

\*\*\*\*\* 大学数学与信息科学学院学士  
学位论文

广义似然比检验原理及其应用

Generalized likelihood ratio test  
principle and its application

姓 名：\*\*\*\*\*

学 号：\*\*\*\*\*

学 院：数学与信息科学学院

专 业：统 计 学

指导老师：\*\*\*\*\* (教 授)

完成时间：\*\*\*\* 年\*\* 月\*\* 日

## 广义似然比检验原理及其应用

\*\*\*

**【摘要】** 广义似然比检验法是 Neyman和 Pearson 在 1928 年提出的，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。这是一个应用很广的方法，而且，它构造出的检验常具有种种最优性质。 本文首先叙述了假设检验的一些基本概念和思想，然后详细的阐述了广义似然比检验的原理， 分别对单个正态总体、两个正态总体、二项分布、指数分布等做了具体的推导过程，最后简单叙述了似然比检验的极限分布及广义似然比检验法的相合性。

**【关键词】** 广义似然比检验 正态总体 二项分布

## General likelihood ratio test principle and its application

\*\*\*\*\*

【 **abstract** 】 General likelihood ratio test was proposed by Neyman and Pearson in 1928 , whose position in the hypothesis test is equivalent to maximum likelihood estimate in the point estimation. This is a widely used method and it is the test constructed that often has the optimal properties. This paper firstly describes some basic concepts and ideas of the hypothesis test, then describe the generalized likelihood ratio test principle in detail, making detailed derivation of a single normal overall, two normal overall, binomial distribution, exponential distribution etc respectively, finally simply narrate the likelihood ratio test limit distribution and the consistency of general likelihood ratio test

【 **keywords** 】 General likelihood ratio test   Normal overall   Binomial distribution

## 目录

引言 .....	1
1 定义、原理 .....	1
2 Neyman--Pearson 原理 .....	3
3 广义似然比检验法在一些分布中的应用 .....	5
3.1 广义似然比检验法在单个正态总体的应用 .....	5
3.1.1 单个正态总体均值 .....	5
3.1.2 单个正态总体方差 .....	9
3.2 广义似然比检验法在两个正态总体的应用 .....	12
3.2.1 两个正态总体方差 .....	12
3.2.2 两个正态总体均值 .....	16
3.3 广义似然比检验法在二项分布的应用 .....	21
3.4 广义似然比检验法在指数分布的应用 .....	22
3.5 Stein 例子 .....	24
4 似然比检验的极限分布概略 .....	25
5 简述关于广义似然比检验的相合性 .....	26
6 关于广义似然比检验法的几点看法 .....	27
7 总结 .....	28
参考文献 .....	29

## 引言

广义似然比检验法是 Neyman 和 Pearson 在 1928 年提出的，他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。这是一个应用很广的方法，而且，它构造出的检验常具有种种最优性质。

Jerzy Neyman (1894 - 1981) 及 Egon sharpe Pearson (1895 — 1980) 在一系列的杰出的文章中澄清了推断理论，特别是有关显著性检验的基本原理——其合理性以往是常被批评。早期的显著性检验为关于二项变量之间或均值之间的，它们被 K. Pearson 推广至  $\chi^2$  检验，被 R.A. Fisher 推广到 F-检验，推广了 Student T-检验。Neyman 和 E.S. Pearson 看出，为了更有效，应该考虑与待检验的零假设相对应的备选假设。他们在这样的检验中设立两种误差并因素导致了他们的基本引理，似然比检验，及势的概念；他们顺便验证了大多数常见的显著性检验的应用；他们还引进了置信限；但是他们的体系从未被 Fisher 所承认。

本文通过首先叙述了假设检验的一些基本概念和思想后，详细的阐述了广义似然比检验的原理上，分别对单个正态总体、两个正态总体、二项分布、指数分布等做了具体的推导过程，最后简单叙述了似然比检验的极限分布及广义似然比检验法的相合性。

## 1 定义、原理

假设检验的基本概念及思想。

原假设、备择假设

令  $(X, A, P)$  是一个统计空间，其中， $P = (p_\theta : \theta \in \Theta)$ ， $\Theta$  是参数空间。令  $\Theta_0$  是  $\Theta$  的一个真子集，我们关心参数  $\theta$  是否属于  $\Theta_0$ 。假设检验的基本思路是，先作一个假设

$$H_0 : \theta \in \Theta_0 \quad (1.1)$$

然后通过样本构造统计量，借助统计量来判断这个假设是否正确。通常称  $H_0$  (1.1) 是零假设或原假设，把  $H_a : \theta \in \Theta - \Theta_0$  叫做对立假设（或叫备择假设），称用来判断假设是否正确的统计量为检验统计量。

检验法：

给出一个规则，通过已知的样本值  $x_1, \dots, x_n$  进行明确的表态，是接受假设  $H_0$  还是拒绝  $H_0$ 。

否定域：

设  $S$  是所有可能的样本值  $(x_1, \dots, x_n)$  ( $n$  固定) 组成的集合（样本空间），空间  $S$  的一个划分： $S = S_1 \cup S_2$  ( $S_1$  与  $S_2$  不相交)。当  $(x_1, \dots, x_n) \in S_1$  时接受  $H_0$ ，当  $(x_1, \dots, x_n) \in S_2$  时拒绝  $H_0$ ，这  $S_1$  叫接受域， $S_2$  叫否定域。

因为，故只要知道了否定域，就知道了检验法。每一个检验法对应一个否定域；反之，任给定  $S$  的一个子集  $W$ ，则有一个检验法以  $W$  作为它的否定域。故研究检验法就相当于研究否定域。

关于两类错误

零假设  $H_0$  在客观上只有两种可能性：真、假。样本值  $(x_1, \dots, x_n)$  也只有两种可能性：属于否定域  $W$  不属于  $W$ 。若采用  $W$  作否定域，则在观察到样本值  $(x_1, \dots, x_n)$  时只可能有下列四种情况：

(1)  $H_0$  真，而属于  $W$ ;

(2)  $H_0$  真, 而不属于  $W$ ;

(3)  $H_0$  假, 而属于  $W$ ;

(4)  $H_0$  假, 而不属于  $W$

根据规则, 在情形 (1) 应拒绝  $H_0$ , 在情形 (2) 应接受  $H_0$ , 在情形 (3) 应拒绝  $H_0$ , 在情形 (4) 应接受  $H_0$ , 在情形 (2) 及 (3) 对  $H_0$  的表态与客观事实相符。但 (1) 及 (4) 两种情形下, 表态犯了错误: 与客观事实不相符。

第一类错误: 情形 (1) 中, 在原假设为真的情形下, 样本值落入否定域。这个事件发生的概率为犯第一类错误的概率, 记为:

$$\begin{aligned} P_w(\theta) &= P(\text{拒绝 } H_0 | \theta) \\ &= P((X_1, \dots, X_n) \in W | \theta) \quad (\theta \in \bar{\Theta}_0) \end{aligned}$$

