

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



数理统计

目录

第一章	统计量与抽样分布
第二章	点估计
第三章	区间估计
第四章	假设检验
第五章	分布的检验
第六章	Bayes统计与统计判决理论
第七章	非参数统计
第八章	抽样调查

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

假设检验

假设检验是统计学中最具特色的部分，其统计意味甚浓。从建立假设，寻找检验统计量，构造拒绝域（或计算 p 值），知道最后做出判断等各个步骤上都能体现多种统计思想的亮点。

假设检验方法和理论的发展，始于20世纪初。大致上按照时间顺序，其过程经历了以下重大事件：

- （1）K.Pearson 的拟合优度检验(1900年)
- （2）R.A.Fisher 的显著性检验(20世纪20年代)
- （3）J.Neyman和E.S.Pearson的NP理论(1928年开始)
- （4）A.Wald的统计判决理论(1950年)
- （5）Bayes方法

Fisher显著性检验思想

例 4.1.1 (女士品茶试验) 一种饮料由牛奶与茶按一定比例混合而成, 可以先倒茶后牛奶(记为TM), 也可以反过来(记为MT). 某女士声称她可以鉴别是TM还是MT. 为此Fisher设计了如下试验来检验她的说法是否可信. 准备8杯饮料, TM与MT各一半, 把它们随机地排在一起, 让她品尝, 并告诉她TM与MT各四杯, 然后让她指出哪四杯是TM. 假设她全说对了, 你相信她有这个能力吗?

Fisher显著性检验思想

解 对此问题，Fisher先引进一个假设：

$$H: \text{该女士并无鉴别能力.} \quad (4.1.1)$$

当 H 正确时，她只能随机地指出四杯是TM，而从8杯中选4杯共有 $\binom{8}{4} = 70$ 种选法，其中仅有一种判断是正确的。于是，当 H 正确时，这个女士选对四杯的概率为 $1/70$ 。此时，我们必须承认，下述两个情况必发生其一：

1. H 不成立，即该女士有一定的鉴别力；
2. 发生了一件概率只有 $\frac{1}{70}$ 的事件。

显然，由于第二件事件发生比较稀奇，因而我们有相当的理由认为第一种情况发生了，或者说，该女士4杯全选对这一事件是一个不利于假设 H 的显著性证据。这样的推理过程就叫做显著性检验。

如果该女士只说对了三杯，则表面上看，4杯说对了3杯，成绩已经不错，但我们可以计算一下，纯粹出于碰巧而得到这个以至更好成绩的机会会有多大，即当 H 成立时，从70种不同的挑法中，挑对3杯以上的概率为 $17/70 = 0.243$ 。显然发生一个概率为0.243的事件并不稀奇。因此，试验结果没有提供不利于 H 的显著性证据。

□

Fisher显著性检验

虽然Fisher提出的显著性检验在实际得到了广泛的应用，但是对于同一个问题，可以有很多不同的检验方法。于是，自然就有一个择优的问题，于是Neyman 和E. Pearson提出了优势检验，从而弥补了显著性检验的这一不足。

基本概念

例 4.1.2 某工厂在把产品卖给商店之前，与商店约定：如果此批产品的次品率小于某个给定值，如0.05，则商店接受这批产品，否则拒收。同时，双方同意从此批产品中随机地抽取 n 件产品以估计其次品率 p 。现以 X 记抽取 n 件产品中的次品数，则商店如何根据 X 的大小来决定是否接受这批产品。

对于此问题，如我们假设 $X \sim B(n, p)$ ，则上述问题就可以归结成如下的统计问题：设有来自总体 $B(n, p)$ 的样本 X ，我们如何根据 X 的观测值对命题

$$H : p \leq 0.05$$

作出“对”或“不对”的判断。

在统计中，我们把这种需要根据样本去推断其“正确”与否的命题，就称为一个假设或统计假设。当然，假设可以是关于参数的，也可以是关于分布的。

基本概念：原假设与备择假设

对于例4.1.1，我们可以把假设检验问题总结如下：

设有样本 $\mathbf{X}$ ，取值于样本空间 $\mathcal{X}$ ，且知道样本来自某一个参数分布族 $\{F(x, \theta) : \theta \in \Theta\}$ ，其中 Θ 为参数空间。设 $\Theta_0 \subset \Theta$ ，且 $\Theta_0 \neq \emptyset$ ，则命题 $H_0 : \theta \in \Theta_0$ 称为一个假设或零假设(null hypothesis)。如记 $\Theta_1 = \Theta - \Theta_0$ ，则命题 $H_1 : \theta \in \Theta_1$ 称为 H_0 的对立假设或备选假设(alternative hypothesis)。于是，我们感兴趣的假设就是

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1, \quad (4.1.2)$$

我们称此假设的检验问题为假设检验问题。

基本概念：原假设与备择假设

实际上，对于例4.1.1，我们可以把其零假设写成如下形式：

$$H_0 : P\{X = i\} = \frac{\binom{4}{i} \binom{4}{4-i}}{\binom{8}{4}},$$

其中 X 表示该女士猜对的杯数，此时的零假设表示 X 服从超几何分布，它没有一个参数。

对于例4.1.2，参数空间 $\Theta = (0, 1)$, $\Theta_0 = (0, 0.05]$ ，就与例4.1.1有所不同，由于假设完全可由参数表示出来。

对于假设(4.1.2)，如果 Θ_0 只有一个点，则我们称之为简单(simple)零假设，否则就称为复杂(composite)或复合原假设。同样，对于备选假设也有简单与复合之别。当 H_0 为简单假设时，其形式可写成 $H_0 : \theta = \theta_0$ 。此时的备选假设有两种可能：

$$H'_1 : \theta \neq \theta_0, \quad H''_1 : \theta < \theta_0 \text{ 或 } \theta > \theta_0.$$

我们称 $H_0 \longleftrightarrow H'_1$ 为双边的， $H_0 \longleftrightarrow H''_1$ 为单边的。

基本概念：拒绝域

另外，对于假设(4.1.2)的检验就是指这样的一个法则或策略：当有了具体的样本后，由该法则或策略就可决定是接受 H_0 还是否定 H_0 ，即检验就等价于把样本空间 $\mathcal{X}$ 划分成两个互不相交的部分 W 和 $\bar{W}$ ，当样本属于 $\bar{W}$ 时，接受 H_0 ；否则拒绝 H_0 。于是，我们称 W 为该检验的拒绝域，而 $\bar{W}$ 为接受域。

基本概念：两类错误

由于上述两种错误决策受随机样本的影响，故它们也具有随机性，于是，我们定义犯第一、二类错误概率如下：

第一类错误概率： $\alpha = P_{\theta}\{\mathbf{X} \in W\}$, $\theta \in \Theta_0$ ，也记为 $P\{\mathbf{X} \in W|H_0\}$,

第二类错误概率： $\beta = P_{\theta}\{\mathbf{X} \in \bar{W}\}$, $\theta \in \Theta_1$ ，也记为 $P\{\mathbf{X} \in \bar{W}|H_1\}$.

决策	H_0 为真	H_1 为真
接受 H_0	正确	Type II
拒绝 H_0	Type I	正确

为了方便记忆，我们分别称第一、二类错误为拒真与纳伪.

基本概念：势函数

既然一个检验都有如上两类错误，那我们能否找到一个检验，使其犯两类错误的概率都尽可能地小呢？实际上，我们做不到这一点。为了说明其原因，我们先引进如下的势函数(power function)或功效函数的概念。

定义 4.1.1 (势函数) 对于假设(4.1.2)的一个检验方法 ψ ，其拒绝域记为 W ，则我们称

$$\beta_{\psi}(\theta) = P_{\theta}\{\mathbf{X} \in W\}, \forall \theta \in \Theta, \quad (4.1.3)$$

为此检验的势函数。

从上一定义可以看出，当 $\theta \in \Theta_0$ 时，此检验犯第一类错误的概率等于其势函数 $\beta_{\psi}(\theta)$ ；而当 $\theta \in \Theta_1$ 时，检验犯第二类错误的概率等于 $1 - \beta_{\psi}(\theta)$ 。

基本概念：势函数

例 4.1.3 如何判断厂家欺骗顾客：如果买到标重是500克的袋糖，其实际重量为400克，我能告厂家欺诈吗？

解 如以 μ 表示厂家生产此种袋糖的平均重量，则我们要检验的假设为

$$H_0 : \mu = 500 \longleftrightarrow H_1 : \mu < 500.$$

顾客为了打官司，就要搜集证据，假设他又买了 n 袋糖，称其重量为 $X_1, \dots, X_n$ ，显然如果 $\bar{X} = \sum_{i=1}^n X_i/n$ 远小于500，这场官司就有胜的把握，即当 $\bar{X} - 500 < c$ 时，我们可以拒绝 H_0 ，于是，取拒绝域为 $\{\mathbf{X} : \bar{X} - 500 < c\}$ 是合理的. 至于如何确定常数 c ，就要依赖于错误概率的选取.

如假设总体服从 $N(\mu, \sigma_0^2)$ ，且 σ_0 已知，则此时的检验统计量为 $T(X) = \sum_{i=1}^n X_i/n$ ，拒绝域为 $W = \{\mathbf{x} : \bar{x} < 500 + c\}$ ，于是，由定义4.1.1知，这个检验的势函数为 $\beta(\mu) = P_\mu\{\bar{X} < 500 + c\}$. 这样其第一、二类错误概率分别为

$$\alpha = P_{\mu=500}\{\bar{X} < 500 + c\} = P_{\mu=500}\left\{\frac{\bar{X} - 500}{\sigma_0/\sqrt{n}} < \frac{c}{\sigma_0/\sqrt{n}}\right\} = \Phi(\sqrt{n}c/\sigma_0), \quad (*1)$$

$$\beta = P_\mu\{\bar{X} > 500 + c\} = P_\mu\left\{\frac{\bar{X} - \mu}{\sigma_0/\sqrt{n}} > \frac{500 + c - \mu}{\sigma_0/\sqrt{n}}\right\} = 1 - \Phi\left(\frac{500 + c - \mu}{\sigma_0/\sqrt{n}}\right). \quad (*2)$$

比较一下(*1)与(*2)可知，这两个错误概率均依赖于常数 c ，并且对于固定的样本容量，我们找不到一个 c ，使得二者均尽可能小. □

基本概念：显著性水平

既然我们不可能同时控制一个检验的第一、二类错误概率，那我们通常的作法是仅限制第一类错误概率，即我们有如下的定义.

定义 4.1.2 (显著性水平) 对于检验 ψ 和事先给定的 $\alpha \in (0, 1)$ ，如果它满足

$$P_{\theta}\{\mathbf{X} \in W\} \leq \alpha, \forall \theta \in \Theta_0, \quad (4.1.4)$$

则称 α 是检验 ψ 的水平或显著性水平，也称 ψ 为显著性水平 α 的检验.

一般地，我们取 $\alpha = 0.01, 0.05, 0.1$ 这三个值，其原因在于方便实际应用者查表. 在女士品茶试验中，当时我们所讲的阈值就是这里的显著性水平.

显著性检验

另外，当我们只控制一个检验的第一类错误概率时，就称这样的检验为显著性检验。一般情况下，求取某假设的显著性检验的步骤如下：

- ❶ 根据实际问题，建立统计假设 $H_0 \longleftrightarrow H_1$.
- ❷ 选取一个合适的统计量 $T(\mathbf{X})$ ，使当 H_0 成立时， T 的分布已知，且与参数 θ 无关(称此分布为统计量 T 的零分布).
- ❸ 根据 H_0 及 H_1 的特点，确定拒绝域 W 的形状.
- ❹ 对于给定的显著性水平 α ，确定拒绝域 W .
- ❺ 由样本观测值 $\mathbf{x}$ ，计算统计量 $T(\mathbf{X})$ 的值 $T(\mathbf{x})$ ，由 $T(\mathbf{x})$ 是否属于 W ，作出最终判断.

