mu = 10620$时：

$$t = \frac{\bar{X} - 10620}{S/\sqrt{10}} \sim t(9)$$

由前面结论可知，该假设检验的拒绝域为$t \leq -t_{0.1}(9) = -1.383$

由于$t = \frac{10600-10620}{81/\sqrt{10}} = -0.79 > -1.383$，不在拒绝域内，所以接受$H_0$

主要内容

- 假设检验介绍
- 正态总体参数的假设检验
  1) 均值的假设检验
  2) 方差的假设检验
- 似然比检验

方差σ^2的检验——均值μ已知

- 假设 $H_0: \sigma^2 = \sigma_0^2$，$H_1: \sigma^2 \neq \sigma_0^2$；　　　（双边检验）

  - 当H_0成立时，由前面章节学习可知

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2\,(n)$$

因此，在μ已知的情况下，可以用如下检验统计量

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$$

- 由$P\{\chi^2 \le \chi^2_{1-\alpha/2}(n)\} = \frac{\alpha}{2}$和$P\{\chi^2 \ge \chi^2_{\alpha/2}(n)\} = \frac{\alpha}{2}$可知，拒绝域为

$$\chi^2 \ge \chi^2_{\alpha/2}(n) \text{ 或 } \chi^2 \le \chi^2_{1-\alpha/2}(n)$$

即：当统计量$\chi^2 \ge \chi^2_{\alpha/2}(n)$ 或 $\chi^2 \le \chi^2_{1-\alpha/2}(n)$，则拒绝原假设；否则接受原假设

方差σ^2的检验——均值μ未知

- 假设 $H_0: \sigma^2 = \sigma_0^2$，$H_1: \sigma^2 \neq \sigma_0^2$；（双边检验）
  - 这时我们不能采用统计量$\chi^2 = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}$，因为$\mu$未知
  - 当H_0为真时，由第六章知识可知$\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma_0^2} \sim \chi^2(n-1)$，即：

$$\frac{(n-1)S^2}{\sigma_0^2} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma_0^2} \sim \chi^2(n-1)$$

因此，在μ未知的情况下，可以用如下检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \text{ 或 } \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma_0^2}$$

- 由

$$P\{\chi^2 \le \chi^2_{1-\alpha/2}(n-1)\} = \frac{\alpha}{2}\text{和}P\{\chi^2 \ge \chi^2_{\alpha/2}(n-1)\} = \frac{\alpha}{2}$$

可知拒绝域为

$$\chi^2 \ge \chi^2_{\alpha/2}(n-1) \text{ 或 } \chi^2 \le \chi^2_{1-\alpha/2}(n-1)$$

即：当统计量$\chi^2 \ge \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 \le \chi^2_{1-\alpha/2}(n-1)$，则拒绝原假设；否则接受原假设

例：某厂生产的某种型号的电池，其寿命长期以来服从方差$\sigma^2 = 5000$(小时^2)的正态分布，现有一批这种电池，随机地取26只电池，测出其寿命的样本方差$s^2 = 9200$(小时^2)．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的变化？（设$\alpha = 0.02$）

解：要检验假设 $H_0: \sigma^2 = 5000$，$H_1: \sigma^2 \neq 5000$

检验统计量：

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

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{0.01}(25) = 44.314$$

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.99}(25) = 11.524$$

因此，可以知道拒绝域为

$$\frac{(n-1)s^2}{\sigma_0^2} \leq 11.524 \text{ 或 } \frac{(n-1)s^2}{\sigma_0^2} \geq 44.314$$

因为

$$\frac{(n-1)s^2}{{\sigma_0}^2} = \frac{25\times 9200}{5000} = 46 > 44.314$$

检验统计量的值落在拒绝域内，所以拒绝H_0

即：这批电池的寿命的波动性较以往的有显著的变化

主要内容

- 假设检验介绍
- 正态总体参数的假设检验
- 似然比检验

引言

- 似然比检验是Neymen和Pearson在1928年提出的构造假设检验的一般方法. 它在假设检验中的地位, 相当于极大似然估计在点估计中的地位. 它可视为 Fisher 的极大似然原理在假设检验问题中的体现.