这里用  $P(A|\theta)$  表示  $X$  的分布的参数的真值  $\theta$  是时事件  $A$  发生的概率 (或用记号

$P_\theta(A)$  表示)。这里  $P_w(\theta)$  叫做  $W$  的功效函数。

第二类错误: 情形 (4) 中, 在原假设为假的情形下, 样本值落入接受域。这个事件发生的概率为犯第二类错误的概率。

在检验中, 我们自然希望找到犯两类错误的概率都很小, 这是在做假设检验中一直围绕的中心。

检验水平  $\alpha$ :

在原假设下为真的情形下, 样本值落入否定域, 这事件发生的最大概率, 记为:  $\sup P_w(\theta)$

检验问题的表示

设总体分布函数  $f(x, \theta)$ , 其中  $\theta \in \bar{\Theta}$ , 这里  $\bar{\Theta}$  为参数空间, 记假设检验问题

$$H_0: \theta \in \bar{\Theta}_0 \leftrightarrow H_a: \theta \in \bar{\Theta}_1$$

这里  $H_0$  是待检验假设,  $H_a$  是备择假设,  $\bar{\Theta}_0$  与  $\bar{\Theta}_1$  不相交, 在很多情形下,  $\bar{\Theta}_1 = \bar{\Theta} - \bar{\Theta}_0$ 。

定义:

一致最大功效否定域:

若  $W$  的水平为  $\alpha$  而且对而且对一切水平不超过  $\alpha$  的否定域  $\tilde{W}$  均有

$$P_W(\theta) \geq P_{\tilde{W}}(\theta) \quad (\text{一切 } \theta \in \bar{\Theta}_1)$$

则称  $W$  是检验水平为  $\alpha$  的一致最大功效否定域。

无偏否定域：

若对一  $\theta \in \bar{\Theta}_1$  切，有

$$P_W(\theta) \geq \alpha$$

则称  $W$  是检验水平为  $\alpha$  的无偏否定域。

一致最大功效无偏否定域：

若  $W$  是水平为  $\alpha$  的无偏否定域，而且对任何水平为  $\alpha$  的无偏否定域  $\tilde{W}$  恒有

$$P_W(\theta) \geq P_{\tilde{W}}(\theta) \quad (\text{一切 } \theta \in \bar{\Theta}_1)$$

则称  $W$  是检验水平为  $\alpha$  的一致最大功效无偏否定域。

下面在正态总体中推导出来的都是一致最大功效无偏否定域。

基本原理及思想：

通常认为在概率很小的随机事件中一次试验中几乎是不可能发生的（即小概率原理），也就是说通过概率大小来判断  $\theta^2$  和  $\theta_0$  的差异。一般情况下，样本值越多则检验得出的结论信赖度越大，这与大数定理是一致的。

## 2 Neyman--Pearson 原理

广义似然比检验法是 Neyman 和 Pearson 在 1928 年提出的，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。这是一个应用很广的方法，而且，它构造出的检验常具有种种最优性质。当样本量  $n$  较大时，广义似然比检验否定域往往相当好（即第二类错误的概率比较小），虽然不一定是一致的最大功效的。多数情况下可以证明  $\lim_n P_{W_0}(\theta) = 1$ （对一切  $\theta \in \bar{\Theta}_1$ ），这里  $\theta \in \bar{\Theta}_1$  是备

择假设， $P_{W_0}(\theta)$  是参数真值是  $\theta$  时检验法拒绝零假设的概率。

设  $X$  的分布密度是  $f(x, \theta)$ ,  $\theta = (\theta_1, \dots, \theta_m) \in \bar{\Theta} \subset R^m$ （对离散型分布可进行类似的讨论）。

设  $\bar{\Theta}_0$  是  $\bar{\Theta}$  的非空真子集，研究检验问题：

$$H_0: \theta \in \bar{\Theta}_0 \leftrightarrow H_a: \theta \in \bar{\Theta} - \bar{\Theta}_0$$



设  $\tilde{x} = (x_1 \cdots x_n)$  是样本  $X = (X_1 \cdots X_n)$  的值。似然函数  $L(\tilde{x}, \theta) = \prod_{i=1}^n f(x_i, \theta)$  ,  
令

$$L(\tilde{\theta}_0) = \sup_{\theta \in \Theta} L(\tilde{x}, \theta)$$

$$L(\tilde{\theta}) = \sup_{\theta \in \Theta} L(\tilde{x}, \theta)$$

定义：称

$$\lambda(\tilde{x}) = \frac{L(\tilde{\theta})}{L(\tilde{\theta}_0)} \quad (2.1)$$

为样本值  $\tilde{x} = (x_1 \cdots x_n)$  的广义似然比。

显然  $\lambda(\tilde{x}) \geq 1$  , 直观上看, 若  $H_0$  是真的, 则  $\theta$  的最大似然估计应该很大可能属于  $\bar{\Theta}_0$  , 从而  $\lambda(\tilde{x})$  应该接近于 1 ; 反之, 若  $\lambda(\tilde{x})$  的值太大就应该否定  $H_0$  这样应取否定域

$$W_0 = \{\tilde{x} : \lambda(\tilde{x}) > \lambda_0\} \quad 1 \leq \lambda(\tilde{x}) \quad \theta \rightarrow \bar{\theta}_0 \quad n \rightarrow \infty \quad \lambda(\tilde{x}) \rightarrow 1$$

( 2.2 )

其中  $\lambda$  满足

$$\sup_{\theta \in \bar{\Theta}_0} P(X \in W_0 | \theta) = a \quad (2.3)$$

这里  $a$  是预先给定的检验水平 (  $0 < a < 1$  )  
从上面的讨论可得到如下结果：

$$1 \leq \lambda(\tilde{x}) ,$$

当  $\theta$  距  $\bar{\theta}_0$  较远时  $\lambda(\tilde{x})$  较大,  $\theta$  距  $\bar{\theta}_0$  较近时  $\lambda(\tilde{x})$  较小,

当  $H_0 : \theta \in \bar{\Theta}_0 \leftrightarrow H_a : \theta \in \bar{\Theta} - \bar{\Theta}_0$  为真, 则但  $n \rightarrow \infty$  时,  $\lambda(\tilde{x}) \rightarrow 1$ 。

这样的检验法叫做广义似然比检验法, 简称似然比检验法。这个方法在许多情形下常可导出有实用价值的具体否定域。 在使用广义似然比检验法时, 关键在于能否找出  $\lambda_0$  满足 ( 2.3 )

我们指出, 广义似然比  $\lambda$  是充分统计量函数, 从而否定域 ( 2.2 ) 常常是由充分统计量来确定, 实际上, 若  $\varphi(X)$  是充分统计量 ( 可以是向量 ) , 则有

$$L(\tilde{x}, \theta) = g[\varphi(\tilde{x}), \theta] \cdot h(\tilde{x}) , 从而$$

$$\lambda(x) = \frac{\sup_{\theta \in \Theta} g[\varphi(x), \theta]}{\sup_{\theta \in \Theta} g[\varphi(x), \theta]} = I[\varphi(x)]$$