假设检验的思想

- (1) 反证法的思想;
- (2) 小概率事件在一次实验中不可能发生;
- (3) “无罪假定”

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

正态总体参数的假设检验

正态分布是最常用的分布，正态总体参数的检验也是最常遇到的检验问题，接下来我们将从以下几个方面谈论正态总体参数的假设检验：

一，单正态总体的检验

- 1，单正态总体均值的检验
- 2，单正态总体方差的检验

二，两正态总体的检验

- 1，两正态总体均值的检验
- 2，两正态总体方差的检验

单正态总体参数的假设检验

例 4.2.1 某糖厂生产某种袋装糖, 如果装袋机工作正常, 则每袋的标准重量为100(克). 根据以往经验, 当生产正常时, 其各袋重量的标准差为 $\sigma = 1.15$ (克). 为了检测此台袋装机是否工作正常, 现从中随机地抽取9袋糖, 测其重量如下(单位: 克):

99.3, 98.7, 100.5, 101.2, 98.3, 99.7, 99.5, 102.1, 100.5.

请问此台包装机工作是否正常?

单正态总体参数的假设检验

§4.2.1 单样本正态总体均值的检验

设 $X_1, \dots, X_n$ 为来自 $N(\mu, \sigma^2)$ 的iid样本, 现在我们感兴趣的是关于其均值 μ 的如下假设:

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu \neq \mu_0, \quad (4.2.1)$$

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu < \mu_0, \quad (4.2.2)$$

$$H_0 : \mu = \mu_0 \longleftrightarrow H_1 : \mu > \mu_0, \quad (4.2.3)$$

$$H_0 : \mu \leq \mu_0 \longleftrightarrow H_1 : \mu > \mu_0, \quad (4.2.4)$$

$$H_0 : \mu \geq \mu_0 \longleftrightarrow H_1 : \mu < \mu_0, \quad (4.2.5)$$

其中 μ_0 是已知的常数. 上述假设有单边的、双边的、简单的及复合的, 另外, σ^2 的已知与否对上述假设检验问题是有影响的. 于是, 我们就 σ^2 已知与未知两种情况来考虑上述五种假设的检验问题.

单正态总体参数的假设检验

一、当 $\sigma^2 = \sigma_0^2$ 已知时

❶ 对于假设(4.2.1)，我们感兴趣的参数在于总体均值 μ 的大小，而在第二章我们讲过，对于正态总体而言，样本均值 $\bar{X}$ 是 μ 的一个很好的点估计。于是，当假设(4.2.1)的 H_0 成立，即 $\mu = \mu_0$ 时， $\bar{X}$ 应与 μ_0 相差不多。而当 H_1 成立时， $\bar{X}$ 与 μ_0 应相差较大。这样，我们可以用 $|\bar{X} - \mu_0|$ 的大小来反映假设(4.2.1)，并且当 $|\bar{X} - \mu_0| > c$ 时，我们有理由拒绝 H_0 ，即认为 H_1 成立。

为了标准化，因为 $\text{Var}(\bar{X} - \mu_0) = \sigma_0^2/n$ ，故我们可以取此假设的检验统计量为

$$U(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}. \quad (4.2.6)$$

根据上述分析，此检验的拒绝域为

$$W = \{\mathbf{x} : |U(\mathbf{x})| > c\}.$$

再根据显著性检验的特点，上述拒绝域中的常数 c 由其事先给定的显著性水平 α 确定，即我们要求此检验的第一类错误概率不大于 α ，也就是说，其常数 c 满足：

$$P_{H_0}\{|U| > c\} \leq \alpha.$$

单正态总体参数的假设检验

又由于当 H_0 成立时, $U(\mathbf{X}) \sim N(0, 1)$, 故上述拒绝域中的常数 c 可取为 $c = u_{\alpha/2}$, 由于此常数是拒绝与接受零假设的分水岭, 故我们称之为检验的临界值(critical value).

总之, 我们把此时关于假设(4.2.1)的显著性检验总结如下:

- 检验统计量: $U(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$,
- 检验的拒绝域: $\{\mathbf{x} : |U(\mathbf{x})| > u_{\alpha/2}\}$,

通过上面分析可知, 此检验即为假设(4.2.1)的水平为 α 的显著性检验, 由于此时检验统计量 U 的零分布是标准正态分布, 且是双边的, 故我们称之为双边的 u 检验.

对于例4.2.1, 我们可以利用上述检验来处理之. 此时, $\mu_0 = 100, \sigma_0 = 1.15$, 因为

$$|U| = \left| \frac{\sqrt{9}(99.98 - 100)}{1.15} \right| = |-0.052| < 1.96 = u_{0.05/2},$$

故在显著性水平0.05下, 我们没有充分的理由拒绝 H_0 , 也就是说, 我们不能认为此台包装机工作不正常.

单正态总体参数的假设检验

② 对于假设(4.2.2)，由于它与假设(4.2.1)的区别在于备选假设的不同，故我们仍可采用关于假设(4.2.1)的检验统计量，但此时根据其备选假设的特点，我们取其拒绝域为

$$W = \{\mathbf{x} : U(\mathbf{x}) < c\}.$$

又由于在 H_0 下，统计量 $U(\mathbf{X}) \sim N(0, 1)$ ，故在控制其第一类错误概率为 α 下，其拒绝域中的常数 c 可取为 $c = -u_\alpha$.

于是，关于假设(4.2.2)的水平为 α 的显著性检验为：

- 检验统计量： $U(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$,
- 检验的拒绝域： $\{\mathbf{x} : U(\mathbf{x}) < -u_\alpha\}$,

我们称之为单边的 u 检验.

单正态总体参数的假设检验

③ 对于假设(4.2.3)，用类似于假设(4.2.2)的方法，可以得到其水平为 α 的显著性检验为：

- 检验统计量： $U(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$,
- 检验的拒绝域： $\{\mathbf{x} : U(\mathbf{x}) > u_\alpha\}$,

也称之为单边的 u 检验.

单正态总体参数的假设检验

④ 对于假设(4.2.4)，由于此时我们感兴趣仍然是均值参数 μ ，故我们仍可以用上述的检验统计量(4.2.6)，虽然此时它的零分布并不是标准正态 $N(0, 1)$ 。另外，根据备选假设的特点，此检验的拒绝域可取为

$$W = \{\mathbf{X} : U(\mathbf{X}) > c\}.$$

单正态总体参数的假设检验

为了控制此检验的第一类错误概率为给定的 α ，我们要求临界值 c 满足

$$\alpha(\mu) = P_{H_0}\{U(\mathbf{X}) > c\} \leq \alpha.$$

事实上， $c = u_\alpha$ 即满足上式的要求. 实际上，因为当 $H_0 : \mu \leq \mu_0$ 成立时，样本均值 $\bar{X} \sim N(\mu, \sigma_0^2/n)$ ，故

$$\begin{aligned}\alpha(\mu) &= P_{H_0}\{U(\mathbf{X}) > u_\alpha\} \\&= P_{H_0}\left\{\frac{\sqrt{n}(\bar{X} - \mu + \mu - \mu_0)}{\sigma_0} > u_\alpha\right\} \\&= P_{H_0}\left\{\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma_0} > u_\alpha - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right\} \\&= 1 - \Phi\left(u_\alpha - \frac{\sqrt{n}(\mu - \mu_0)}{\sigma_0}\right) \quad (\text{注意到此时}\mu \leq \mu_0) \\&\leq 1 - \Phi(u_\alpha) = \alpha.\end{aligned}$$

综上所述，假设(4.2.4)的水平为 α 的显著性检验同假设(4.2.3)的水平为 α 的显著性检验.

单正态总体参数的假设检验

⑥ 对于假设(4.2.5), 用类似于假设(4.2.4)方法, 可以说明, 它的水平为 α 的显著性检验同假设(4.2.2) 的水平为 α 的显著性检验.

例 4.2.2 假设某种型号三极管的放大倍数服从正态分布, 其期望值不小于100(hfe). 为了检验其生产情况是否正常, 现从中随机地抽取30只三极管, 测得其放大倍数的平均值为95. 请问最近生产的三极管的放大倍数是不是降低了? (假设总体方差 $\sigma^2 = 50$)

解 对于此问题, 我们感兴趣的假设为(4.2.5), 且 $\mu_0 = 100, \sigma_0^2 = 50, n = 30, \bar{X} = 95$. 对于 $\alpha = 0.05$, 我们有

$$U = \frac{\sqrt{30}(\bar{X} - \mu_0)}{\sqrt{\sigma_0^2}} = \frac{\sqrt{30}(95 - 100)}{\sqrt{50}} = -3.9 < -1.645 = -u_{0.05},$$

故我们有理由拒绝零假设, 即认为最近生产的三极管的放大倍数降低了. □

单正态总体参数的假设检验

二、当 σ^2 未知时

在上一段的内容中，我们始终假设正态总体的方差是已知的，但在许多实际问题，总体方差可能是未知的。于是，本段将考虑当总体方差未知时，关于上述五个假设的显著性检验问题。

对于假设(4.2.1)，由于样本均值 $\bar{X}$ 依然是总体均值的一个好的点估计，故知，当 H_0 成立时， $\bar{X}$ 与 μ_0 应比较接近，而当 H_1 成立时，二者应相差较大。但是，此时样本均值 $\bar{X}$ 的零分布 $N(\mu_0, \sigma^2/n)$ 并非完全已知。另外，我们注意到，样本方差 S_n^2 是总体方差 σ^2 的一个好的点估计，且当 H_0 成立时，有

$$T(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S_n} \sim t(n-1), \quad (4.2.7)$$

于是，我们可采用它作为检验统计量。

单正态总体参数的假设检验

这样，用类似于上一段的方法，我们可以得到此时假设(4.2.1)的水平为 α 的显著性检验如下：

- 检验统计量： $T(\mathbf{X}) = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S_n}$,
- 检验的拒绝域： $\{\mathbf{x} : |T(\mathbf{x})| > t_{\alpha/2}(n-1)\}$,

由于此时检验统计量 T 的零分布是 t 分布，且是双边的，故我们称之为双边的 t 检验。

对于其余四个假设，我们均可用类似于上一段的方法得到其水平为 α 的显著性检验，只是这里所用的检验统计量为(4.2.7)式的 $T(\mathbf{X})$ ，且相应的临界值为 $t(n-1)$ 分布的相应分位数。

综上所述，我们把关于单样本正态总体均值的水平为 α 的显著性检验总结在表4.1中。

单正态总体参数的假设检验

表4.1 单样本正态总体均值的显著性检验

方差	假设	检验统计量	拒绝域	名字
$\sigma^2 = \sigma_0^2$	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu \neq \mu_0$	$U = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma_0}$	$\{ U > u_{\alpha/2}\}$	双边 u 检验
	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu < \mu_0$		$\{U < -u_{\alpha}\}$	单边 u 检验
	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu > \mu_0$		$\{U > u_{\alpha}\}$	单边 u 检验
	$H_0: \mu \leq \mu_0 \longleftrightarrow H_1: \mu > \mu_0$		$\{U > u_{\alpha}\}$	单边 u 检验
	$H_0: \mu \geq \mu_0 \longleftrightarrow H_1: \mu < \mu_0$		$\{U < -u_{\alpha}\}$	单边 u 检验
σ^2 未知	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu \neq \mu_0$	$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S_n}$	$\{ T > t_{\alpha/2}(n-1)\}$	双边 t 检验
	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu < \mu_0$		$\{T < -t_{\alpha}(n-1)\}$	单边 t 检验
	$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu > \mu_0$		$\{T > t_{\alpha}(n-1)\}$	单边 t 检验
	$H_0: \mu \leq \mu_0 \longleftrightarrow H_1: \mu > \mu_0$		$\{T > t_{\alpha}(n-1)\}$	单边 t 检验
	$H_0: \mu \geq \mu_0 \longleftrightarrow H_1: \mu < \mu_0$		$\{T < -t_{\alpha}(n-1)\}$	单边 t 检验

单正态总体参数的假设检验

§4.2.2 单样本正态总体方差的检验

假设 $X_1, \dots, X_n$ 为来自正态总体 $N(\mu, \sigma^2)$ 的iid样本, 我们感兴趣的假设为

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 \neq \sigma_0^2, \quad (4.2.8)$$

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 < \sigma_0^2, \quad (4.2.9)$$

$$H_0 : \sigma^2 = \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 > \sigma_0^2, \quad (4.2.10)$$

$$H_0 : \sigma^2 \leq \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 > \sigma_0^2, \quad (4.2.11)$$

$$H_0 : \sigma^2 \geq \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 < \sigma_0^2. \quad (4.2.12)$$

类似于均值的检验, 我们也分均值已知与未知两种情况来考虑上述问题的显著性检验.