- 由这种检验方法具有良好的性质.它对分布族的形式没有什么特殊的要求，适用面广

- 设有样本分布组$\mathcal{F} = \{f(x,\theta), \theta \in \Theta\}$，其中$\Theta$为参数空间. 令$\boldsymbol{X} = (X_1, \cdots, X_n)$为自 $\mathcal{F}$中抽取的$i.i.d.$样本，样本的概率函数$f(\boldsymbol{x},\theta)$若视为θ的函数，称为似然函数. 与极大似然估计相似，若$f(\boldsymbol{x},\theta_1) < f(\boldsymbol{x},\theta_2)$，则我们认为真参数为$\theta_2$的“似然性”比$\theta_1$的“似然性”大

- 由于假设检验在“$\theta \in \Theta_0$与$\theta \in \Theta_1$”这两者中选其一，我们自然考虑以下两个量

$$L_{\Theta_0}(\boldsymbol{x}) = \sup_{\theta \in \Theta_0} f(\boldsymbol{x},\theta), \quad L_{\Theta_1}(\boldsymbol{x}) = \sup_{\theta \in \Theta_1} f(\boldsymbol{x},\theta)$$

考虑其比值$\frac{L_{\Theta_1}(\boldsymbol{x})}{L_{\Theta_0}(\boldsymbol{x})}$，若比值较大，说明真参数在$\Theta_1$内的“似然性”较大，故倾向于否定假设“$\theta \in \Theta_0$”. 反之，若此比值较小，我们倾向于接受假设“$\theta \in \Theta_0$”

- 若记

$$\lambda(\boldsymbol{x}) = L_\Theta(\boldsymbol{x})/L_{\Theta_0}(\boldsymbol{x})$$

其中$L_\Theta(\boldsymbol{x}) = \sup_{\theta\in\Theta} f(\boldsymbol{x},\theta)$. 由于$\lambda(\boldsymbol{x})$与 $L_{\Theta_1}(\boldsymbol{x})/L_{\Theta_0}(\boldsymbol{x})$同增或同减，我们可以用$\lambda(\boldsymbol{x})$代替比值 $L_{\Theta_1}(\boldsymbol{x})/L_{\Theta_0}(\boldsymbol{x})$. 这样做的好处是 $L_\Theta(\boldsymbol{x}) = \sup_{\theta\in\Theta} f(\boldsymbol{x},\theta)$的计算比 $L_{\Theta_1}(\boldsymbol{x}) = \sup_{\theta\in\Theta_1} f(\boldsymbol{x},\theta)$ 要容易

似然比检验的定义

- 设样本$\boldsymbol{X}$有概率函数$f(\boldsymbol{x},\theta),\theta\in\Theta$，而$\Theta_0$为参数空间$\Theta$的真子集，对于检验问题$H_0:\theta\in\Theta_0\ vs\ H_1:\theta\in\Theta_1$，称统计量

$$\lambda(\boldsymbol{x})=\sup_{\theta\in\Theta} f(\boldsymbol{x},\theta)\Big/\sup_{\theta\in\Theta_0} f(\boldsymbol{x},\theta)$$

为关于该检验问题的似然比. 由下述定义的检验函数

$$\varphi(\boldsymbol{x})=\begin{cases}1, & \text{当}\lambda(\boldsymbol{x})>c\\ 0, & \text{当}\lambda(\boldsymbol{x})\le c\end{cases}$$

称为似然比检验（LRT），有些文献也称为广义似然比检验

- c为待定常数，通过选择常数c使得检验具有给定的水平α

似然比检验的步骤

- 四个步骤
  1) 求似然函数$f(\boldsymbol{x},\theta)$，指明参数空间Θ和Θ_0
  2) 算出$L_\Theta(\boldsymbol{x})=\sup\limits_{\theta\in\Theta} f(\boldsymbol{x},\theta)$和 $L_{\Theta_0}(\boldsymbol{x})=sup_{\theta\in\Theta_0} f(\boldsymbol{x},\theta)$
  3) 求出$\lambda(\boldsymbol{x})$或与其等价的统计量分布
  4) 确定c的值使得检验水平为指定的α

- 第三步中的$\lambda(\boldsymbol{x})$的分布很多时候难以获得，因而难以根据该分布来取得常数c的值

- 但若 $\lambda(\boldsymbol{x})$ 为某个统计量$T(\boldsymbol{x})$的严格单调递增函数，即：$\lambda(\boldsymbol{x}) = g\big(T(\boldsymbol{x})\big)$且$g(\cdot)$为严格单调递增函数，那么原始基于$\lambda(\boldsymbol{x})$的检验函数

$$\varphi(\boldsymbol{x}) = \begin{cases} 1, & \text{当}\lambda(\boldsymbol{x}) > c' \\ 0, & \text{当}\lambda(\boldsymbol{x}) \le c' \end{cases}$$

可以等价的写成关于统计量$T(\boldsymbol{x})$的形式

$$\varphi(\boldsymbol{x}) = \begin{cases} 1, & \text{当}T(\boldsymbol{x}) > c \\ 0, & \text{当}T(\boldsymbol{x}) \le c \end{cases}$$