所以否定域

$$W_0 = \{x: \lambda(x) > \lambda_0\} = \{x: \varphi(x) \in B\}$$

这里 B 是适当的集合。

### 3 广义似然比检验法在一些分布中的应用

由于正态总体在实际工作中最常见，数学处理上也比较成熟，下面首先讨论广义似然比检验法在正态总体的各种检验问题，其次讨论二项分布、指数分布及 Stein 例子。

#### 3.1 广义似然比检验法在单个正态总体的应用

##### 3.1.1 单个正态总体均值

设  $X \sim N(\mu, \sigma^2)$  已知， $\mu \in \Theta = (-\infty, \infty)$ ，研究下列检验问题：

$$H_0: \mu = \mu_0 \Leftrightarrow H_a: \mu \neq \mu_0$$

给定  $\alpha = 0.05$ ，样本是  $X_1, \dots, X_n$ ，如何检验  $H_0$ ？

现利用广义似然比检验法推导否定域：

求  $x$  的似然函数

$$\begin{aligned} L(x, \mu) &= \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2 \right] \right) \end{aligned}$$

得到

$$L(\mathcal{G}_0) = L(x, \mu_0)$$

$$L(\mathcal{G}) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right)$$

知广义似然比

$$\lambda = e^{\frac{n}{2\sigma^2}(\bar{x} - \mu_0)^2},$$

故否定域

$$W_0 = \{x: \lambda(x) > \lambda_0\} = \{x: |\bar{x} - \mu_0| > C\} \quad (3.1)$$

其中 C 满足

$$P(|\bar{X} - \mu_0| > C | \mu_0) = 0.05$$

由于  $\mu = \mu_0$  时,  $\frac{\sqrt{n}}{\sigma}(|\bar{X} - \mu_0|) : N(0,1)$ , 因此利用总体分布表知  $C = \frac{1.96}{\sqrt{n}}\sigma$ , 这样

就可以计算出否定域 (3.1)。

数学上可以证明, 否定域 (3.1) 是一致最大功效无偏的。

设  $X \sim N(\mu, \sigma^2)$ ,  $\Theta = \{(\mu, \sigma^2) : -\infty < \mu < \infty, 0 < \sigma^2 < \infty\}$ , 方差  $\sigma^2$  未知,

记  $\Theta_0 = \{(\mu, \sigma^2) : \sigma^2 > 0\}$ , 设  $\tilde{x} = (x_1 \cdots x_n)$  是样本  $X = (X_1 \cdots X_n)$  的值。

研究下列检验问题:

$$H_0: \mu = \mu_0 \Leftrightarrow H_a: \mu \neq \mu_0$$

现利用广义似然比检验法推导否定域:

求  $\tilde{x}$  的似然函数

$$\begin{aligned} L(\tilde{x}, \theta) &= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + nx(\bar{x} - \mu)^2 \right] \right) \end{aligned}$$

得到

$$\begin{aligned} L(\Theta) &= \sup_{\theta} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \\ L(\Theta_0) &= \sup_{\theta \in \Theta_0} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \mu_0)^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \end{aligned}$$

知广义似然比

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

其中

$$T = \frac{\sqrt{n(n-1)}(\bar{x} - \mu_0)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

显然  $\lambda$  是  $T^2$  的严格增函数，故广义似然比检验的否定域为：

$$W_0 = \{(x_1, \dots, x_n) : |T| > C\} \quad (3.2)$$

其中  $C$  满足

$$P(\tilde{X} \in W_0 | (\mu_0, \sigma)) = \alpha \quad (\text{任意 } \sigma)$$

在  $\mu = \mu_0$  时，统计量  $T$  服从  $n-1$  个自由度的  $t$  分布，因此 (3.2) 中的  $C$  可从  $t$  分布表中查出。

由数学上可以证明，否定域 (3.2) 是一致最大功效无偏的。（参看文献 [8]）

设  $\tilde{X} \sim N(\mu, \sigma^2)$ ， $\sigma^2$  未知， $x_1, \dots, x_n$  是样本  $\tilde{X} = (X_1, \dots, X_n)$  的值。

研究下列检验问题：

$$H_0 : \mu \leq \mu_0 \Leftrightarrow H_a : \mu > \mu_0$$

现在用广义似然比检验法推导否定域：

求  $\tilde{x}$  的似然函数

$$\begin{aligned} L(\tilde{x}, \theta) &= \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + nx(\bar{x} - \mu)^2 \right] \right) \end{aligned}$$

得到

$$L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}}$$

$$L(\theta_0) = \sup_{\theta \in \Theta} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \mu_0)^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}}$$

知广义似然比

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

其中

$$T = \frac{\sqrt{n(n-1)}(\bar{x} - \mu_0)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (3.3)$$

显然  $\lambda$  是  $T^2$  的严格增函数，故否定域

$$W_0 = \{(x_1, \dots, x_n) : T > C_1\} \quad (3.4)$$

$C_1$  由方程

$$\int_{C_1}^{\infty} t_{n-1}(y) dy = a$$

确定。上式的  $t_{n-1}(y)$  是  $n-1$  个自由度的  $t$  分布的分布密度，则  $C_1$  可以通过  $t$  分布表中查出。

否定域 (3.4) 在直观上是合理的，数学上也可证明，它是一致最大功效无偏的。相类似的，研究下列检验问题：

$$H_0 : \mu \geq \mu_0 \Leftrightarrow H_a : \mu < \mu_0$$

也可用广义似然比类似的推导否定域：

$$W_0 = \{(x_1, \dots, x_n) : T < C_2\} \quad (3.5)$$

其中  $T$  的表达式参见 (3.3)， $C_2$  满足方程

$$\int_{C_2}^{\infty} t_{n-1}(y) dy = 1 - a$$

$C_2$  可以通过  $t$  分布表中查出。

数学上也可证明，否定域 (3.5) 是一致最大功效无偏的。

## 3.1.2 单个正态总体方差

设  $X \sim N(\mu, \sigma^2)$ ,  $\mu$  未知, 给定  $\alpha \in (0, 1)$ , 设  $x_1, \dots, x_n$  是样本  $X = (X_1 \cdots X_n)$  的值。

研究下列检验问题：

$$H_0: \sigma^2 = \sigma_0^2 \Leftrightarrow H_a: \sigma^2 \neq \sigma_0^2$$

现利用广义似然比检验法推导否定域为：

求  $x$  的似然函数

$$\begin{aligned} L(\tilde{x}, \theta) &= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + nx(\bar{x} - \mu)^2 \right] \right) \end{aligned}$$

得到

$$\begin{aligned} L(\hat{\theta}_0) &= \sup_{\theta \in \Theta} L(\tilde{x}, \theta) = \left[ \frac{1}{2\pi\sigma_0^2} \right]^{\frac{n}{2}} \exp \left( -\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right) \\ L(\hat{\theta}) &= \sup_{\theta} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}} \end{aligned}$$

知广义似然比

$$\begin{aligned} \lambda &= \left[ \frac{2\pi n \sigma_0^2}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} \exp \left( -\frac{n}{2} + \frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right) \\ &= \binom{n}{u}^{\frac{n}{2}} \exp \left( \frac{u - n}{2} \right) \end{aligned}$$