单正态总体参数的假设检验

一、当 $\mu = \mu_0$ 已知时

❶ 关于假设(4.2.8)，由于此时 σ^2 的一个很好的点估计为 $\sum_{i=1}^n (X_i - \mu_0)^2$ ，且当 H_0 成立时， $\sum_{i=1}^n (X_i - \mu_0)^2 / \sigma^2 \sim \chi^2(n)$ ，于是，我们可以选取检验统计量为

$$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}, \quad (4.2.13)$$

另外，由此时零假设与备选假设的特点可知，其拒绝域为

$$W = \{\chi^2 < c_1\} \cup \{\chi^2 > c_2\},$$

其中 $c_1 < c_2$ 为两个待定的常数，且由检验的显著性水平 α 来确定，即它们满足

$$P_{H_0}\{\mathbf{X} \in W\} \leq \alpha,$$

即

$$P_{H_0}\{\chi^2 < c_1\} + P_{H_0}\{\chi^2 > c_2\} \leq \alpha. \quad (4.2.14)$$

单正态总体参数的假设检验

虽然在 H_0 下, $\chi^2 \sim \chi^2(n)$, 但我们并不能由(4.2.14)式确定两个未知常数. 于是, 一个常用的方法即是如下选取 c_1, c_2 :

$$P_{H_0}\{\chi^2 < c_1\} \leq \alpha/2, \quad P_{H_0}\{\chi^2 > c_2\} \leq \alpha/2. \quad (4.2.15)$$

这样, 由 χ^2 的零分布及(4.2.15)式, 我们得到满足(4.2.14)式的临界值 c_1, c_2 如下:

$$c_1 = \chi_{1-\alpha/2}^2(n), \quad c_2 = \chi_{\alpha/2}^2(n).$$

于是, 我们得到假设(4.2.8)的水平为 α 的一个显著性检验为

- 检验统计量 $\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域: $W = \{\chi^2 < \chi_{1-\alpha/2}^2(n)\} \cup \{\chi^2 > \chi_{\alpha/2}^2(n)\}$.

单正态总体参数的假设检验

❷ 关于假设(4.2.9)，我们可以利用类似于(4.2.8)的方法，求得其水平为 α 的显著性检验为

- 检验统计量 $\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域: $W = \{\chi^2 < \chi_{1-\alpha}^2(n)\}$.

❸ 关于假设(4.2.10)，我们也可以利用类似于(4.2.8)的方法，求得水平为 α 的显著性检验为

- 检验统计量 $\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}$.
- 检验的拒绝域: $W = \{\chi^2 > \chi_{\alpha}^2(n)\}$.

单正态总体参数的假设检验

④ 关于假设(4.2.11), 虽然此时的零假设为复合的, 但我们感兴趣的仍然是总体方差的大小, 故我们仍采用(4.2.13)式的 χ^2 统计量作为此假设问题的检验统计量, 且根据此时备选假设的特点, 其拒绝域为

$$W = \{\chi^2 > c\},$$

其中常数 c 由控制第一类错误概率得到, 即

$$P_{H_0}\{\chi^2 > c\} \leq \alpha.$$

但我们注意到, 在 H_0 下, (4.2.13)式的 χ^2 统计量的分布并不是 $\chi^2(n)$. 然而, 我们有

$$\begin{aligned}\alpha &\geq P_{H_0}\{\chi^2 > c\} \\ &= P_{H_0}\left\{\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2} > c\right\} \\ &= P_{H_0}\left\{\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} > \frac{c\sigma_0^2}{\sigma^2}\right\},\end{aligned}\tag{4.2.16}$$

由于在 H_0 成立时, $\sigma^2 \leq \sigma_0^2$, 故只要我们取 c 满足

$$\alpha \geq P_{H_0}\left\{\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} > c\right\},\tag{4.2.17}$$

单正态总体参数的假设检验

则因为

$$P_{H_0} \left\{ \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} > c \right\} \geq P_{H_0} \left\{ \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} > \frac{c\sigma_0^2}{\sigma^2} \right\},$$

故它一定满足(4.2.16)式.

由于在 H_0 成立时, (4.2.17)式中的 $\frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma^2} \sim \chi^2(n)$, 故我们可以取 $c = \chi_\alpha^2(n)$. 于是, 我们可知, 关于假设(4.2.10)的水平为 α 的显著性也是假设(4.2.11)的水平为 α 的显著性检验.

⑥ 关于假设(4.2.12), 我们同样可以利用关于假设(4.2.9)的水平为 α 的显著性检验作此假设的显著性检验.

单正态总体参数的假设检验

二、当 μ 未知时

此时，我们可以用类似于导出 μ 已知时的检验方法得到五个假设的水平为 α 的显著性检验。只是注意到，此时所用的检验统计量为

$$\chi^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2}, \quad (4.2.18)$$

此时各假设的临界值与 μ 已知时非常类似，只需要把那里的自由度 n 换成现在的自由度 $n - 1$ 即可。总之，我们把上述两种情况下关于单样本正态总体方差的显著性检验总结见表4.2。

单正态总体参数的假设检验

表4.2 单样本正态总体方差的显著性检验

均值	假设	检验统计量	拒绝域
$\mu = \mu_0$	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{\sum_{i=1}^n (X_i - \mu_0)^2}{\sigma_0^2}$	$\{\chi^2 < \chi_{1-\alpha/2}^2(n)\} \cup \{\chi^2 > \chi_{\alpha/2}^2(n)\}$
	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 < \sigma_0^2$		$\{\chi^2 < \chi_{1-\alpha}^2(n)\}$
	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 > \sigma_0^2$		$\{\chi^2 > \chi_{\alpha}^2(n)\}$
	$\sigma^2 \leq \sigma_0^2 \longleftrightarrow \sigma^2 > \sigma_0^2$		$\{\chi^2 > \chi_{\alpha}^2(n)\}$
	$\sigma^2 \geq \sigma_0^2 \longleftrightarrow \sigma^2 < \sigma_0^2$		$\{\chi^2 < \chi_{1-\alpha}^2(n)\}$
μ 未知	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 \neq \sigma_0^2$	$\chi^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2}$	$\{\chi^2 < \chi_{1-\alpha/2}^2(n-1)\} \cup \{\chi^2 > \chi_{\alpha/2}^2(n-1)\}$
	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 < \sigma_0^2$		$\{\chi^2 < \chi_{1-\alpha}^2(n-1)\}$
	$\sigma^2 = \sigma_0^2 \longleftrightarrow \sigma^2 > \sigma_0^2$		$\{\chi^2 > \chi_{\alpha}^2(n-1)\}$
	$\sigma^2 \leq \sigma_0^2 \longleftrightarrow \sigma^2 > \sigma_0^2$		$\{\chi^2 > \chi_{\alpha}^2(n-1)\}$
	$\sigma^2 \geq \sigma_0^2 \longleftrightarrow \sigma^2 < \sigma_0^2$		$\{\chi^2 < \chi_{1-\alpha}^2(n-1)\}$

两正态总体参数的假设检验

两样本问题在实际中具有广泛的应用，下面我们先看一个例子.

例 4.3.1 为研究正常成年男女血液中红细胞平均数之差别，现在某地随机地抽取正常成年男子156人，正常成年女子74人，测得男性红细胞平均数为465.13万/mm³，标准差为54.8万/mm³，女性红细胞平均数为422.16万/mm³，样本标准差为49.2万/mm³. 请问此地区正常成年人的红细胞数与性别有关吗？($\alpha = 0.01$)

例4.3.1就是一个两样本问题，为了解答此问题，我们先考虑一般情形下的假设检验问题.

设 $X_1, \dots, X_m$ 为来自正态总体 $N(\mu_1, \sigma_1^2)$ 的iid 样本， $Y_1, \dots, Y_n$ 为来自正态总体 $N(\mu_2, \sigma_2^2)$ 的iid 样本，且全样本是独立的.

两正态总体参数的假设检验

§4.3.1 两样本正态总体均值的显著性检验

此时，我们感兴趣的假设为

$$H_0 : \mu_1 = \mu_2 \longleftrightarrow H_1 : \mu_1 \neq \mu_2, \quad (4.3.1)$$

$$H_0 : \mu_1 \leq \mu_2 \longleftrightarrow H_1 : \mu_1 > \mu_2, \quad (4.3.2)$$

$$H_0 : \mu_1 \geq \mu_2 \longleftrightarrow H_1 : \mu_1 < \mu_2. \quad (4.3.3)$$

两正态总体参数的假设检验

一、当 σ_1^2, σ_2^2 均已知时

❶ 对于假设(4.3.1), 由于 $\bar{X}, \bar{Y}$ 分别是 μ_1, μ_2 的一个很好的点估计, 故我们有理由用两样本均值差 $\bar{X} - \bar{Y}$ 来反映 μ_1 与 μ_2 的区别. 考虑到假设的特点, 我们知道, 当 $|\bar{X} - \bar{Y}| \geq c$ 时, 我们有理由拒绝 H_0 , 其中常数 c 由此检验的第一类错误概率 α 来决定.

由于当 H_0 成立时, 统计量

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0, 1), \quad (4.3.4)$$

且其拒绝域形如

$$W = \{|U| \geq c\},$$

故为控制其第一类错误概率为 α , 于是常数 c 应满足

$$P_{H_0}\{|U| \geq c\} \leq \alpha,$$

两正态总体参数的假设检验

这样，由(4.3.4)知，我们选取临界值 $c = u_{\alpha/2}$.

综合上述分析，我们得到此时假设(4.3.1)的水平为 α 的显著性检验为：

- 检验统计量： $U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$,
- 拒绝域： $\{|U| \geq u_{\alpha/2}\}$,

我们称之为两样本的双边 u 检验.

两正态总体参数的假设检验

② 对于假设(4.3.2)，我们可以利用上面的方法得到其水平 α 的显著性检验为

- 检验统计量: $U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$,

- 拒绝域: $\{U \geq u_\alpha\}$,

我们称之为两样本的单边 u 检验.

两正态总体参数的假设检验

我们可以证明，此检验的水平确为 α . 事实上，在 H_0 成立时，我们有

$$\begin{aligned} & P_{H_0}\{U \geq u_\alpha\} \\ &= P_{H_0}\left\{\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \geq u_\alpha\right\} \\ &= P_{H_0}\left\{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \geq u_\alpha - \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}\right\} \\ &\leq P_{H_0}\left\{\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \geq u_\alpha\right\} \quad (H_0 : \mu_1 \leq \mu_2) \\ &= \alpha. \quad (\text{因为其中的随机变量为标准正态}) \end{aligned}$$

两正态总体参数的假设检验

③ 对于假设(4.3.3)，我们利用类似于关于假设(4.3.2)的方法得到其水平为 α 的显著性检验为

- 检验统计量: $U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$,
- 拒绝域: $\{U \leq -u_\alpha\}$,

我们也称之为两样本的单边 u 检验.

两正态总体参数的假设检验

二、当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时

❶ 对于假设(4.3.1), 我们仍可以用 $\bar{X} - \bar{Y}$ 来估计 $\mu_1 - \mu_2$, 但是, 我们注意到, 当 H_0 成立时, $\bar{X} - \bar{Y} \sim N(0, (m+n)\sigma^2/mn)$, 而其中 σ^2 未知. 由第二章内容可知, 此时我们可以利用总的样本方差

$$S_{mn}^{*2} = \frac{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2}{m+n-2} \quad (4.3.5)$$

来估计 σ^2 , 且在 H_0 成立时, $(m+n-2)S_{mn}^{*2}/\sigma^2 \sim \chi^2(m+n-2)$. 于是, 一个自然的检验统计量为

$$T = \frac{(\bar{X} - \bar{Y})/\sigma\sqrt{(m+n)/(mn)}}{\sqrt{(m+n-2)S_{mn}^{*2}/(\sigma^2(m+n-2))}} = \sqrt{\frac{mn}{m+n}} \frac{\bar{X} - \bar{Y}}{S_{mn}^*}, \quad (4.3.6)$$

且当 H_0 成立时, $T \sim t(m+n-2)$.

