

第三版前言

本书第二版出版至今已有 16 年了,感谢各位同仁一直以来给予我们的帮助与鼓励,感谢高等教育出版社给我们又一次再版的机会,修正已发现的疏漏和不当之处。趁再版之际,第三章增加了置换检验内容(§3.9)。置换检验是国际著名统计学家 Fisher 于 20 世纪 30 年代提出的一个非参数检验方法。置换检验的计算量大,因而限制了它的应用与推广。但如今计算机技术飞速发展,置换检验将会得到越来越广泛的应用。感谢岳荣先教授的仔细审阅,并推荐本书第三版早日出版。

当然此书还会有不当之处,恳请读者指正。我们十分感谢,并将及时改正,旨在提高本书质量。

茆诗松 王静龙 濮晓龙
2021 年 9 月于华东师范大学

第二版前言

本书出版已有六年多,不少教师选用此书作为概率论与数理统计专业的研究生教材。他们在指出不足之处的同时,还给了我们不少鼓励。特别地,上海理工大学叶慈南教授把教学中发现的本书不足之处一一向我们指正。在此向各位同仁一并表示我们的谢意。此外,我们还要特别感谢高等教育出版社给我们一个再版的机会。

趁此次再版之际,我们修正了原书中所有已发现的不足之处。在保持原书结构的同时,适当增加一些内容,供教师选用。所增内容有五处,它们是:矩的近似(§ 1.4.5)、稳健估计(§ 2.8)、秩检验(§ 3.8)、Bayes 假设检验(§ 5.5.6)和 Bayes 统计计算软件——Winbugs 简介(§ 6.4.6)。希望这些内容能对夯实研究生的统计基础知识起到积极的作用。

不当之处敬请读者指正。

茆诗松 王静龙 濮晓龙

2005 年 12 月于华东师范大学统计系

第一版前言

本书是为统计学专业及相关专业的学生和统计工作者编写的教科书。阅读此书需要有高等数学基础和概率论与数理统计基础知识。读完本书即可进入数理统计各分支的学习和研究。基于这样的要求,我们在本书中着力于数理统计的基本概念、基本方法和基本理论,充分反映数理统计的现代发展,力求做到理论与实际的结合,为读者进入理论研究领域和实际应用领域打下扎实的基础。

全书共分六章,依次为基本概念、点估计、假设检验、区间估计、统计决策理论与 Bayes 分析和统计计算方法。前五章的前身是一份讲义,曾在华东师范大学统计系研究生的高等数理统计课程上使用了十多年。虽经多次修改,总感不足。这次趁出版之际,对前五章作了较大的修改,充实了一些新的内容,另外在叙述上也作了不少改进,使内容有点新意,也更易理解。书中丰富的例子着力说明统计思想和统计应用领域,配置的习题足够让读者得到各种基本训练,掌握本书内容。完成这些习题就能品尝到统计学特有的味道。

本书的出版是在上海市学位委员会“上海研究生专项经费”资助下实现的,在他们的大力支持和倡导下,我们充满信心地完成这本书的充实、改写和完善工作。在此对上海市学位委员会表示衷心的感谢。另外对我校研究生院培养处徐钧涛副教授、高等教育出版社张小萍和翁咏梅二位女士表示衷心的感谢,没有他(她)们的帮助与关心,此书不可能很快出版。

本书的编写和修改得到我系广大师生的帮助,特别是梁小筠教授和程依明副教授。另外还有尤进红、刘忠和何基报三位博士生为本书部分章节的打印、修改做了很多工作,在此一并表示衷心感谢。

本书由茆诗松主编。第一、五章由茆诗松执笔,第三、四章由王静龙执笔,第二、六章由濮晓龙执笔,最后由茆诗松统稿。由于编者水平有限,错谬之处在所难免,恳请国内同行及广大读者批评指正。

茆诗松 王静龙 濮晓龙

1998 年 4 月于华东师范大学统计系

第三章 假设检验

§ 3.1 基本概念

常用的 U 检验, t 检验和 F 检验都具有足够的直观性和可信感, 本章将对这些检验的优良性作深入的研究. 那么, 一个检验“好”的标准是什么? 对于给定的“好坏”标准, 是否存在最优的检验方法? 如果存在, 又怎样求得? 这些问题的深入研究就要涉及假设检验理论的一些基本问题. 为此, 先引进一些基本概念.

§ 3.1.1 假设

定义 3.1 设 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ 为一统计结构, 则 $\mathcal{P}$ 的非空子集称为假设. 在参数分布族 $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ 时, Θ 的非空子集称为假设.

设样本空间 $\mathcal{X}$ 中的样本点 x 是我们的观测值, 除了参数估计, 统计推断中的又一类重要问题是, 检验观测值 x 与给定的假设是否存在矛盾. 这类问题称为假设检验问题.

在一个假设检验问题中常涉及两个假设. 所要检验的假设称为原假设, 记为 H_0 . 与 H_0 不相容的假设称为备择假设, 记为 H_1 . 关于统计结构 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ 的原假设和备择假设分别记为

$$H_0: P \in \mathcal{P}_0 \quad \text{对} \quad H_1: P \in \mathcal{P}_1.$$

这里 $\mathcal{P}_0$ 和 $\mathcal{P}_1$ 是 $\mathcal{P}$ 的两个互不相交的非空子集. 在参数分布族 $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ 场合, 原假设和备择假设分别记为

$$H_0: \theta \in \Theta_0 \quad \text{对} \quad H_1: \theta \in \Theta_1.$$

这里 Θ_0 和 Θ_1 是 Θ 的两个互不相交的非空子集. 给定 H_0 和 H_1 就等于给定一个检验问题, 记为检验问题 (H_0, H_1) . 而 (Θ_0, Θ_1) 称为参数假设检验问题, 其他称为非参数假设检验问题.