其中

$$u = \frac{1}{\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 > 0$$

由  $\lambda$  为  $u$  的增函数故否定域

$$W_0 = \{(x_1 \cdots x_n): \mu > C_2 \text{ 或 } < C_1\} \quad (3.6)$$

$C_1 < C_2$  满足：

$$P(X \in W_0 | \mu, \sigma_0) = a \quad (3.7)$$

能够证明，在  $H_0$  成立的条件下  $\mu = \mu(X_1, \dots, X_n)$  服从  $n-1$  个自由度的  $\chi^2$  分布，(3.7) 等价于

$$\int_0^{C_1} g_{n-1}(y) dy + \int_{C_2}^{\infty} g_{n-1}(y) dy = a \quad (3.8)$$

其中  $g_n(y)$  是  $n$  个自由度的  $\chi^2$  分布的分布密度函数。

否定域 (3.6) 在直观上合理的，数学上也可证明，如果  $C_1, C_2$  满足 (3.8) 及

$$\int_{C_1}^{C_2} g_{n-1}(y) dy = 1 - a$$

则有否定域 (3.6) 是一致最大功效无偏的。

在实际中常常取  $C_1 < C_2$ ，满足

$$\begin{aligned} \int_0^{C_1} g_{n-1}(y) dy &= \frac{a}{2} \\ \int_{C_2}^{\infty} g_{n-1}(y) dy &= \frac{a}{2} \end{aligned}$$

这样的  $C_1, C_2$  可通过查  $\chi^2$  分布表很容易得到。

设  $X \sim N(\mu, \sigma^2)$ ，给定  $a \in (0, 1)$ ，设  $x_1, \dots, x_n$  是样本  $X = (X_1, \dots, X_n)$  的值。

研究下列检验问题：

$$H_0: \sigma^2 \leq \sigma_0^2 \Leftrightarrow H_a: \sigma^2 > \sigma_0^2$$

现利用广义似然比检验法推导否定域：

求  $x$  的似然函数

$$\begin{aligned} L(x, \theta) &= \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \\ &= \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^n \exp \left( -\frac{1}{2\sigma^2} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 + nx(\bar{x} - \mu)^2 \right] \right) \end{aligned}$$

得到

$$L(\theta_0) = \sup_{\theta \in \Theta} L(x, \theta) = \left[ \frac{1}{2\pi\sigma_0^2} \right]^{\frac{n}{2}} \exp \left( -\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right)$$



$$L(\hat{\theta}) = \sup_{\theta} L(\tilde{x}, \theta) = \left[ \frac{n}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} e^{-\frac{n}{2}}$$

知广义似然比

$$\begin{aligned} \lambda &= \left[ \frac{2\pi n \sigma_0^2}{2\pi \sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} \exp \left( -\frac{n}{2} + \frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 \right) \\ &= \left( \frac{n}{u} \right)^{\frac{n}{2}} \exp \left( \frac{u-n}{2} \right) \quad u = \frac{1}{\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 > 0 \end{aligned}$$

由  $\lambda$  为  $u$  的增函数故否定域

$$W_0 = \{(x_1 \cdots x_n) : \mu > C_2\} \quad (3.9)$$

其中

$$\mu = \frac{1}{\sigma_0^2} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (3.10)$$

$C_2$  满足：

$$\int_{C_2}^{\infty} g_{n-1}(y) dy = \alpha$$

其中  $g_{n-1}(y)$  是  $n-1$  个自由度的  $\chi^2$  分布的密度， $C_2$  可通过查  $\chi^2$  分布表很容易得到。

数学上也可证明，否定域 (3.9) 是一致最大功效无偏的。

相类似的，对于检验问题：

$$H_0 : \sigma^2 \geq \sigma_0^2 \Leftrightarrow H_a : \sigma^2 < \sigma_0^2$$

也可用广义似然比类似的导出合适的否定域：

$$W_0 = \{(x_1 \cdots x_n) : \mu < C_1\} \quad (3.11)$$

其中  $u$  的表达式参见 (3.10)， $C_1$  满足

$$\int_0^{C_1} g_{n-1}(y) dy = \alpha$$

数学上也可证明，否定域 (3.11) 是一致最大功效无偏的。



## 3.2 广义似然比检验法在两个正态总体的应用

### 3.2.1 两个正态总体方差

研究两个正态总体的假设检验问题。设  $X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  $X$  的样本值是  $X = (X_1 \cdots X_{n_1})$ ,  $Y$  的样本值是  $Y = (Y_1 \cdots Y_{n_2})$ , 假设两个样本相互独立, 给定  $\alpha \in (0, 1)$ 。

研究下列检验问题：

$$H_0: \sigma_1^2 = \sigma_2^2 \Leftrightarrow H_a: \sigma_1^2 \neq \sigma_2^2$$

现利用广义似然比检验法推导否定域：

密度函数

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{1}{2} \left[ \frac{(x - \mu_1)^2}{\sigma_1^2} + \frac{(y - \mu_2)^2}{\sigma_2^2} \right] \right\}$$

求似然函数

$$L(x, y) = \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_2^2} \right] \right\}$$

得到

$$\begin{aligned} L(\bar{\Theta}) &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{\sigma_2^2} \right] \right\} \\ &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2) \\ L(\bar{\Theta}_0) &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{\sigma^2} \right] \right\} \\ &= \left( \frac{1}{2\pi\sigma\sigma} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2) \end{aligned}$$

由

$$\sigma^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_1 + n_2}$$

$$\sigma_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1} \quad \sigma_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_2}$$

知广义似然比

$$\begin{aligned} \lambda &= \left( \frac{\sigma_1^2}{\sigma^2} \right)^{\frac{n_1}{2}} \cdot \left( \frac{\sigma_2^2}{\sigma^2} \right)^{\frac{n_2}{2}} \\ &= \left( \frac{n_1}{n_1 + n_2} \right)^{\frac{n_1}{2}} \left( \frac{n_2}{n_1 + n_2} \right)^{\frac{n_2}{2}} \left[ 1 + \frac{n_1 - 1}{n_2 - 1} F(x, y) \right]^{\frac{n_2}{2}} \left[ 1 + \frac{n_2 - 1}{n_1 - 1} F(x, y) \right]^{\frac{n_1}{2}} \end{aligned}$$

其中

$$F(x, y) = \frac{S_1^2}{S_2^2} \quad S_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1 - 1} \quad S_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_2 - 1}$$

由上式，可知  $\lambda$  是关于  $F$  的一个凸函数，故否定域

$$W_0 = \{(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) : F < C_1 \text{ 或 } > C_2\} \quad (3.12)$$

其中

$$F = \frac{\frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{(n_1 - 1)}}{\frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{(n_2 - 1)}} \quad (3.13)$$

其中

$$\bar{x} = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i, \quad \bar{y} = \frac{1}{n_2} \sum_{i=1}^{n_2} y_i$$

由于

$$\begin{aligned} \frac{1}{\sigma_1^2} \sum_{i=1}^{n_1} (X_i - \bar{X})^2 &\sim \chi^2(n_1 - 1) \\ \frac{1}{\sigma_2^2} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2 &\sim \chi^2(n_2 - 1) \end{aligned}$$