另外, 为满足其显著性水平为 α , 我们可取拒绝域为

$$W = \{|T| \geq t_{\alpha/2}(m+n-2)\}. \quad (4.3.7)$$

综上所述, 此时双边假设(4.3.1)的水平 α 的显著性检验统计量由(4.3.6)式给出, 而其拒绝域由(4.3.7)给出. 我们称此检验法为两样本的 t 检验.

两正态总体参数的假设检验

② 对于假设(4.3.2), 我们仍可以利用检验统计量(4.3.6), 且可以证明其拒绝域可取为

$$\{T \geq t_{\alpha}(m+n-2)\}.$$

另外, 我们称此检验法为单边的两样本 t 检验.

③ 对于假设(4.3.3), 我们同样可以利用(4.3.6)给出的检验统计量, 且其拒绝域为

$$\{T \leq -t_{\alpha}(m+n-2)\}.$$

它也被称作单边的两样本 t 检验.

对于例4.3.1中的数据, 如假设两总体的方差相等且未知, 则我们可以利用本段方法进行显著性检验. 实际上, 此时我们感兴趣的是要检验(4.3.1)的双边假设. 由已知条件知: $S_{1m}^2 = 54.8$, $S_{2n}^2 = 49.20$, 于是, 因为

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{(m-1)S_{1m}^2 + (n-1)S_{2n}^2}} \sqrt{\frac{mn(m+n-2)}{m+n}} = 5.7 > 2.6 = t_{0.01/2}(228),$$

故我们有理由拒绝 H_0 , 即认为该地区成年男女性间的红细胞有显著的差别.

两正态总体参数的假设检验

三、一般情况

关于一般情况下两样本正态总体均值的显著性检验问题，我们在这里仅不加证明地给出其水平近似为 α 的显著性检验. 请有兴趣的读者参见陈希孺(1997)和茆诗松、王静龙(1990).

❶ 当 m, n 都充分大时

对于假设(4.3.1), (4.3.2), (4.3.3), 我们采用的统计量均为

$$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_{1m}^2}{m} + \frac{S_{2n}^2}{n}}}, \quad (4.3.8)$$

而它们的拒绝域分别为

$$\{|U| \geq u_{\alpha/2}\}, \quad \{U > u_{\alpha}\}, \quad \{U < -u_{\alpha}\}. \quad (4.3.9)$$

两正态总体参数的假设检验

② 当 m, n 都不是很大时

对于上述三个假设, 此时我们采用的统计量仍为(4.3.8), 只是由于样本比较小, 故无法利用大样本近似. 正如我们在讲置信区间时一样, 当 $\mu_1 = \mu_2$ 时, (4.3.8)式的统计量近似服从自由度为 r 的 t 分布, 其中

$$r = \frac{S_{mn}^4}{\frac{S_{1m}^4}{m^2(m-1)} + \frac{S_{2n}^4}{n^2(n-1)}}, \quad (4.3.10)$$

它与(3.3.5)一样, 其中 $S_{mn}^2 = \frac{S_{1m}^2}{m} + \frac{S_{2n}^2}{n}$.

另外, 对于假设(4.3.1), (4.3.2), (4.3.3)的拒绝域, 分别为

$$\{|U| \geq t_{\alpha/2}(r)\}, \quad \{U > t_{\alpha}(r)\}, \quad \{U < -t_{\alpha}(r)\}. \quad (4.3.11)$$

综上所述, 我们把两样本正态总体均值的显著性检验总结如下:

两正态总体参数的假设检验

表4.3 两样本正态总体均值的显著性检验

方差		假设	检验统计量	拒绝域
σ_1^2, σ_2^2 已知		$H_0: \mu_1 = \mu_2 \longleftrightarrow H_1: \mu_1 \neq \mu_2$ $H_0: \mu_1 \leq \mu_2 \longleftrightarrow H_1: \mu_1 > \mu_2$ $H_0: \mu_1 \geq \mu_2 \longleftrightarrow H_1: \mu_1 < \mu_2$	$U = \frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}}$	$\{ U \geq u_{\alpha/2}\}$ $\{U \geq u_{\alpha}\}$ $\{U \leq -u_{\alpha}\}$
$\sigma_1^2 = \sigma_2^2$ 未知		$H_0: \mu_1 = \mu_2 \longleftrightarrow H_1: \mu_1 \neq \mu_2$ $H_0: \mu_1 \leq \mu_2 \longleftrightarrow H_1: \mu_1 > \mu_2$ $H_0: \mu_1 \geq \mu_2 \longleftrightarrow H_1: \mu_1 < \mu_2$	$T = \sqrt{\frac{mn}{m+n}} \frac{\bar{X} - \bar{Y}}{S_{mn}^*}$	$\{ T \geq t_{\alpha/2}(m+n-2)\}$ $\{T \geq t_{\alpha}(m+n-2)\}$ $\{T \leq -t_{\alpha}(m+n-2)\}$
σ_1^2 σ_2^2 未知	m, n 充分大时	$H_0: \mu_1 = \mu_2 \longleftrightarrow H_1: \mu_1 \neq \mu_2$ $H_0: \mu_1 \leq \mu_2 \longleftrightarrow H_1: \mu_1 > \mu_2$ $H_0: \mu_1 \geq \mu_2 \longleftrightarrow H_1: \mu_1 < \mu_2$	$U = \frac{\bar{X} - \bar{Y}}{\sqrt{S_{1m}^2/m + S_{2n}^2/n}}$	$\{ U \geq u_{\alpha/2}\}$ $\{U \geq u_{\alpha}\}$ $\{U \leq -u_{\alpha}\}$
	m, n 不都充分大	$H_0: \mu_1 = \mu_2 \longleftrightarrow H_1: \mu_1 \neq \mu_2$ $H_0: \mu_1 \leq \mu_2 \longleftrightarrow H_1: \mu_1 > \mu_2$ $H_0: \mu_1 \geq \mu_2 \longleftrightarrow H_1: \mu_1 < \mu_2$	$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_{1m}^2/m + S_{2n}^2/n}}$	$\{ T \geq t_{\alpha/2}(r)\}$ $\{T \geq t_{\alpha}(r)\}$ $\{T \leq -t_{\alpha}(r)\}$

两正态总体参数的假设检验

§4.3.2 两样本正态总体方差的显著性检验

从上一节内容可以看出，两样本均值的检验方法与总体方差的大小有关。本节则将考虑两样本方差的检验问题。此时，我们感兴趣的假设为下面三个：

$$H_0 : \sigma_1^2 = \sigma_2^2 \longleftrightarrow H_1 : \sigma_1^2 \neq \sigma_2^2, \quad (4.3.12)$$

$$H_0 : \sigma_1^2 \leq \sigma_2^2 \longleftrightarrow H_1 : \sigma_1^2 > \sigma_2^2, \quad (4.3.13)$$

$$H_0 : \sigma_1^2 \geq \sigma_2^2 \longleftrightarrow H_1 : \sigma_1^2 < \sigma_2^2. \quad (4.3.14)$$

下面，我们将分均值已知与未知两种情况考虑上述三种假设的显著性检验。

两正态总体参数的假设检验

一、当 μ_1, μ_2 均已知时

❶ 对于假设(4.3.12)，由于此时 σ_1^2 与 σ_2^2 的合理估计分别为

$$S_{1m}'^2 = \frac{1}{m} \sum_{i=1}^m (X_i - \mu_1)^2, \quad S_{2n}'^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \mu_2)^2,$$

故一个合理的检验统计量为

$$F = \frac{S_{1m}'^2}{S_{2n}'^2} = \frac{\sum_{i=1}^m (X_i - \mu_1)^2 / m}{\sum_{i=1}^n (Y_i - \mu_2)^2 / n}, \quad (4.3.15)$$

并且由备选假设的特点知道，此时的拒绝域为

$$\{F \leq c_1\} \cup \{F \geq c_2\}, \quad (c_1 \leq c_2)$$

为控制其第一类错误概率不大于 α ，故我们要求 c_1, c_2 满足

$$P_{H_0}\{F \leq c_1\} + P_{H_0}\{F \geq c_2\} \leq \alpha.$$

两正态总体参数的假设检验

为了均衡上下两侧，我们通常要求 c_1, c_2 满足

$$P_{H_0}\{F \leq c_1\} \leq \frac{\alpha}{2}, \quad P_{H_0}\{F \geq c_2\} \leq \frac{\alpha}{2}. \quad (4.3.16)$$

因为在 H_0 成立时，(4.3.15)式的检验统计量 $F \sim F(m, n)$ ，故我们可以取 c_1, c_2 如下：

$$c_1 = F_{1-\alpha/2}(m, n), \quad c_2 = F_{\alpha/2}(m, n). \quad (4.3.17)$$

于是，我们得到假设(4.3.12)的水平为 α 的显著性检验为

- 检验统计量： $F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2 / m}{\sum_{i=1}^n (Y_i - \mu_2)^2 / n}$,
- 拒绝域： $\{F \leq F_{1-\alpha/2}(m, n)\} \cup \{F \geq F_{\alpha/2}(m, n)\}$,

我们称之为双边的 F 检验.

两正态总体参数的假设检验

② 对于假设(4.3.13), 我们可以证明, 其水平为 α 的显著性检验为

- 检验统计量: $F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2 / m}{\sum_{i=1}^n (Y_i - \mu_2)^2 / n}$,
- 拒绝域: $\{F \geq F_{\alpha}(m, n)\}$,

我们称之为单边的 F 检验.

③ 对于假设(4.3.14), 其水平为 α 的显著性检验为

- 检验统计量: $F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2 / m}{\sum_{i=1}^n (Y_i - \mu_2)^2 / n}$,
- 拒绝域: $\{F \leq F_{1-\alpha}(m, n)\}$,

我们也称之为单边的 F 检验.

两正态总体参数的假设检验

二、当 μ_1, μ_2 均未知时

对于上面的三个假设，我们采用的统计量为

$$F = \frac{S_{1m}^2}{S_{2n}^2} = \frac{\sum_{i=1}^m (X_i - \bar{X})^2 / (m - 1)}{\sum_{i=1}^n (Y_i - \bar{Y})^2 / (n - 1)}, \quad (4.3.18)$$

其拒绝域只需把上一段中的自由度换成 $m - 1, n - 1$ 即可.

两正态总体参数的假设检验

于是，我们把两样本正态总体方差的显著性总结如下：

表4.4 两样本正态总体方差的显著性检验

讨厌参数	假设	检验统计量	拒绝域
μ_1, μ_2 已知	$\sigma_1^2 = \sigma_2^2 \longleftrightarrow \sigma_1^2 \neq \sigma_2^2$	$F = \frac{\sum_{i=1}^m (X_i - \mu_1)^2 / m}{\sum_{i=1}^n (Y_i - \mu_2)^2 / n}$	$\{F \leq F_{1-\alpha/2}(m, n)\} \cup$ $\{F \geq F_{\alpha/2}(m, n)\}$
	$\sigma_1^2 \leq \sigma_2^2 \longleftrightarrow \sigma_1^2 > \sigma_2^2$		$\{F \geq F_{\alpha}(m, n)\}$
	$\sigma_1^2 \geq \sigma_2^2 \longleftrightarrow \sigma_1^2 < \sigma_2^2$		$\{F \leq F_{1-\alpha}(m, n)\}$
μ_1, μ_2 未知	$\sigma_1^2 = \sigma_2^2 \longleftrightarrow \sigma_1^2 \neq \sigma_2^2$	$F = \frac{\sum_{i=1}^m (X_i - \bar{X})^2 / (m-1)}{\sum_{i=1}^n (Y_i - \bar{Y})^2 / (n-1)}$	$\{F \leq F_{1-\alpha/2}(m-1, n-1)\} \cup$ $\{F \geq F_{\alpha/2}(m-1, n-1)\}$
	$\sigma_1^2 \leq \sigma_2^2 \longleftrightarrow \sigma_1^2 > \sigma_2^2$		$\{F \geq F_{\alpha}(m-1, n-1)\}$
	$\sigma_1^2 \geq \sigma_2^2 \longleftrightarrow \sigma_1^2 < \sigma_2^2$		$\{F \leq F_{1-\alpha}(m-1, n-1)\}$

两正态总体参数的假设检验

例 4.3.3 有甲乙两台机床加工不同产品，从此两台机床加工的产品中随机地抽取若干件，测得其产品直径(单位: mm)如下:

甲: 30.5, 29.8, 29.7, 20.4, 20.1, 20.0, 19.0, 19.9

乙: 16.7, 17.8, 17.5, 16.8, 16.4, 17.6, 16.2

假设甲乙两台机床加工产品的直径分别服从正态分布 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$. 问: 这两台机床加工的精度有无显著性差异? ($\alpha = 0.05$)

解 对于此问题，此时的假设为(4.3.12)式的双边假设，且均值未知. 因为

$$F = \frac{\sum_{i=1}^m (X_i - \bar{X})^2 / (m - 1)}{\sum_{i=1}^n (Y_i - \bar{Y})^2 / (n - 1)} = 69.13 > 5.7 = F_{0.025}(7, 6),$$

故我们有理由认为这两台机床加工的精度有显著的区别.