- 这样只要求的统计量$T(\boldsymbol{x})$的分布，并根据该分布来确定常数c的值，使得检验水平等于指定的α

例：设$\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$是从正态分布族$\{N(\mu, \sigma^2), -\infty < \mu < +\infty, \sigma^2 > 0\}$中抽取的随机样本，求

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

水平为α的似然比检验（α和μ_0已知）

解：记$\boldsymbol{\theta} = (\mu, \sigma^2)$，则$\theta$的似然函数为

$$f(\boldsymbol{x}, \boldsymbol{\theta}) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}$$

参数空间Θ和Θ_0为

$$\Theta = \{\boldsymbol{\theta} = (\mu, \sigma^2): -\infty < \mu < +\infty, \sigma^2 > 0\}$$

$$\Theta_0 = \{\boldsymbol{\theta} = (\mu, \sigma^2): \mu = \mu_0, \sigma^2 > 0\}$$

在参数空间Θ上，μ和σ^2的极大似然估计为

$$\hat{\mu} = \bar{X}, \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

在参数空间Θ_0上，σ^2的极大似然估计为

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu_0)^2$$

因此，

$$L_{\Theta}(x) = \sup_{\boldsymbol{\theta}\in\Theta} f(x,\boldsymbol{\theta}) = f\left(\boldsymbol{x},\hat{\mu},\hat{\sigma}^{\mathbf{2}}\right) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}}\left(\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{-n/2}$$

$$L_{\Theta_0}(x) = \sup_{\boldsymbol{\theta}\in\Theta_0} f(x,\boldsymbol{\theta}) = f(\boldsymbol{x},\mu_0,\tilde{\sigma}^2) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}}\left(\sum_{i=1}^{n}(x_i - \mu_0)^2\right)^{-n/2}$$

从而有

$$\lambda(\boldsymbol{x}) = \left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \Big/ \sum_{i=1}^{n}(x_i - \mu_0)^2\right]^{-n/2} = \left[1 + n(x_i - \mu_0)^2 \Big/ \sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{n/2}$$

$$= \left(1 + \frac{1}{n-1}T^2\right)^{n/2}$$

其中$T = \frac{\sqrt{n}(\bar{X}-\mu_0)}{S}$，$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

由于$\lambda(\boldsymbol{x})$为$|T|$的严格递增函数，故检验的否定域

$$D=\{\boldsymbol{X}=(X_1,\cdots,X_n):\lambda(\boldsymbol{X})>c'\}=\{\boldsymbol{X}:|T|>c\}$$

由检验水平要求可知

$$P(|T|>c|H_0)=\alpha$$

而当H_0成立时，可知$T\sim t_{n-1}$. 因此，可知

$$c=t_{n-1}(\alpha/2)$$

因此，

$$\varphi(\boldsymbol{x})=\begin{cases}1, & \text{当}T>t_{n-1}(\alpha/2)\\ 0, & \text{当}T\le t_{n-1}(\alpha/2)\end{cases}$$

是检验水平为α的似然比检验

例：设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$为自均分布总体$U(0,\theta)$中抽取的随机样本，求

$$H_0:\theta\le\theta_0 \Leftrightarrow H_1:\theta>\theta_0$$

水平为α的似然比检验，其中α和θ_0给定.

解：此时似然函数为

$$f(\boldsymbol{x},\theta)=\begin{cases}\theta^{-n}, & 0<x_{(n)}<\theta,\\ 0, & \text{其他}.\end{cases}$$

参数空间$\Theta=(0,\infty),\Theta_0=(0,\theta_0]$. 由于$X_{(n)}=\max\{X_1,\cdots,X_n\}$为$\theta$的MLE，故有

$$L_\Theta(\boldsymbol{x})=\sup_{\theta\in\Theta}f(\boldsymbol{x},\theta)=\left(x_{(n)}\right)^{-n}$$

和

$$L_{\Theta_0}(\boldsymbol{x})=\sup_{\theta\in\Theta_0}f(\boldsymbol{x},\theta)=\begin{cases}L_\Theta(\boldsymbol{x}), & 0<x_{(n)}\le\theta_0,\\ 0, & x_{(n)}>\theta_0.\end{cases}$$

因此有

$$\lambda(\boldsymbol{x})=\begin{cases}1, & 0<x_{(n)}\le\theta_0,\\ \infty, & x_{(n)}>\theta_0\end{cases}$$

由于$\lambda(\boldsymbol{x})$为$T(\boldsymbol{x})=x_{(n)}$的非降函数，故检验的否定域为

$$D=\{\boldsymbol{X}=(X_1,\cdots,X_n)\colon X_{(n)}>c\}.$$

注：为使检验水平等于α，将集合$G=\{\boldsymbol{x}=(x_1,\cdots,x_n)\colon \lambda(\boldsymbol{x})=1\}$分为两部分$G_1=\{\boldsymbol{x}\colon c<x_{(n)}\le\theta_0\}$，$G_2=G-G_1$，则$G_1\cup\{\boldsymbol{x}\colon \lambda(\boldsymbol{x})=\infty\}=\{\boldsymbol{x}\colon x_{(n)}>c\}$为上述检验问题的否定域是合理的.