这二者相互独立，所以在  $H_0$  成立下统计量  $F$  是服从  $n_1 - 1, n_2 - 1$  的  $F$  分布。通过

$F$  分布的分位数表容易找到  $C_1 < C_2$  满足

$$\int_0^{C_1} f_{n_1-1, n_2-1}(u) du + \int_{C_2}^{\infty} f_{n_1-1, n_2-1}(u) du = a \quad (3.14)$$

则在  $H_0$  成立下  $P(F < C_1, F < C_2) = a$  , 从而否定域 ( 3.12 ) 的检验水平是  $a$ 。

数学上可以证明：如果  $C_1 < C_2$  满足 ( 3.14 ) 及下式

$$\int_{a_1}^{a_2} B_{\frac{1}{2}(n_1+1), \frac{1}{2}(n_2-1)}(y) dy = 1 - a$$

其中

$$a_i = \frac{n_i - 1}{n_2 - 1} C_i / \left( \frac{n_i - 1}{n_2 - 1} C_i + 1 \right), i = 1, 2$$

$B_{r,s}(y)$  是  $\beta$  分布的密度函数，故否定域 ( 3.12 ) 是一致最大功效无偏的。

在实际中常常取  $C_1, C_2$  满足

$$\int_0^{C_1} f_{n_1-1, n_2-1}(y) dy = \int_{C_2}^{\infty} f_{n_1-1, n_2-1}(y) dy = \frac{a}{2}$$

注意

$$\begin{aligned} \int_0^{C_1} f_{n_1-1, n_2-1}(y) dy &= P(F < C_1 | H_0) = P\left(\frac{1}{F} > \frac{1}{C_1} | H_0\right) \\ &= \int_{C_1^{-1}}^{\infty} f_{n_2-1, n_1-1}(y) dy \end{aligned}$$

所以  $C_1, C_2$  均可以利用  $F$  分布表得到。

对于两个正态总体的检验问题：

$$H_0: \sigma_1^2 \leq \sigma_2^2 \Leftrightarrow H_a: \sigma_1^2 > \sigma_2^2$$

现利用广义似然比检验法推导否定域：

密度函数

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$$

求似然函数

$$L(x, y) = \left(\frac{1}{2\pi\sigma_1\sigma_2}\right)^{n_1+n_2} \exp\left\{-\frac{1}{2}\left[\frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_2^2}\right]\right\}$$

得到

$$\begin{aligned}
 L(\bar{\Theta}) &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{\sigma_2^2} \right] \right\} \\
 &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2) \\
 L(\bar{\Theta}_0) &= \left( \frac{1}{2\pi\sigma\sigma} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{\sigma^2} \right] \right\} \\
 &= \left( \frac{1}{2\pi\sigma\sigma} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2)
 \end{aligned}$$

由

$$\begin{aligned}
 \sigma^2 &= \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_1 + n_2} \\
 \sigma_1^2 &= \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1} \quad \sigma_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_2}
 \end{aligned}$$

知广义似然比

$$\begin{aligned}
 \lambda &= \left( \frac{\sigma_1^2}{\sigma^2} \right)^{\frac{n_1}{2}} \cdot \left( \frac{\sigma_2^2}{\sigma^2} \right)^{\frac{n_2}{2}} \\
 &= \left( \frac{n_1}{n_1 + n_2} \right)^{\frac{n_1}{2}} \left( \frac{n_2}{n_1 + n_2} \right)^{\frac{n_2}{2}} \left[ 1 + \frac{n_1 - 1}{n_2 - 1} F(x, y) \right]^{\frac{n_2}{2}} \left[ 1 + \frac{n_2 - 1}{n_1 - 1} F(x, y) \right]^{\frac{n_1}{2}}
 \end{aligned}$$

其中

$$F(x, y) = \frac{S_1^2}{S_2^2} \quad S_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2}{n_1 - 1} \quad S_2^2 = \frac{\sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_2 - 1}$$

由上式，可知  $\lambda$  是关于  $F$  的一个凸函数，故否定域

$$W_0 = \{F > C_2\} \quad (3.15)$$

这里  $F$  的表达式参见 (3.13)， $C_2$  由下式确定：

$$\int_{C_2}^{\infty} f_{n_1-1, n_2-1}(y) dy = \alpha$$

$C_2$  可由  $F$  分布表查得。

数学上也可证明，否定域 ( 3.15 ) 是一致最大功效无偏的。  
相类似的，研究下列检验问题：

$$H_0: \sigma_1^2 \geq \sigma_2^2 \Leftrightarrow H_a: \sigma_1^2 < \sigma_2^2$$

现利用广义似然比检验法推导否定域：

$$W_0 = \{F < C_1\} \quad (3.16)$$

这里  $F$  的表达式参见 ( 3.13 )， $C_1$  由下式确定：

$$\int_{C_1}^{\infty} f_{n_2-1, n_1-1}(y) dy = \alpha,$$

可利用  $F$  分布表得到  $C_1$  值。

数学上也可证明，否定域 ( 3.16 ) 是一致最大功效无偏的。

### 3.2.2 两个正态总体均值

现在研究  $\sigma_1^2 = \sigma_2^2$  时的下列检验问题

$$H_0: \mu_1 = \mu_2 \Leftrightarrow H_a: \mu_1 \neq \mu_2$$

现利用广义似然比检验法推导否定域：

密度函数

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{1}{2} \left[ \frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right] \right\}$$

求似然函数

$$\begin{aligned} L(x, y) &= \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_2^2} \right] \right\} \\ &= \left( \frac{1}{2\pi\sigma^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma^2} \right] \right\} \end{aligned}$$

当  $\mu_1 = \mu_2 = \mu$  有  $\bar{x} = \frac{n_1\bar{x} + n_2\bar{y}}{n_1 + n_2}$

$$\sigma_0^2 = \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2}{n_1 + n_2} \quad \sigma^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_1 + n_2}$$

得到

$$\begin{aligned} L(\bar{\Theta}) &= \left( \frac{1}{2\pi\sigma^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma^2} \right] \right\} \\ &= \left( \frac{n_1+n_2}{2\pi \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1+n_2) \\ L(\bar{\Theta}_0) &= \left( \frac{1}{2\pi\sigma_0^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_0^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_0^2} \right] \right\} \\ &= \left( \frac{n_1+n_2}{2\pi \sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1+n_2) \end{aligned}$$

知广义似然比

$$\lambda = \left( \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2} \right)^{n_1+n_2} = \left( \frac{S_1}{S_2} \right)^{n_1+n_2}$$

其中

$$\begin{aligned} S_1 &= \sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2 \\ S_2 &= \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2 \end{aligned}$$

由

$$S_1 = S_2 + n_1(\bar{x} - \mu_1)^2 + n_2(\bar{y} - \mu_2)^2$$

记

$$T^* = \frac{|\bar{x} - \bar{y}|}{\sqrt{S_2}}$$

易知  $\lambda$  是  $T^*$  的增函数，故否定域

$$W_0^* = \{(x, y) : |T^*| > C\}$$

此时，当  $H_0$  成立时，有

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{S_2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} = T^* \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \sim t(n_1 + n_2 - 2)$$