□

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

似然比检验

我们再分析一下前面讲的显著性检验，它处理问题的思路是：先假设总体服从某参数分布族 $f(x, \theta)$ ，之后根据实际问题提出假设 $H_0 \longleftrightarrow H_1$ ，然后再利用样本及总体信息提出检验统计量及拒绝域，最后通过控制第一类错误概率求得具体的拒绝域。从表4.1-4.5可以看出，我们考虑的假设均是特定的区间型的，而不是下面的非常一般的假设：

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1 = \Theta - \Theta_0. \quad (4.5.1)$$

对于有如(4.5.1)的假设，我们很难由显著性检验的步骤确定拒绝域。但如从似然的角度看，我们则可以得到如下的似然比检验方法。

似然比统计量

定义 4.5.1 (似然比统计量) 设 $X_1, \dots, X_n$ 为来自分布族 $\mathcal{F} = \{f(x, \theta) : \theta \in \Theta\}$ 的iid样本, 对于感兴趣的假设(4.5.1), 令

$$\lambda(\mathbf{X}) = \frac{\sup_{\theta \in \Theta_0} f(\mathbf{X}, \theta)}{\sup_{\theta \in \Theta} f(\mathbf{X}, \theta)}. \quad (4.5.2)$$

则我们称统计量 $\lambda(\mathbf{X})$ 为假设(4.5.1)的似然比(Likelihood Ratio), 有时也称之为广义似然比.

从 $\lambda(\mathbf{X})$ 的定义不难看出, 如果 $\lambda(\mathbf{X})$ 的值很小, 则说明 $\theta \in \Theta_0$ 的可能性要比 $\theta \in \Theta$ 的可能性小, 于是, 我们有理由认为 H_0 不成立. 这样, 我们有如下的似然比检验.

似然比检验

定义 4.5.2 (似然比检验) 当采用(4.5.2)式的似然比统计量 $\lambda(\mathbf{X})$ 作为假设(4.5.1)的检验统计量, 且取其拒绝域为 $\{\lambda(\mathbf{x}) \leq c\}$ 时, 其中临界值 c 满足

$$P_{\theta}\{\lambda(\mathbf{X}) \leq c\} \leq \alpha, \forall \theta \in \Theta_0, \quad (4.5.3)$$

则称此检验为水平 α 的似然比检验(Likelihood Ratio Test, 简记为LRT).

注 4.5.1 如果似然比统计量 $\lambda(\mathbf{X})$ 的零分布未知, 则我们很难由(4.5.3)式确定似然比检验的临界值. 但如果此时存在一个统计量 $T(\mathbf{X})$ 关于 $\lambda(\mathbf{X})$ 是单调的且它的零分布已知, 则我们可以给出一个基于 $T(\mathbf{X})$ 的显著性检验.

似然比检验

例 4.5.1 设 $X_1, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的iid样本, μ, σ^2 均未知. 试求假设

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

的水平为 α 的似然比检验.

似然比检验

解 此时样本分布为

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

其参数空间为

$$\Theta_0 = \{(\mu_0, \sigma^2) : \sigma^2 > 0\}, \quad \Theta = \{(\mu, \sigma^2) : \mu \in R, \sigma^2 > 0\}.$$

利用微分法，我们容易求得

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

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

似然比检验

于是，其似然比统计量为

$$\begin{aligned}\lambda(\mathbf{x}) &= \frac{\sup_{\theta \in \Theta_0} f(\mathbf{x}, \theta)}{\sup_{\theta \in \Theta} f(\mathbf{x}, \theta)} = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2} \right)^{n/2} \\ &= \left[\frac{1}{\frac{\sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right]^{n/2} = \left[\frac{1}{1 + n \frac{(\bar{x} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right]^{n/2} \\ &= \left[1 + \frac{T^2}{n-1} \right]^{-n/2},\end{aligned}$$

其中 $T = \frac{\sqrt{n}(\bar{x} - \mu_0)}{S_n}$.

另外，从上式可知，此时的似然比统计量与传统的 t 统计量的平方成反比，于是，两个检验统计量的拒绝域有如下关系：

$$\{\lambda(\mathbf{x}) \leq c\} \iff \{|T(\mathbf{x})| \geq d\}.$$

又因为当 H_0 成立时， $T \sim t(n-1)$ ，故我们取 $d = t_{\alpha/2}(n-1)$ 即可控制其第一类错误概率不超过 α 。由此可见，此时的似然比检验与我们前面讲过的双边 t 检验完全等价。□

似然比检验

例 4.5.2 设 $X_1, \dots, X_n$ 为来自具有如下PDF

$$f(x, \mu) = e^{-(x-\mu)}, \quad x \geq \mu, \quad \mu \in R$$

的总体的iid样本, 试求假设

$$H_0 : \mu = 0 \longleftrightarrow H_1 : \mu \neq 0$$

的水平 α 的似然比检验.

似然比检验

解 此时的样本分布为

$$f(\mathbf{x}, \mu) = e^{-\sum_{i=1}^n (x_i - \mu)} I_{(x_{(1)} \geq \mu)},$$

且参数空间为 $\Theta = R, \Theta_0 = \{0\}$.

于是, 可以求得

$$\sup_{\theta \in \Theta_0} f(\mathbf{x}, \mu) = e^{-\sum_{i=1}^n x_i}, \quad \sup_{\theta \in \Theta} f(\mathbf{x}, \mu) = e^{-\sum_{i=1}^n x_i + nx_{(1)}}.$$

这样, 此时的似然比统计量为

$$\lambda(\mathbf{X}) = e^{-nX_{(1)}} = e^{-\frac{1}{2}(2nX_{(1)})},$$

它等价于统计量 $2nX_{(1)}$, 且其拒绝域为 $\{2nX_{(1)} \geq c\}$.

又因为在 H_0 成立时, $2nX_{(1)} \sim \chi^2(2)$ (计算其PDF即知), 故其临界值 $c = \chi_{\alpha}^2(2)$. 即此时的似然比检验统计量为 $2nX_{(1)}$, 其拒绝域为 $\{2nX_{(1)} \geq \chi_{\alpha}^2(2)\}$. □

似然比检验

例 4.5.3 设总体分布为

$$P\{X = i\} = p_i, \quad i = 1, \dots, r,$$

其中 $p_i > 0, \sum_{i=1}^r p_i = 1$. 设有 n 个来自此分布的 iid 样本 $X_1, \dots, X_n$, 试求

$$H_0 : p_i = p_{i0}, \quad i = 1, \dots, r \quad (4.5.4)$$

的似然比检验, 其中 p_{i0} 已知, 且 $\sum_{i=1}^r p_{i0} = 1$.

似然比检验

解 对于此问题, 我们注意到, 其总体是一个离散分布, 且参数空间 Θ 是 $r - 1$ 维, 而 Θ_0 只包含一个点.

另外, 为方便, 我们以 n_i 表示样本中取值 i 的个数, 即 $\sum_{i=1}^r n_i = n$. 此时样本分布为

$$P_{\theta}\{\mathbf{X} = \mathbf{x}\} = p_1^{n_1} \cdots p_r^{n_r},$$

由此可求得

$$\sup_{\theta \in \Theta} P_{\theta}\{\mathbf{X} = \mathbf{x}\} = \prod_{i=1}^r \left(\frac{n_i}{n}\right)^{n_i}, \quad \sup_{\theta \in \Theta_0} P_{\theta}\{\mathbf{X} = \mathbf{x}\} = \prod_{i=1}^r p_{i0}^{n_i}.$$

于是, 其似然比统计量为

$$\lambda(\mathbf{X}) = \prod_{i=1}^r \left(\frac{np_{i0}}{n_i}\right)^{n_i}.$$

似然比检验

另外, 由于

$$\begin{aligned}\ln \lambda(\mathbf{X}) &= \sum_{i=1}^r n_i \ln \frac{np_{i0}}{n_i} \\&= - \sum_{i=1}^r [np_{i0} + (n_i - np_{i0})] \ln \left(1 + \frac{n_i - np_{i0}}{np_{i0}} \right) \\&\approx - \sum_{i=1}^r [np_{i0} + (n_i - np_{i0})] \left[\frac{n_i - np_{i0}}{np_{i0}} - \frac{1}{2} \left(\frac{n_i - np_{i0}}{np_{i0}} \right)^2 + o_p(n^{-2}) \right] \\&\approx - \frac{1}{2} \sum_{i=1}^r \frac{(n_i - np_{i0})^2}{np_{i0}} + o_p(n^{-1}),\end{aligned}$$

所以,

$$-2 \ln \lambda(\mathbf{X}) \approx \sum_{i=1}^r \frac{(n_i - np_{i0})^2}{np_{i0}}. \quad (4.5.5)$$

又因为 $-2 \ln \lambda(\mathbf{X})$ 是 $\lambda(\mathbf{X})$ 的严减函数, 故似然比检验等价于用检验统计量 $-2 \ln \lambda(\mathbf{X})$, 且拒绝域为 $\{-2 \ln \lambda(\mathbf{X}) \geq c\}$.

似然比检验

总之，对于此问题，我们可采用检验统计量 $\sum_{i=1}^r \frac{(n_i - np_{i0})^2}{np_{i0}}$ ，且其拒绝域为

$$\left\{ (n_1, \dots, n_r) : \sum_{i=1}^r \frac{(n_i - np_{i0})^2}{np_{i0}} \geq c \right\}. \quad (4.5.6)$$

□

我们在这里看不出这个统计量有什么特别之处，但是我们在后面的第6章可以看出，这是K. Pearson引进的 χ^2 拟合优度检验统计量，它在许多实际问题中都得到了广泛的应用。值得注意的是，此时检验统计量的精确零分布很难导出，也就是说我们无法给出一个精确的临界值 c 使得 $P_{H_0}\{-2\ln\lambda(\mathbf{X}) \geq c\} = \alpha$ 。在当今众多的统计应用问题中，我们通常都是难以得到似然比检验统计量的精确零分布的。此时，统计大样本理论可以帮助我们得到某种近似（极限）分布，这就是著名的Wilks定理：假设在 H_0 下我们对 k 维参数向量 $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)^T$ 有 r 个约束（则 $\boldsymbol{\theta}$ 中只有 $k - r$ 个分量可以是自由的）。在适当条件下， $-2\ln\lambda(\mathbf{X})$ 的极限零分布是 χ_r^2 分布，即收敛到自由度是备选假设和原假设下自由参数个数之差的卡方分布。该定理是由Wilks在1938年的数理统计年刊上完整阐述的，后人也通常称似然比检验统计量的这一特点为Wilks现象。

似然比检验

定理 2.7.9 设参数空间 Θ 是开区间, 总体PDF $f(x; \theta)$ 满足

- (1) 在参数真值 θ_0 的邻域内, $\partial^i \ln f / \partial \theta^i$, $i = 1, 2, 3$ 存在;
- (2) 在参数真值 θ_0 的邻域内, $|\partial^3 \ln f / \partial \theta^3| \leq H(x)$, 且 $EH(X) < \infty$;
- (3) 在参数真值 θ_0 处,

$$E_{\theta_0} \left[\frac{f'(X, \theta_0)}{f(X, \theta_0)} \right] = 0, \quad E_{\theta_0} \left[\frac{f''(X, \theta_0)}{f(X, \theta_0)} \right] = 0, \quad I(\theta_0) = E_{\theta_0} \left[\frac{f'(X, \theta_0)}{f(X, \theta_0)} \right]^2 > 0.$$

记 $\hat{\theta}_n$ 为当 $n \rightarrow \infty$ 时似然方程的相合解, 则

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N(0, I^{-1}(\theta_0)). \quad (2.7.9)$$

定理 4.5.1 假设定理2.7.9中的条件成立且 H_0 由(4.5.1)表示, 则在 H_0 下, $-2 \ln \lambda(\mathbf{X}) \xrightarrow{d} \chi_r^2$.