由于$T = X_{(n)}$的密度函数为$g(t) = \frac{nt^{n-1}}{\theta^n} \cdot I_{(0,\theta)}(t)$，故由

$$\alpha = P\left(X_{(n)} > c \middle| \theta = \theta_0\right) = \int_c^{\theta_0} \frac{nt^{n-1}}{\theta_0^n} \mathrm{d}t = 1 - \left(\frac{c}{\theta_0}\right)^n,$$

解出$c = \theta_0 \sqrt[n]{1-\alpha}$，故否定域为

$$D = \left\{\boldsymbol{X}: X_{(n)} > \theta_0 \sqrt[n]{1-\alpha}\right\}.$$

检验的功效函数为

$$\beta_\varphi(\theta) = P_\theta\left(X_{(n)} > \theta_0 \sqrt[n]{1-\alpha}\right) = \int_{\theta_0 \sqrt[n]{1-\alpha}}^{\theta} \frac{nt^{n-1}}{\theta^n} \mathrm{d}t$$

$$= \frac{1}{\theta^n}\left[\theta^n - \theta_0^n(1-\alpha)\right] = 1 - (1-\alpha)\left(\frac{\theta_0}{\theta}\right)^n$$

它是θ的单调增函数，故有

$$\beta_\varphi(\theta) \le \beta_\varphi(\theta_0), \qquad \theta \le \theta_0,$$

因此以D为否定域的检验水平为α. 因此

$$\varphi(\boldsymbol{x})=\begin{cases}1, & X_{(n)}>\theta_0\sqrt[n]{1-\alpha},\\ 0, & X_{(n)}\le\theta_0\sqrt[n]{1-\alpha}\end{cases}$$

为上述检验问题的水平为α的似然比检验.

例：设$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$取自下列指数分布总体，其密度函数为

$$f(x,\lambda)=\frac{1}{\lambda}\exp\left\{-\frac{x}{\lambda}\right\}I_{(0,\infty)}(x).$$

求检验问题

$$H_0:\lambda=\lambda_0 \quad \leftrightarrow \quad H_1:\lambda\neq\lambda_0$$

的检验水平为α的似然比检验，此处λ_0和α给定.

解：参数λ的似然函数为

$$L(\lambda,\boldsymbol{x})=\begin{cases}\lambda^{-n}\exp\{-n\bar{x}/\lambda\}, & 当x_1,\cdots,x_n>0,\\ 0, & 其他,\end{cases}$$

参数空间和H_0对应的参数空间的子集分别为$\Theta = (0, \infty)$和$\Theta_0 = \{\lambda: \lambda = \lambda_0\}$. 由于$\lambda$的极大似然估计为$\hat{\lambda} = \overline{X}$，故在$\Theta$和$\Theta_0$上似然函数的最大值分别为

$$L_\Theta(\boldsymbol{x}) = \sup_{\lambda \in \Theta} L(\lambda, \boldsymbol{x}) = \frac{e^{-n}}{\overline{x}^n},$$

$$L_{\Theta_0}(\boldsymbol{x}) = L(\lambda_0, \boldsymbol{x}) = \lambda_0^{-n} \exp\{-n\overline{x}/\lambda_0\}.$$

记$T = \sum_{i=1}^n x_i = n\overline{x}$，则似然比为

$$\lambda(T) = \frac{L_\Theta(\boldsymbol{x})}{L_{\Theta_0}(\boldsymbol{x})} = \frac{n^n \lambda_0^n}{e^n T^n} \exp\{T/\lambda_0\} = c \cdot g(T).$$

此处$c = \left(\frac{n\lambda_0}{e}\right)^n, g(T) = T^{-n} e^{T/\lambda_0}$. 显见当$T \to \infty$和$T \to 0$时$g(T) \to \infty$，$g(T)$的形状与下图类似.