故可得否定域

$$W_0 = \{(x_1 \cdots x_{n_1}, y_1 \cdots y_{n_2}) : |T| > C\} \quad (3.17)$$

其中

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \quad (3.18)$$

给定检验水平  $\alpha$ ，在  $H_0$  成立下统计量  $T$  是服从  $n_1 + n_2 - 2$  的  $t$  分布，查  $t$  分布表有  $C$  满足

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

此时，否定域 (3.17) 的检验水平是  $\alpha$ 。

否定域 (3.17) 在直观上甚为合理，数学上可以证明，它是一致最大功效无偏的。

现在研究  $\sigma_1^2 = \sigma_2^2$  时的下列假设检验问题：

$$H_0 : \mu_1 \geq \mu_2 \leftrightarrow H_a : \mu_1 < \mu_2$$

现利用广义似然比检验法推导否定域：

密度函数

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left\{ -\frac{1}{2} \left[ \frac{(x - \mu_1)^2}{\sigma_1^2} + \frac{(y - \mu_2)^2}{\sigma_2^2} \right] \right\}$$

求似然函数

$$L(x, y) = \left( \frac{1}{2\pi\sigma_1\sigma_2} \right)^{n_1 + n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_1^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_2^2} \right] \right\}$$



$$= \left( \frac{1}{2\pi\sigma^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma^2} \right] \right\}$$

当  $\mu_1 = \mu_2 = \mu$  有  $\mu = \frac{n_1\bar{x} + n_2\bar{y}}{n_1 + n_2}$

$$\sigma_0^2 = \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2}{n_1 + n_2} \quad \sigma^2 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2}{n_1 + n_2}$$

得到

$$L(\bar{\Theta}) = \left( \frac{1}{2\pi\sigma^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma^2} \right] \right\}$$

$$= \left( \frac{n_1 + n_2}{2\pi \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2)$$

$$L(\bar{\Theta}_0) = \left( \frac{1}{2\pi\sigma_0^2} \right)^{n_1+n_2} \exp \left\{ -\frac{1}{2} \left[ \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2}{\sigma_0^2} + \frac{\sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sigma_0^2} \right] \right\}$$

$$= \left( \frac{n_1 + n_2}{2\pi \sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2} \right)^{n_1+n_2} \exp -\frac{1}{2} (n_1 + n_2)$$

知广义似然比

$$\lambda = \left( \frac{\sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2}{\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2} \right)^{n_1+n_2} = \left( \frac{S_1}{S_2} \right)^{n_1+n_2}$$

其中

$$S_1 = \sum_{i=1}^{n_1} (x_i - \mu_1)^2 + \sum_{i=1}^{n_2} (y_i - \mu_2)^2$$



$$S_2 = \sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{i=1}^{n_2} (y_i - \bar{y})^2$$

由

$$S_1 = S_2 + n_1(\bar{x} - \mu_1)^2 + n_2(\bar{y} - \mu_2)^2$$

记

$$T^* = \frac{|\bar{x} - \bar{y}|}{\sqrt{S_2}}$$

易知  $\lambda$  是  $T^*$  的增函数，故否定域

$$W_0^* = \{(x, y) : |T^*| > C\}$$

此时，当  $H_0$  成立时，有

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{S_2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} = T^* \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \sim t(n_1 + n_2 - 2)$$

故否定域

$$W_0 = \{(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) : T > C_2\} \quad (3.19)$$

其中  $T$  参见 (3.18)， $C_2$  满足

$$\int_{C_2}^{\infty} t_{n_1+n_2-2}(y) dy = \alpha$$

数学上可以证明，否定域 (3.19) 它是一致最大功效无偏的。  
相类似的，研究下列检验问题：

$$H_0 : \mu_1 \leq \mu_2 \leftrightarrow H_a : \mu_1 > \mu_2$$

现利用广义似然比检验法推导否定域：

$$W_0 = \{(x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) : T < C_1\} \quad (3.20)$$

其中  $T$  参见 (3.18)， $C_1$  满足

$$\int_{-\infty}^{C_1} t_{n_1+n_2-2}(y) dy = \alpha$$

数学上可以证明，否定域 (3.20) 是一致最大功效无偏的。

关于研究  $\sigma_1^2 \neq \sigma_2^2$  时，检验下列问题

$$H_0 : \mu_1 = \mu_2 \leftrightarrow H_a : \mu_1 \neq \mu_2$$

用似然比相当复杂，这是著名的 Behrens-Fisher 问题，这里暂且不讨论。

### 3.3 广义似然比检验法在二项分布的应用

#### 二项分布参数检验

设  $X \sim B(n, p)$ , 令随机变量  $X$  的取值 0 或 1,  $\underline{x} = (x_1 \cdots x_n)$  是样本  $\underline{X} = (X_1 \cdots X_n)$  的值, 且

$$p = P(X = x) = 1 - P(X = x)$$

随机变量  $X$  的概率函数是

$$P(X = x) = p^x (1 - p)^{1-x} \quad (x = 1, 0)$$

研究下列检验问题:

$$H_0: p = p_0 \Leftrightarrow p \neq p_0 \quad \text{给定 } \alpha$$

现利用广义似然比检验法推导否定域:

求  $\underline{x}$  的似然函数

$$\begin{aligned} L(\underline{x}; p) &= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} = p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \\ &= p^y (1-p)^{n-y} \end{aligned} \quad \text{记 } y = \sum_{i=1}^n x_i$$

得到

$$\begin{aligned} L(\hat{p}) &= \sup_p L(\underline{x}; p) = \left(\frac{y}{n}\right)^y \left(1 - \frac{y}{n}\right)^{n-y} \quad \left(\hat{p} = \frac{y}{n}\right) \\ L(\hat{p}_0) &= \sup_{p_0} L(\underline{x}; p_0) = p_0^y (1-p_0)^{n-y} \end{aligned}$$

知广义似然比

$$\lambda = \frac{\sup_p L(\underline{x}; p)}{\sup_{p_0} L(\underline{x}; p_0)} = \frac{\left(\frac{y}{n}\right)^y \left(1 - \frac{y}{n}\right)^{n-y}}{p_0^y (1-p_0)^{n-y}}$$

由于当  $p_0 < \frac{y}{n}$ ,  $p_0 > \frac{y}{n}$  时, 数学上可以证明,  $\lambda$  关于  $y$  分别是增函数和减函数, 故否定域

$$W_0' = \left\{ \underline{x}: \begin{array}{l} y > C_1 \\ y < C_2 \end{array} \right\}$$

其中  $C_1, C_2$  为常数, 满足

$$\begin{cases} P(y > C_1 | p_0) = \sum_{y=0}^{C_1} C_n^y p_0^y (1-p_0)^{n-y} = a \\ P(y < C_2 | p_0) = \sum_{y=0}^{C_2} C_n^y p_0^y (1-p_0)^{n-y} = a \end{cases} \quad (3.21)$$

此概率在容量大的情形下，计算相当繁杂。这里我采取近似正态的方法。  
记

$$U_n(y) = \frac{y - np_0}{\sqrt{np_0(1-p_0)}}$$

则当  $n \rightarrow \infty$  时， $U_n(y) \xrightarrow{L} N(0,1)$ ，显然这个  $U_n(y)$  容易算。