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

假设检验基本概念

针对假设 $H_0 \longleftrightarrow H_1$, 一个检验函数或检验法则或检验, 就是设法把样本空间 $\mathcal{X}$ 划分为两个互不相交的可测集:

$$\mathcal{X} = W \cup \bar{W}.$$

当观测值 $\mathbf{x} \in W$ 时, 拒绝原假设 H_0 ; 当 $\mathbf{x} \in \bar{W}$ 时, 就不能拒绝原假设 H_0 . 其中的 W 称为拒绝域. 即此时的检验或检验函数被定义成如下的函数.

对于拒绝域 $W \subset \mathcal{X}$, 定义

$$\phi(\mathbf{X}) = \begin{cases} 1, & \mathbf{X} \in W, \\ 0, & \mathbf{X} \in \bar{W}, \end{cases}$$

它是 W 的示性函数, 且仅取0, 1两个值. 反之, 如果对于一个仅取0, 1两个值的函数 $\phi(\mathbf{X})$, 则 $W = \{\mathbf{X} : \phi(\mathbf{X}) = 1\}$ 也可作为拒绝域. 这样的函数 $\phi(x)$ 就称为检验函数或检验, 其势函数为

$$\beta_{\phi}(\theta) = P_{\theta}\{\mathbf{X} \in W\} = E_{\theta}\phi(\mathbf{X}). \quad (5.1.1)$$

检验

上述定义是我们在上一章中所采用的，但是对于本章要讲的优势检验而言，我们有必要重新定义检验如下：

定义 5.1.1 (检验) 设 $\phi(\mathbf{X})$ 是定义在 $\mathcal{X}$ 上的可测函数，满足 $0 \leq \phi(\mathbf{x}) \leq 1$ ，则称 $\phi(\mathbf{X})$ 为检验函数，简称检验。如果 $\phi(\mathbf{x})$ 仅取0, 1两值时，则称为非随机化检验，否则，就称为随机化检验，其势函数为 $\beta_{\phi}(\theta) = E_{\theta}\phi(\mathbf{X})$ 。

从上述定义可以看出，上一章我们称拒绝域的示性函数为一个检验，而本章则不然：由定义5.1.1所定义的检验可以在 $[0, 1]$ 间取值，它可能不是拒绝域的示性函数。

对于一个随机化检验 $\phi(x)$ ，当其取值为1时，我们拒绝原假设；当其取值为0时，我们不能拒绝原假设；当其取值为 $\delta \in (0, 1)$ 时，我们如下处理：取一个来自于 $b(1, \delta)$ 的随机数，如此随机数为1，则拒绝原假设，否则不能拒绝原假设。

另外，当检验统计量为连续随机变量时，我们采用的是非随机化检验。当检验统计量为离散随机变量时，就有可能采用随机化检验。这一点，我们将在后面接触到。

假设检验的充分性原则

既然我们用两类错误概率来衡量一个检验的好坏，则我们可以利用其势函数把检验分类。

定义 5.1.2 设 $\phi_1(\mathbf{X})$ 和 $\phi_2(\mathbf{X})$ 是某检验问题 $H_0 \longleftrightarrow H_1$ 的检验函数，如果它们的势函数相同，即

$$E_{\theta}\phi_1(\mathbf{X}) = E_{\theta}\phi_2(\mathbf{X}), \forall \theta \in \Theta,$$

则称检验函数 $\phi_1(\mathbf{X})$ 和 $\phi_2(\mathbf{X})$ 等价。

前面的知识告诉我们，充分统计量在估计中非常有用，由于它完全包含了样本中关于参数的信息，那它在检验中又如何呢？见下面的定理。

定理 5.1.1 设 $X_1, \dots, X_n$ 是来自分布族 $\{f(x, \theta) : \theta \in \Theta\}$ 的样本。 $T(\mathbf{X})$ 是参数 θ 的充分统计量。则对于任意一个检验函数 $\phi(\mathbf{X})$ ，存在另一个只依赖于 $T(\mathbf{X})$ 的检验函数与它等价。

这个定理告诉我们，当 θ 的充分统计量存在时，关于此参数的检验问题，我们仅需在由充分统计量构成的检验函数中寻找就可以了。这就是假设检验中的“充分性原则”。

MP检验

定义 5.1.3 (MP检验) 对于参数分布族 $\mathcal{F} = \{f(x, \theta) : \theta \in \Theta\}$, 设我们感兴趣的假设为

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta = \theta_1 \ (\theta_1 \neq \theta_0), \quad (5.1.2)$$

并设有两个水平为 α 的检验函数 $\phi_1(\mathbf{X}), \phi_2(\mathbf{X})$, 即满足

$$E_{\theta_0} \phi_i(\mathbf{X}) \leq \alpha, \quad i = 1, 2, \quad (5.1.3)$$

如果

$$\beta_{\phi_1}(\theta_1) \geq \beta_{\phi_2}(\theta_1), \quad (5.1.4)$$

则称检验 ϕ_1 比 ϕ_2 有效. 如果检验 ϕ_1 对于任一个水平小于等于 α 的检验 ϕ_2 , (5.1.4)式均成立, 则称 ϕ_1 是假设(5.1.2)的水平 α 的最优势检验(Most Powerful Test, 简记为MPT).

从上述定义可以看出, 一个MP检验, 不仅控制其第一类错误概率, 而且也控制其第二类错误概率, 且比一般水平 α 的显著性检验有效. 那这种好的检验存在吗? 请见下节的N-P引理.

Neyman-Pearson基本引理

定理 5.1.2 (Neyman-Pearson基本引理) 对于参数分布族 $\{f(x, \theta) : \theta \in \Theta = \{\theta_0, \theta_1\}\}$, 则关于检验问题(5.1.2), 我们有如下结论:

(1) 对给定的 $\alpha \in (0, 1)$, 存在一个检验函数 $\phi(x)$ 及常数 $k \geq 0$, 使得

$$\phi(x) = \begin{cases} 1, & f(x; \theta_1) > k f(x; \theta_0), \\ 0, & f(x; \theta_1) < k f(x; \theta_0). \end{cases} \quad (5.2.1)$$

且

$$E_{\theta_0} \phi(X) = \alpha. \quad (5.2.2)$$

(2) 由(5.2.1)和(5.2.2)式确定的检验函数 $\phi(x)$ 是检验问题(5.1.2)的水平为 α 的MPT.

(3) 如果 $\phi'(x)$ 是此检验问题的水平 α 的MPT, 则一定存在常数 $k \geq 0$, 使得 $\phi'(x)$ 满足(5.2.1)式. 又如 $\phi'(x)$ 满足 $E_{\theta_1} \phi'(X) < 1$, 则它也满足(5.2.2)式.

Neyman-Pearson基本引理

- 从N-P引理的证明可以看出，MPT是似然比统计量 $\lambda(X) = f(X; \theta_1)/f(X; \theta_0)$ 的函数，这种检验我们也称之为似然比检验；
- 如果似然比 $\lambda(X)$ 的分布是连续的，则假设(5.1.2)的MPT是非随机化的检验；如果 $\lambda(X)$ 的分布是离散的，则假设(5.1.2)的MPT有可能是随机化的；
- 对于非参数假设，N-P引理仍是正确的。

MPT

例 5.2.1 设 $X_1, \dots, X_n$ 是来自 $N(\mu, 1)$ 的 IID 样本, 试考虑假设

$$H_0 : \mu = 0 \longleftrightarrow H_1 : \mu = \mu_1 (> 0) \quad (5.2.3)$$

的水平 α 的 MPT.

MPT

解 此时样本的联合PDF为

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

由此可求得似然比统计量为

$$\lambda(x) = \exp\{n\mu_1 \bar{x} - n\mu_1^2/2\},$$

由此可见, $\lambda(x)$ 与 $\bar{x}$ 成正比, 于是, 由N-P引理知, 水平 α 的MPT为

$$\phi(X) = \begin{cases} 1, & \bar{X} > k, \\ 0, & \bar{X} < k, \end{cases}$$

并满足 $E_{\mu=0}\phi(X) = \alpha$. 又由于在 H_0 成立时, $\bar{X} \sim N(0, 1/n)$, 故上面MPT中的 k 满足

$$\alpha = P_{H_0}\{\bar{X} > k\} = P_{H_0}\{\sqrt{n}\bar{X} > \sqrt{n}k\} = 1 - \Phi(\sqrt{n}k),$$

即 $k = u_\alpha / \sqrt{n}$.

MPT

综上所述，知此假设的水平 α 的MPT为

$$\phi(X) = \begin{cases} 1, & \bar{X} > u_\alpha/\sqrt{n}, \\ 0, & \bar{X} < u_\alpha/\sqrt{n}. \end{cases} \quad (5.2.4)$$

□

关于例5.2.1，我们有如下几点注解：

- (5.2.4)式的MPT只与 α 及 μ_1 的符号有关，而与 μ_1 的具体数值无关.
- 如果把例5.2.1中的备选假设改为 $H_1 : \mu = \mu_1 < 0$ ，则其MPT的拒绝域为

$$W = \{x : \bar{x} \leq -u_\alpha/\sqrt{n}\},$$

它也仅与 μ_1 的符号有关，而与其具体数值无关.

- 若上述总体的分布为 $N(\mu, \sigma_0^2)$ ，其中 σ_0^2 已知，且我们感兴趣的假设为

$$H_0 : \mu = \mu_0 \longleftrightarrow H_0 : \mu = \mu_1 (> \mu_0),$$

MPT

则可能通过变换 $y_i = (x_i - \mu_0)/\sigma_0$ 将此假设化为例5.2.1中的假设, 因此其水平为 α 的MPT的拒绝域为

$$W = \left\{ \frac{\bar{x} - \mu_0}{\sigma_0} \geq \frac{u_\alpha}{\sqrt{n}} \right\}.$$

- 如我们感兴趣的假设为如下的简单对复合假设:

$$H_0 : \mu = 0 \longleftrightarrow H_2 : \mu > 0, \quad (5.2.5)$$

则由于(5.2.4)式的检验 $\phi(X)$ 是假设(5.2.3)的水平 α 的MPT, 故对于(5.2.3)的任一水平为 α 的检验 $\phi'(X)$, 均有

$$E_{\mu_1} \phi(X) \geq E_{\mu_1} \phi'(X). \quad (**1)$$

我们注意到, (5.2.4)式的 $\phi(X)$ 与 μ_1 无关, 故它也是(5.2.5)的水平 α 的一个检验, 且对于(5.2.5)的任一水平为 α 的检验 $\phi''(X)$, 由于它也是(5.2.3)的水平为 α 的一个检验, 故由(**1)式知,

$$E_{\mu_1} \phi(X) \geq E_{\mu_1} \phi''(X), \quad \forall \mu_1 > 0, \quad (**2)$$

故由(**2)式知, $\phi(X)$ 也是(5.2.5)的水平 α 的MPT, 关于 $\mu > 0$ 是一致的.