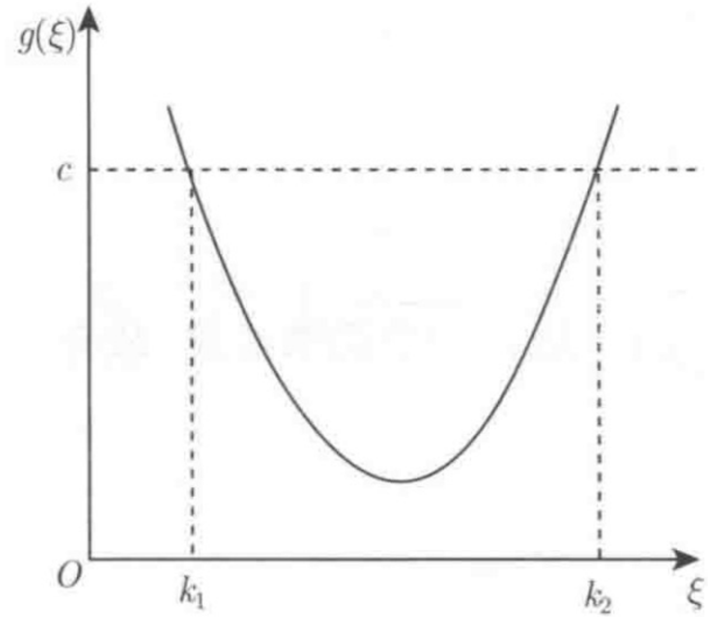
因此有

$$\varphi(\boldsymbol{x}) = \begin{cases} 1, & 当\lambda(T) > c \\ 0, & 当\lambda(T) \le c \end{cases} = \begin{cases} 1, & 当T < k_1或T > k_2, \\ 0, & 其他. \end{cases}$$

由推论2.4.5可知，当H_0成立时，$\frac{2T}{\lambda_0} \sim \chi^2_{2n}$. 为确定临界值$k_1$和$k_2$，令

$$P(T < k_1 | H_0) = P\left(\frac{2T}{\lambda_0} < \frac{2k_1}{\lambda_0}\right) = \frac{\alpha}{2},$$

$$P(T > k_2 | H_0) = P\left(\frac{2T}{\lambda_0} > \frac{2k_1}{\lambda_0}\right) = \frac{\alpha}{2},$$



得到

$$k_1 = \frac{\lambda_0}{2}\chi^2_{2n}\left(1 - \frac{\alpha}{2}\right), \qquad k_2 = \frac{\lambda_0}{2}\chi^2_{2n}\left(\frac{\alpha}{2}\right).$$

因此上述检验问题的检验水平为α的似然比检验为

$$\varphi(\boldsymbol{x})=\begin{cases}1, & \text{当}\lambda(T)<\dfrac{\lambda_0}{2}\chi_{2n}^2\left(1-\dfrac{\alpha}{2}\right)\text{或}\lambda(T)>\dfrac{\lambda_0}{2}\chi_{2n}^2\left(\dfrac{\alpha}{2}\right),\\ 0, & \text{其他}.\end{cases}$$

小结

假设检验的两类错误

真实情况 (未知)	所作决策	
	接受H_0	拒绝H_0
H_0为真	正确	犯第I类错误
H_0不真	犯第II类错误	正确

1. 关于μ的检验

Z检验法（σ^2已知）

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$\lvert Z \rvert \geq z_{\alpha/2}$
$\mu \geq \mu_0$	$\mu < \mu_0$		$Z \leq -z_{\alpha}$
$\mu \leq \mu_0$	$\mu > \mu_0$		$Z \geq z_{\alpha}$

t检验法（σ^2未知）

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$	$\lvert t \rvert \geq t_{\alpha/2}(n-1)$
$\mu \geq \mu_0$	$\mu < \mu_0$		$t \leq -t_{\alpha}(n-1)$
$\mu \leq \mu_0$	$\mu > \mu_0$		$t \geq t_{\alpha}(n-1)$

2. 关于σ^2的检验

χ^2检验法（μ已知）

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
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\chi^2 = \dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2} \sim \chi^2(n)$	$\chi^2 \le \chi^2_{1-\frac{\alpha}{2}}(n)$ 或 $\chi^2 \ge \chi^2_{\frac{\alpha}{2}}(n)$
$\sigma^2 \ge \sigma_0^2$	$\sigma^2 < \sigma_0^2$		$\chi^2 \le \chi^2_{1-\alpha}(n)$
$\sigma^2 \le \sigma_0^2$	$\sigma^2 > \sigma_0^2$		$\chi^2 \ge \chi^2_{\alpha}(n)$

χ^2检验法（μ未知）

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 $\chi^2\ge\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2\ge\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\le\chi^2_{1-\alpha}(n-1)$
$\sigma^2\le\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\ge\chi^2_{\alpha}(n-1)$