则 (3.21) 可以近似地转化为

$$\begin{cases} P(y > C_1 | p_0) = P(U_n(y) > U_1 | p_0) = \frac{a}{2} \\ P(y < C_2 | p_0) = P(U_n(y) < U_2 | p_0) = \frac{a}{2} \end{cases}$$

显然否定域

$$W_0 = \begin{cases} U_n(y) > U_1 \\ U_n(y) < U_2 \end{cases}$$

当  $H_0$  成立时，统计量  $U_n(y) \xrightarrow{L} N(0,1)$ ，故  $U_1, U_2$  可由查正态分布表得出。

### 3.4 广义似然比检验法在指数分布的应用

#### 指数分布参数检验

设  $X$  服从指数分布，设  $\underline{x} = (x_1 \cdots x_n)$  是样本  $X = (X_1 \cdots X_n)$  的值，随机变量  $X$  的概率函数是

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right), & x > 0 \\ 0, & x \leq 0 \end{cases} \quad \theta \in (0, \infty)$$

研究下列检验问题：

$$H_0: \theta = \theta_0 \leftrightarrow H_a: \theta \neq \theta_0 \quad \text{给定 } a$$

现利用广义似然比检验法推导否定域：

求  $\underline{x}$  的似然函数

$$L(\theta) = \left(\frac{1}{\theta}\right)^n \prod_{i=1}^n \exp\left(-\frac{x_i}{\theta}\right) = \left(\frac{1}{\theta}\right)^n \exp\left(-\frac{\sum_{i=1}^n x_i}{\theta}\right)$$

得到

$$L(\hat{\theta}) = \left(\frac{n}{\sum_{i=1}^n x_i}\right)^n \exp(-n)$$

$$L(\theta_0) = \left(\frac{1}{\theta_0}\right)^n \exp\left(-\frac{\sum_{i=1}^n x_i}{\theta_0}\right) \quad \left(\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i\right)$$

知广义似然比

$$\begin{aligned} \lambda &= \frac{L(\hat{\theta})}{L(\theta_0)} = \left(\frac{n\theta_0}{\sum_{i=1}^n x_i}\right)^n \exp\left(\frac{\sum_{i=1}^n x_i}{\theta_0} - n\right) \\ &= \left(\frac{2n\theta_0}{2\sum_{i=1}^n x_i}\right)^n \exp\left(\frac{2\sum_{i=1}^n x_i}{2\theta_0} - n\right) \\ &= \left(\frac{2n}{e}\right)^n \left(\frac{\theta_0}{2\sum_{i=1}^n x_i}\right)^n \exp\left\{\frac{1}{2} \left(\frac{2\sum_{i=1}^n x_i}{\theta_0}\right)\right\} \end{aligned}$$

取  $t = \frac{2\sum_{i=1}^n x_i}{\theta_0}$

则

$$\lambda(t) = \left(\frac{2n}{e}\right)^n \left(\frac{1}{t}\right)^n \exp\left(\frac{t}{2}\right)$$

当  $H_0$  成立时。统计量  $t : \chi^2(2n)$ ，设其密度函数为  $f(x)$ ，

在  $(0, 2n)$ ， $(2n, +\infty)$  时，数学上可以证明， $\lambda$  关于  $y$  分别是严格单调增函数和严

格单调减函数，故否定域

$$W_0 = \left\{ x: \begin{array}{l} t > C_1 \\ t < C_2 \end{array} \right\}$$

其中  $C_1, C_2$  为常数，且有下列式子确定

$$\begin{cases} \int_{C_2}^{C_1} f(x) dx = 1 - a \\ \lambda(C_1) = \lambda(C_2) \end{cases} \Rightarrow \begin{cases} \int_{C_2}^{C_1} \frac{1}{2^n \Gamma(n)} \exp\left(-\frac{x}{2}\right) dx = 1 - a \\ C_1^n \exp\left(-\frac{C_1}{2}\right) = C_2^n \exp\left(-\frac{C_2}{2}\right) \end{cases}$$

### 3.5 Stein 例子

对于非正态总体，广义似然比检验法也是可用的（但不是所有情形下的广义似然比检验法都好）。

设  $X$  取值限于集合  $\{2, -2, 1, -1, 0\}$ ,  $P(X = x) = p(x, \theta)$   $0 \leq \theta < 1$ , 其中  $\theta$  是未知数。

$$p(x, 0) = \begin{cases} \frac{a}{2}, & x = \pm 2 \\ \frac{1-2a}{2}, & x = \pm 1 \\ a, & x = 0 \end{cases}$$

$$p(x, \theta) = \begin{cases} \theta C, & x = -2 \\ \frac{1-C}{1-a} \left( \frac{1}{2} - a \right), & x = \pm 1 \\ a \left( \frac{1-C}{1-a} \right), & x = 0 \\ (1-\theta)C, & x = 2 \end{cases} \quad (0 < \theta < 1)$$

其中  $0 < a < \frac{1}{2}$   $\frac{a}{2-a} < C < a$ 。

研究下列检验问题：

$$H_0: \theta = 0 \Leftrightarrow H_a: 0 < \theta < 1$$

现利用广义似然比检验法导出合适的否定域

设样本是  $x_1$  (样本量是 1), 广义似然比为

$$\lambda(x) = \frac{\sup_{0 \leq \theta < 1} p(x, \theta)}{p(x, 0)}$$

易知

$$\lambda(2) = \frac{2C}{a} \quad (\text{因为 } \frac{a}{2} < C)$$

$$\lambda(-2) = \frac{2C}{a}$$

$$\lambda(-1) = \lambda(1) = \frac{1-C}{1-a}$$

$$\lambda(0) = \frac{1-C}{1-a}$$

水平上  $a$  的否定域为  $W_0 = \{x: \lambda(x) > \lambda_0\}$ , 其中  $\lambda_0$  适合

$$P(X_1 \in W_0 | \theta = 0) = a$$

易知取  $\lambda_0 = \frac{1-C}{1-a}$  时

$$P(\lambda(X_1) > \lambda_0 | H_0) = P(X_1 = \pm 2 | \theta = 0) = a$$

故否定域是  $W_0 = \{2, -2\}$ , 功效函数

$$\begin{aligned} P\left(\lambda(X_1) > \frac{1-C}{1-a} | \theta\right) &= P(X_1 = \pm 2 | \theta = 0) = \theta C + (1-\theta)C \\ &= C < a \quad (0 < \theta < 1) \end{aligned}$$

可见这个检验法不是无偏的。

下列检验法更好些, 去否定域  $W_1 = \{0\}$ , 既当且仅当  $X_1 = 0$  时拒绝  $H_0$ 。

此时  $P(X_1 \in W_1 | 0) = P(0, 0) = a$ 。功效函数

$$P(X_1 \in W_1 | \theta) = P(0, \theta) = a \frac{1-C}{1-a} > a \quad (0 < \theta < 1)$$

此检验法是无偏的。

## 4 似然比检验的极限分布概略

与最大似然估计一样, 似然比检验有很好的统计性质。但是似然比检验也有两个比较明显的缺点: 很强的依赖于模型中的密度函数 有时候很难得到零假设下的分布函数

定理: 令  $\theta_0$  是  $\Theta$  的一个内点。当  $H_0: \theta = \theta_0$  零假设为真时, 对数似然比检验