MPT

例 5.2.2 设 $X_1, \dots, X_n$ 为来自均匀分布 $U(0, p)$ 的IID样本, 考虑假设

$$H_0 : p = 1 \longleftrightarrow H_1 : p = p_1 (> 1)$$

的水平 α 的MPT.

MPT

解 此时的似然比统计量为

$$\lambda(x) = \frac{\prod_{i=1}^n f(x; p_1)}{\prod_{i=1}^n f(x; p_0)} = \begin{cases} p_1^{-n}, & 0 < x_{(n)} < 1, \\ \infty, & 1 \leq x_{(n)} < p_1. \end{cases}$$

于是, $\lambda(X)$ 的分布函数

$$h(c) = P_{H_0} \{ \lambda(X) \leq c \}$$

是一个退化分布, 即当 $c < p_1^{-n}$ 时, $h(c) = 0$; 当 $c \geq p_1^{-n}$ 时, $h(c) = 1$. 这样根据N-P引理的证明, 我们取 $k = p_1^{-n}$, 此时有

$$0 = h(k - 0) < 1 - \alpha < h(k) = 1$$

则由N-P引理知, 此时的MPT是如(*2)式的随机化检验, 即

$$\phi(x) = \begin{cases} 1, & 1 \leq x_{(n)} < p_1, \\ \alpha, & 0 < x_{(n)} < 1, \\ 0, & \text{其它.} \end{cases} \quad (**)$$

MPT

由于 $X_{(n)}$ 是连续的, 故我们也可以取如下的非随机检验(也可以看成是把(**)或的随机部分非随机化):

$$\phi(x) = \begin{cases} 1, & c \leq x_{(n)} < p_1, \\ 0, & 0 < x_{(n)} < c, \end{cases}$$

其中常数 c 由 $E_{H_0}\phi(X) = \alpha$ 决定. 实际上, 当 H_0 成立时, $X_{(n)}$ 的PDF为 $p(t) = nt^{n-1}$, $t \in (0, 1)$, 故常数 c 满足

$$\alpha = \int_c^1 nt^{n-1}dt,$$

由此求得 $c = \sqrt[n]{1-\alpha}$, 即这个非随机化的MPT为

$$\phi(x) = \begin{cases} 1, & \sqrt[n]{1-\alpha} \leq x_{(n)} < p_1, \\ 0, & 0 < x_{(n)} < \sqrt[n]{1-\alpha}. \end{cases} \quad (***)$$

我们注意到(**)与(***)式的区别在于把随机化检验非随机化.

□

MPT

对于例5.2.2, 如果我们感兴趣的检验为

$$H_0 : p = p_0 \longleftrightarrow H_1 : p = p_1 (> p_0)$$

则可以通过变换 $y_i = x_i/p_0$ 来把上述问题变成例5.2.2的问题. 此时其水平为 α 的非随机化MP检验为

$$\phi(x) = \begin{cases} 1, & \sqrt[n]{1-\alpha} \leq x_{(n)}/p_0 < p_1/p_0, \\ 0, & 0 < x_{(n)}/p_0 < \sqrt[n]{1-\alpha}. \end{cases}$$

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

一致最优势检验

定义 5.3.1 (一致最优势检验) 对于参数分布族 $\{f(x, \theta) : \theta \in \Theta\}$, 设感兴趣的假设为

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1. \quad (5.3.1)$$

如果对于两个水平为 α 的检验 ϕ_1, ϕ_2 , 即

$$E_{\theta}\phi_i \leq \alpha, \quad \forall \theta \in \Theta_0, \quad i = 1, 2, \quad (5.3.2)$$

且满足

$$\beta_{\phi_1}(\theta) = E_{\theta}\phi_1 \geq E_{\theta}\phi_2 = \beta_{\phi_2}(\theta), \quad \forall \theta \in \Theta_1, \quad (5.3.3)$$

则称检验 ϕ_1 一致优于检验 ϕ_2 . 如果存在一个检验 ϕ_1 , 使对任何水平为 α 的检验 ϕ_2 , 均有(5.3.3)成立, 则称检验 ϕ_1 是假设(5.3.1)的水平 α 的一致最优势(Uniformly Most Powerful, 简记为UMP)检验.

通过对定义5.1.3与定义5.3.1的比较, 我们发现, MP检验仅是UMP检验针对简单假设对简单假设的一种特殊情况. 另外, 这里所说的一致性是关于备选假设下的参数空间而言的.

一致最优势检验

定理 5.3.1 设有如下三个检验问题

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1, \quad (5.3.4)$$

$$H_0 : \theta \in \Theta_{01} \longleftrightarrow H_1 : \theta \in \Theta_1, \quad (5.3.5)$$

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta = \theta_1, \quad (5.3.6)$$

其中 $\Theta_{01} \subset \Theta_0$, $\theta_1 \in \Theta_1$, 并设 $\phi(x)$ 是检验问题(5.3.4)的水平 α 的检验, 则

(1) 如果 $\phi(x)$ 是检验问题(5.3.5)的水平 α 的UMPT, 则 $\phi(x)$ 也是(5.3.4)的水平 α 的UMPT.

(2) $\phi(x)$ 是检验问题(5.3.4)的水平 α 的UMPT的充要条件是, 对任意的 $\theta_1 \in \Theta_1$, $\phi(x)$ 是检验问题(5.3.6)的水平 α 的MPT.

一致最优势检验

定理 5.3.2 设 $\phi(x)$ 是(5.3.4)的水平 α 的检验, 如果对某个 $\theta_0 \in \Theta_0$ 及任意一个 $\theta_1 \in \Theta_1$, $\phi(x)$ 都是假设

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta = \theta_1 \quad (5.3.7)$$

的水平 α 的MPT, 则 $\phi(x)$ 也是(5.3.4)的水平 α 的UMPT.

一致最优势检验

综合定理5.3.2与例5.2.1, 我们知道, 在例5.2.1中给出的MP检验也是相应假设的UMP, 即, 对于正态总体 $N(\mu, 1)$, 假设

$$H_0 : \mu = 0 \longleftrightarrow H_1 : \mu > 0$$

的水平为 α 的UMPT为

$$\phi(X) = \begin{cases} 1, & \bar{X} > u_\alpha/\sqrt{n}, \\ 0, & \bar{X} < u_\alpha/\sqrt{n}. \end{cases}$$

而假设

$$H_0 : \mu = 0 \longleftrightarrow H_1 : \mu < 0$$

的水平为 α 的UMPT为

$$\phi(X) = \begin{cases} 1, & \bar{X} < -u_\alpha/\sqrt{n}, \\ 0, & \bar{X} > -u_\alpha/\sqrt{n}. \end{cases}$$

MLR分布族

为了讨论单边假设的UMP检验,我们先引入如下的单调似然比(Monotone Likelihood Ratio, 简记为MLR)分布族.

定义 5.3.2 (MLR分布族) 一个参数分布族 $\{f(x; \theta) : \theta \in \Theta\}$ 称为关于统计量 $T(X)$ 单增(减)似然比分布族, 如果它满足

- (1) Θ 是直线上的一个区间;
- (2) 对于 $\theta_1 \neq \theta_2$, 有 $f(x, \theta_1) \neq f(x, \theta_2)$;
- (3) 似然比 $\lambda(x) = f(x; \theta_2)/f(x; \theta_1)$ ($\theta_2 > \theta_1$)是统计量 $T(x)$ 的非降(增)函数.

单增或减似然比分布族统称为MLR分布族.

MLR分布族

MLR分布族与单参数指数型分布之间的关系如下：

对于单参数指数型分布

$$f(x; \theta) = c(\theta) \exp\{Q(\theta)T(x)\}h(x),$$

其中 $c(\theta) > 0$ ，如果 $Q(\theta)$ 是 θ 的严格单增(减)函数，则它必为关于其充分统计量的单增(减)似然比分布族。

于是，我们常用的分布，如二项分布、负二项分布、Poisson分布、单参数正态分布、指数分布等均是MLR分布族。另外，虽然均匀分布 $U(0, \theta)$ 并不是单参数指数型分布，但如定义 $0/0 = \infty$ ，则它也是关于其充分统计量 $X_{(n)}$ 的MLR分布族。

单边假设的UMP检验

§5.3.3 单边假设的UMP检验

在本小节，我们将考虑如下单边假设

$$H_0 : \theta \leq \theta_0 \longleftrightarrow H_1 : \theta > \theta_0, \quad (5.3.7)$$

$$H_0 : \theta \geq \theta_0 \longleftrightarrow H_1 : \theta < \theta_0 \quad (5.3.8)$$

的UMP检验，并且假设其分布族为MLR分布族.

单边假设的UMP检验

定理 5.3.4 设单参数分布族 $\{f(x; \theta) : \theta \in \Theta \subset \mathcal{R}\}$ 是关于统计量 $T(x)$ 非降的MLR分布族, 则关于单边假设(5.3.7),

(1) 存在水平为 α 的UMP检验函数

$$\phi(T) = \begin{cases} 1, & T(x) > c, \\ \delta, & T(x) = c, \\ 0, & T(x) < c, \end{cases} \quad (5.3.9)$$

其中常数 δ, c 由下式决定

$$E_{\theta_0} \phi(T(X)) = \alpha. \quad (5.3.10)$$

(2) 势函数 $\beta(\theta) = E_{\theta} \phi(T(X))$ 是非降的, 且在集合 $\{\theta : 0 < \beta(\theta) < 1\}$ 上严格单增.

(3) 如 $\phi'(x)$ 是假设(5.3.7)的一个检验, 且满足 $E_{\theta_0} \phi'(X) = \alpha < 1$, 则

$$E_{\theta} \phi(T(X)) \leq E_{\theta} \phi'(X), \quad \forall \theta < \theta_0.$$

单边假设的UMP检验

例 5.3.4 设 $X_1, \dots, X_n$ 是来自正态总体 $N(0, \sigma^2)$ 的IID样本, 试求假设

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

的水平 α 的UMPT.

单边假设的UMP检验

解 此时样本的联合PDF为

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

于是, $Q(\sigma^2) = -1/2\sigma^2$ 为 σ^2 的单增函数, 故它是关于 $T(x) = \sum x_i^2$ 的单增似然比分布族. 由定理5.3.4知, 此假设的UMPT为

$$\phi(T(x)) = \begin{cases} 1, & T(x) \geq c, \\ 0, & T(x) < c, \end{cases}$$

其中 c 满足 $E_{\sigma_0^2} \phi(T(X)) = \alpha$. 注意到, 当 $\sigma^2 = \sigma_0^2$ 时, $T(X)/\sigma_0^2 \sim \chi^2(n)$, 于是, 常数 c 满足

$$\int_{c/\sigma_0^2}^{\infty} \chi^2(x; n) dx = \alpha,$$

于是, $c = \sigma_0^2 \chi_{\alpha}^2(n)$.

单边假设的UMP检验

当 $\sigma_0 = 1, n = 10, \alpha = 0.1$ 时, 此UMP检验的拒绝域为

$$W = \left\{ x : \sum_{i=1}^n x_i^2 \geq \sigma_0^2 \chi_{\alpha}^2(n) \right\} = \left\{ x : \sum_{i=1}^{10} x_i^2 \geq 15.99 \right\}.$$

另外, 如我们用统计量 $T_1(x) = \sum_{i=1}^{10} (x_i - \bar{x})^2$ 来检验本假设, 则其拒绝域为

$$W_1 = \left\{ x : \sum_{i=1}^n (x_i - \bar{x})^2 \geq \sigma_0^2 \chi_{\alpha}^2(n-1) \right\} = \left\{ x : \sum_{i=1}^{10} (x_i - \bar{x})^2 \geq 14.68 \right\},$$

记此时的检验函数为 ϕ_1 . 我们不难验证它也是本假设的一个水平为 $\alpha = 0.1$ 的检验.

单边假设的UMP检验

由于，上述两个检验的势分别为

$$\beta_{\phi}(\sigma^2) = P \left\{ X : \sum_{i=1}^{10} X_i^2 \geq 15.99 \right\} = 1 - \chi^2(15.99/\sigma^2; 10),$$

$$\beta_{\phi_1}(\sigma^2) = P \left\{ X : \sum_{i=1}^{10} (X_i - \bar{X})^2 \geq 14.68 \right\} = 1 - \chi^2(14.68/\sigma^2; 9),$$

其中 $\chi^2(x; n)$ 表示自由度为 n 的 χ^2 分布随机变量的CDF. 对于不同的 σ^2 ，我们计算如上的势函数，其结果见图5.1.

单边假设的UMP检验

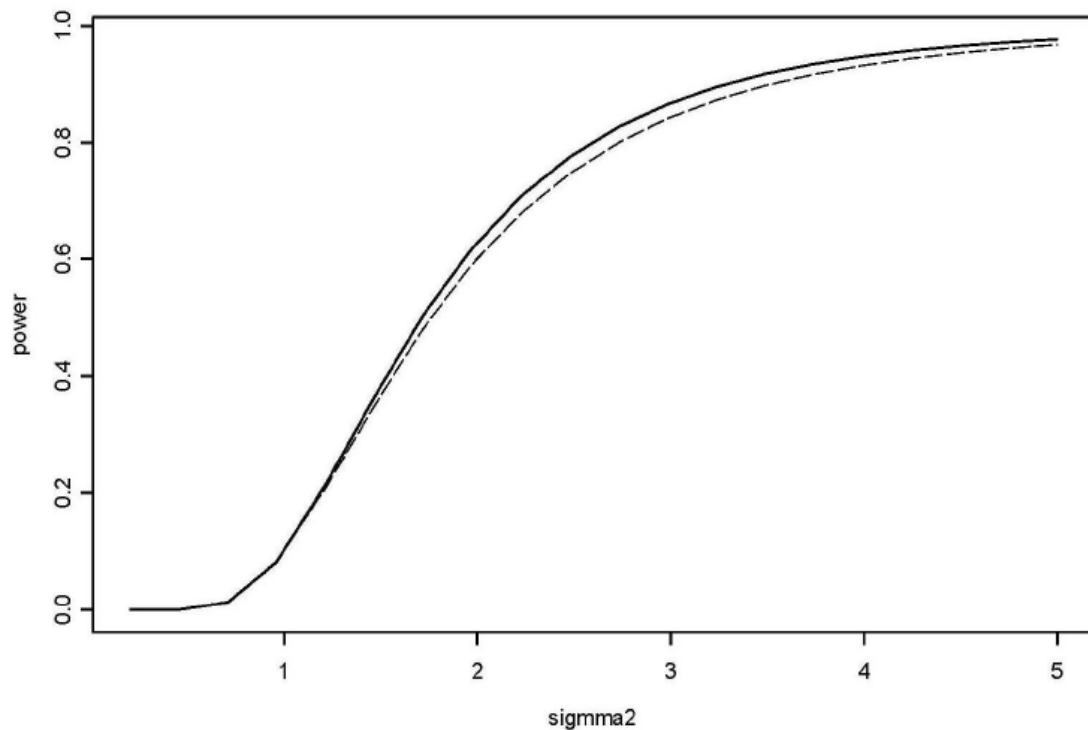


图5.1 例5.3.4中两个检验的势函数的比较

在图5.1中，实线表示UMP检验 ϕ 的势函数，而虚线则表示检验 ϕ_1 的势函数。从图中可以看出，此时的UMP检验 ϕ 的势明显优于检验 ϕ_1 。 □

双边假设的UMP检验

由于我们在一些实际问题中，感兴趣的双边假设可能形如

$$H_0 : \theta \leq \theta_1 \text{ 或 } \theta \geq \theta_2 \longleftrightarrow H_1 : \theta_1 < \theta < \theta_2, \quad (5.3.15)$$

故我们将在本小节考虑(5.3.15)的双边假设的UMP检验问题.

定理 5.3.6 设 $X_1, \dots, X_n$ 是来自单参数指数型分布

$$f(x; \theta) = c(\theta) \exp\{Q(\theta)T(x)\}h(x)$$

的IID样本，其中 θ 为一维实参数， $c(\theta) > 0$ ， $Q(\theta)$ 是 θ 的严格单增函数. 则关于双边假设(5.3.15)，存在水平为 $\alpha \in (0, 1)$ 的UMPT，其形式为

$$\phi(T(x)) = \begin{cases} 1, & c_1 < T(x) < c_2, \\ \delta_i, & T(x) = c_i, i = 1, 2, \\ 0, & T(x) < c_1 \text{ 或 } T(x) > c_2, \end{cases} \quad (5.3.17)$$

其中常数 $c_i, \delta_i, i = 1, 2$ 由下式确定

$$E_{\theta_1} \phi(T(X)) = E_{\theta_2} \phi(T(X)) = \alpha. \quad (5.3.18)$$

双边假设的UMP检验

例 5.3.8 对于正态总体 $N(\mu, 1)$, 考虑检验问题 $H_0 : \mu \leq \mu_1 \text{ 或 } \mu \geq \mu_2 \longleftrightarrow H_1 : \mu_1 < \mu < \mu_2$ 的UMPT.

双边假设的UMP检验

解 其样本联合PDF为

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

它是关于统计量 $T(x) = \bar{x}$ 单增的似然比分布族, 且 $Q(\mu) = n\mu$, 于是其UMPT为

$$\phi(T) = \begin{cases} 1, & c_1 \leq T \leq c_2, \\ 0, & \text{其它,} \end{cases}$$

其中 c_i 满足 $E_{\mu_1} \phi(T) = E_{\mu_2} \phi(T) = \alpha$.

由于当 $\mu = \mu_i$ 时, $\bar{X} \sim N(\mu_i, 1/n)$, 故 c_i 满足

$$\Phi(\sqrt{n}(c_2 - \mu_1)) - \Phi(\sqrt{n}(c_1 - \mu_1)) = \alpha,$$

$$\Phi(\sqrt{n}(c_2 - \mu_2)) - \Phi(\sqrt{n}(c_1 - \mu_2)) = \alpha.$$

目 录

4.1 Fisher的显著性检验

4.2 正态总体参数的假设检验

4.3 似然比检验

4.4 Neyman-Pearson基本引理

4.5 一致最优势检验

4.6 无偏检验与一致最优势无偏检验

无偏检验

假设我们感兴趣的假设为

$$H_0 : \theta \in \Theta_0 \longleftrightarrow H_1 : \theta \in \Theta_1.$$

定义 5.4.1 (无偏检验) 设 $\phi(x)$ 是上述假设的一个检验, 如果其势函数 $E_\theta\phi(X)$ 满足条件

$$E_\theta\phi(X) \leq \alpha, \forall \theta \in \Theta_0, \quad E_\theta\phi(X) \geq \alpha, \forall \theta \in \Theta_1,$$

则称之为水平 α 的无偏检验.

一致最优势无偏检验

定义 5.4.3 对于上述假设, 如果存在一个水平为 α 的无偏检验 $\phi(x)$, 使得对于任何水平为 α 的无偏检验 $\phi'(x)$, 均满足

$$E_{\theta}\phi(X) \geq E_{\theta}\phi'(X), \forall \theta \in \Theta_1,$$

则称检验 $\phi(x)$ 是水平 α 的一致最优势无偏检验, 记为UMPUT.

一致最优势无偏检验

定理 5.4.1 设 $X_1, \dots, X_n$ 是来自单参数指数型分布

$$f(x; \theta) = c(\theta) \exp\{Q(\theta)T(x)\}h(x)$$

的IID样本, 其中 θ 为一维实参数, $c(\theta) > 0$, $Q(\theta)$ 是 θ 的严格单增函数. 则双边假设(5.4.2)

$$H_0 : \theta_1 \leq \theta \leq \theta_2 \longleftrightarrow H_1 : \theta < \theta_1 \text{ 或 } \theta > \theta_2,$$

存在水平为 $\alpha \in (0, 1)$ 的UMPUT, 其检验函数为

$$\phi(T(x)) = \begin{cases} 1, & T(x) < c_1 \text{ 或 } T(x) > c_2, \\ \delta_i, & T(x) = c_i, \ i = 1, 2, \\ 0, & c_1 < T(x) < c_2, \end{cases} \quad (5.4.3)$$

其中常数 $\delta_i, c_i, i = 1, 2$ 由下式决定

$$E_{\theta_1} \phi(T(X)) = E_{\theta_2} \phi(T(X)) = \alpha. \quad (5.4.4)$$

一致最优势无偏检验

我们注意到此定理的结论, 即关于双边假设(5.4.2) 的UMPUT 与双边假设(5.3.15) 的UMPT 非常相似, 只是检验函数中的不等号变号而已.

定理 5.4.2 设 $X_1, \dots, X_n$ 是来自单参数指数型分布

$$f(x; \theta) = c(\theta) \exp\{Q(\theta)T(x)\}h(x)$$

的IID样本, 其中 θ 为一维实参数, $c(\theta) > 0$, $Q(\theta)$ 是 θ 的严格单增函数. 则双边假设(5.4.1)

$$H_0 : \theta = \theta_0 \longleftrightarrow H_1 : \theta \neq \theta_0,$$

存在水平为 $\alpha \in (0, 1)$ 的UMPUT, 其检验函数由(5.4.3)给出, 其中常数 δ_i, c_i , $i = 1, 2$ 由下式决定

$$E_{\theta_0} \phi(T(X)) = \alpha, \quad E_{\theta_0} [T(X) \phi(T(X))] = \alpha E_{\theta_0} T(X). \quad (5.4.5)$$

一致最优势无偏检验

例 5.4.1 设 $X_1, \dots, X_n$ 是来自正态总体 $N(\mu, 1)$ 的 IID 样本, 考虑假设

(1) $H_0 : \mu = 0 \longleftrightarrow H_1 : \mu \neq 0;$

(2) $H_0 : -1 \leq \mu \leq 1 \longleftrightarrow H_1 : \mu < -1 \text{ 或 } \mu > 1;$

的最优势检验.

一致最优势无偏检验

解 容易验证此分布族是满足条件的指数族分布, 且 $T(X) = \bar{X} \sim N(\mu, 1/n)$ 是连续的.

(1) 当 H_0 成立时, $T(X) \sim N(0, 1/n)$, 于是由上面的注5.4.1知, 其水平为 α 的UMPUT形如

$$\phi(T(x)) = \begin{cases} 1, & |\bar{x}| \geq \delta, \\ 0, & |\bar{x}| < \delta, \end{cases}$$

其中 δ 由下式确定

$$P_{H_0}\{\bar{X} < -\delta\} = \alpha/2,$$

故 $\delta = u_{\alpha/2}/\sqrt{n}$.

一致最优势无偏检验

(2) 关于此双边假设, 由定理5.4.1知, 其水平为 α 的UMPUT为

$$\phi(T(x)) = \begin{cases} 1, & \bar{x} < c_1 \text{ 或 } \bar{x} > c_2, \\ 0, & c_1 \leq \bar{x} \leq c_2, \end{cases}$$

其中 c_i 满足

$$E_{\mu=-1}\phi(\bar{X}) = E_{\mu=1}\phi(\bar{X}) = \alpha,$$

即

$$P_{\mu_1}\{c_1 \leq \bar{X} \leq c_2\} = 1 - \alpha = P_{\mu_2}\{c_1 \leq \bar{X} \leq c_2\},$$

也即

$$P_{\mu_1}\{\sqrt{n}(c_1 + 1) \leq \sqrt{n}(\bar{X} + 1) \leq \sqrt{n}(c_2 + 1)\} = 1 - \alpha,$$

$$P_{\mu_2}\{\sqrt{n}(c_1 - 1) \leq \sqrt{n}(\bar{X} - 1) \leq \sqrt{n}(c_2 - 1)\} = 1 - \alpha.$$

即

$$\int_{\sqrt{n}(c_1-1)}^{\sqrt{n}(c_2-1)} \phi(x)dx = 1 - \alpha = \int_{\sqrt{n}(c_1+1)}^{\sqrt{n}(c_2+1)} \phi(x)dx,$$

其中 $\phi(x)$ 是标准正态分的PDF.

假设检验与区间估计的对偶关系

p值

我们在这里

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他回答问题

联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