$\gamma(x) = -2 \ln \lambda(x)$  的极限分布为自由度  $k$  的  $\chi^2$  分布, 即  $\lambda(x) \xrightarrow{L} \chi_k^2$ 。

## 5 简述关于广义似然比检验的相合性

一个检验是相合的，如果当样本容量无限增加时，犯第二种错误的概率趋于零。

条件

对每个  $\theta \in \bar{\Theta}$  和  $P > 0$ ， $f(x, \theta, P)$  为  $B_x$  可测；

对每个  $\theta \in \bar{\Theta}$  及  $\theta_0 \in \bar{\Theta}$ ，存在  $P_{\theta\theta_0} > 0$ ，使当  $0 < P < P_{\theta\theta_0}$  时，有  $\int_x [\log^*(x, \theta, P)] f(x, \theta_0) d\mu < \infty$ 。又当  $r$  充分大（与  $\theta$  有关）时，  
 $\int_x [\log^*(x, r)] f(x, \theta_0) d\mu < \infty$ ；

对任何  $\theta \in \bar{\Theta}$  有  $\int_x |f(x, \theta)| f(x, \theta) d\mu(x) < \infty$ ；

存在一集  $A \in B_x$ ，使对任何  $\theta$  有  $\int_A f(x, \theta) d\mu(x) = 0$ ，且  $\lim_{|\theta| \rightarrow \infty} f(x, \theta') = 0$ ，  
 当  $x \notin A$ ；

当  $\theta_1 \neq \theta_2$  时， $\mu(\{x: f(x, \theta_1) \neq f(x, \theta_2)\}) > 0$ ；

对任何  $\theta \in \bar{\Theta}$  存在一集  $B_\theta \in B_x$ ，致  $\int_{B_\theta} f(x, \theta_0) d\mu = 0$  对任何  $\theta \in \bar{\Theta}_0$ ，且当  $x \in B_\theta$  时，有  $f(x, \theta') \rightarrow f(x, \theta)$ ，对任何  $\theta' \rightarrow \theta$ 。

条件

$$\int_x \frac{\partial f(x, \theta)}{\partial \theta_i} d\mu(x) = 0$$

$I(\theta) = (I_{ij}(\theta))_{k \times k} > 0$  对任何  $\theta \in \bar{\Theta}$ ，其中

$$I_{ij}(\theta) = \int_x \left( \frac{\partial \log f(x, \theta)}{\partial \theta_i} \right) \left( \frac{\partial \log f(x, \theta)}{\partial \theta_j} \right) f(x, \theta) d\mu(x)。$$

存在  $M(x)$ ，致  $\int_x M(x) f(x, \theta) d\mu(x) < \infty$  对任何  $\theta \in \bar{\Theta}$ ，且

$$\left| \frac{\partial^3 \log f(x, \theta)}{\partial \theta_a \partial \theta_b \partial \theta_c} \right| \leq M(x) \quad a, b, c = 1, \dots, k$$

不同的  $\theta$  值相应于  $X$  的不同分布。

条件



设  $X_1, \dots, X_n, \dots$  是变量  $X$  的 iid. 样本,  $X$  有分布族

$$\{f(x, \theta) d\mu(x), \theta = (\theta_1, \dots, \theta_k) \in \bar{\Theta}\} \quad (5.1)$$

分布族

$$\left\{ \begin{aligned} f(x, \varphi) d\mu &= f(x, \varphi_1, \dots, \varphi_r) d\mu(x) \\ &= f(x, g_1(\varphi_1, \dots, \varphi_r), \dots, g_k(\varphi_1, \dots, \varphi_r)) d\mu(x), \varphi \in A \end{aligned} \right\}$$

(5.2)  $A$  为集合

$\bar{\Theta}$  是  $R_k$  中的一个有内点的集合。

分布族 (5.1) 满足条件

分布族 (5.2) 满足条件

定理：若条件 和条件 都成立, 而  $\bar{\Theta}_H$  为  $\bar{\Theta}$  的一个闭子集, 则的广义似然

$$\text{比 } \left\{ \lambda(x) = \frac{L(\hat{\theta})}{L(\hat{\theta}_0)} \right\} \text{ 检验 (2.1) 是相合的。}$$

## 6 关于广义似然比检验法的几点看法

最后, 关于广义似然比检验法, 提出几点看法:

似然比的检验功效

要精确计算似然比检验的功效, 当然需要知道似然比分布。而这一般是难于求出的, 然而, 利用 4 中提出的关于在对立假设下似然比的极限分布, 可对功效值给予一个近似的估计。

似然比检验与无偏性

从似然比检验的相合性可知, 当样本大小很大时, 似然比检验所近似地无偏的。这种“渐近无偏性”是任何相合检验都有的性质。然而, 对固定的样本大小而言, 似然比检验可以不是无偏的。

似然比检验在直观上看是有根据的, 在具体问题上常给出具有良好性质的检验, 而且从理论上也可证明它具有某些良好的大样本性质。但是, 在个别特殊情况下它可以给出很坏的结果, 如 Stein 例子。



## 7 总结

本文通过叙述了假设检验的一些基本概念和思想，然后详细的阐述了广义似然比检验的原理，分别对单个正态总体、两个正态总体、二项分布、指数分布等做了具体的推导过程，最后简单叙述了似然比检验的极限分布及广义似然比检验法的相合性。

广义似然比检验法是 Neyman和 Pearson 在 1928 年提出的，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。这是一个应用很广的方法，而且，它构造出的检验常具有种种最优性质。

寻找一致最大功效（或一致最大功效无偏）的否定域是相当复杂的问题，需要对零假设  $H_0$  的具体结构及总体  $X$  的概率性质进行具体分析。而广义似然比检验法适用的范围较广泛，当样本量  $n$  较大时，广义似然比检验否定域往往相当好（即第二类错误的概率比较小），虽然不一定是一致的最大功效的。

## 参考文献

- [1] 陈希孺. 数理统计引论. 科学出版社, 1981.
- [2] 陈家鼎, 孙山泽, 李东风, 刘力平. 数理统计学讲义 (第二版), 高等教育出版社, 2008
- [3] 史宁中. 统计检验的理论与方法, 2008
- [4] 李贤平. 概率论基础 (第二版) [M]. 北京: 高等教育出版社, 1997.
- [5] 贾俊平, 何晓群, 金勇进. 统计学 (第三版). 中国人民大学出版社, 2007.
- [6] 葛余博. 概率论与数理统计. 清华大学出版社, 2005.
- [7] 茆诗松, 汤银才, 王玲玲. 可靠性统计. 高等教育出版社, 2008.
- [8] Lehmann E L. Testing Statistical Hypotheses. 2nd edition. Wiley, 1986.
- [9] (美) 康斯坦丝·瑞德 著, 姚慕生, 陈克艰, 王顺义 译. 奈曼: 来自生活的统计学家 伟人与时代. 上海科学技术出版社, 2001.
- [10] 华东师范大学数学系. 数学分析 (上) (第三版). 高等教育出版社, 2006
- [11] 华东师范大学数学系. 数学分析 (下) (第三版). 高等教育出版社, 2006