§ 3.1.2 检验, 拒绝域与检验统计量

定义 3.2 在检验问题 (H_0, H_1) 中, 所谓检验法则 (简称检验法, 或检验), 就是设法把样本空间划分为互不相交的两个可测集:

$$\mathcal{R} = W + \overline{W},$$

并作如下规定:

当观测值 $x \in W$ 时, 就拒绝原假设 H_0 , 认为备择假设 H_1 成立;

当观测值 $x \notin W$ (即 $x \in \overline{W}$) 时, 就不拒绝原假设 H_0 .

这里的 W 称为检验的拒绝域.

这样一来, 选定了检验法, 就是确定了拒绝域; 反之, 选定了拒绝域, 也就确定了检验法.

为了确定拒绝域, 往往首先由问题的直观背景出发, 寻找一个统计量, 使得在原假设 H_0 成立时和在备择假设 H_1 成立时, 该统计量的值有差异. 从而使得我们能够根据这个统计量的值的大小选定拒绝域. 我们称这个能从样本空间中划分出拒绝域的统计量为检验统计量.

例 3.1 电话交换台单位时间内接到的呼唤次数服从 Poisson 分布 $P(\lambda)$, $\lambda > 0$. λ 为单位时间内接到的平均呼唤次数. 为了考察该交换台在单位时间内的平均呼唤次数是否不超过 1, 可考虑建立如下两个假设:

$$H_0: \lambda \leq 1 \quad \text{对} \quad H_1: \lambda > 1. \quad (3.1)$$

设 $x = (x_1, x_2, \dots, x_n)$ 是该电话交换台的 n 次记录, 即它是来自 Poisson 分布 $\mathcal{P} = P(\lambda)$ 的样本 $X = (X_1, X_2, \dots, X_n)$ 的一个观测值. 取检验统计量 $T = \sum_{i=1}^n X_i$. 它是 λ 的充分、完备的统计量. 一般来说, 在原假设 H_0 成立时检验统计量的值较小, 而在备择假设 H_1 成立时检验统计量的值较大. 这里有一个临界值 c , 在 $T \geq c$ 时认为样本数据 x 与原假设相矛盾, 因而拒绝原假设; 而在 $T < c$ 时就不拒绝原假设. 该检验的拒绝域可用检验统计量表示如下:

$$W = \left\{ x: \sum_{i=1}^n x_i \geq c \right\}.$$

这个检验是直观和可信的. 如何确定 c 值的问题与检验犯两类错误的概率有关.

§ 3.1.3 两类错误

在进行检验时, 由于样本的随机性, 我们可能作出正确的判断, 也可能作出错误的判断, 正确的判断是原假设 H_0 成立时接受 H_0 , 或原假设 H_0 不成立时拒

绝 H_0 . 而错误的判断是原假设 H_0 成立时但被拒绝, 或原假设 H_0 不成立时但被接受. 为了对检验法的好坏给出一个合理的评选标准, 需要考察一个检验法可能犯错误的概率.

当原假设 H_0 成立时, 样本观测值却落在拒绝域 W 中, 从而拒绝了原假设. 这种错误称为第一类错误. 犯第一类错误的概率为

$$\alpha = P(X \in W), \quad P \in \mathcal{P}_0.$$

在参数结构, 犯第一类错误的概率为

$$\alpha(\theta) = P_\theta(X \in W), \quad \theta \in \Theta_0.$$

另一类错误是, 当原假设 H_0 不成立时, 样本观测值却没有落在拒绝域 W 中, 从而没有拒绝原假设. 这种错误称为第二类错误. 犯第二类错误的概率为

$$\beta = P(X \notin W) = 1 - P(X \in W), \quad P \in \mathcal{P}_1.$$

在参数结构, 犯第二类错误的概率为

$$\beta(\theta) = P_\theta(X \notin W) = 1 - P_\theta(X \in W), \quad \theta \in \Theta_1.$$

譬如例 3.1 中的检验犯第一类错误的概率为

$$\alpha(\lambda) = \sum_{k=c}^{\infty} \frac{(n\lambda)^k}{k!} \exp\{-n\lambda\}, \quad \lambda \leq 1. \quad (3.2)$$

而检验犯第二类错误的概率为

$$\begin{aligned} \beta(\lambda) &= \sum_{k=0}^{c-1} \frac{(n\lambda)^k}{k!} \exp\{-n\lambda\} \\ &= 1 - \sum_{k=c}^{\infty} \frac{(n\lambda)^k}{k!} \exp\{-n\lambda\}, \quad \lambda > 1. \end{aligned} \quad (3.3)$$

这两类错误的概率都可以通过 $P_\theta(X \in W)$ 来表示.

§ 3.1.4 势函数

定义 3.3 称样本观测值落在拒绝域的概率为检验的势函数, 记为

$$g(\theta) = P_\theta(X \in W), \quad \theta \in \Theta.$$

在 $\theta \in \Theta_0$ 时, $g(\theta) = \alpha(\theta)$, $g(\theta)$ 是检验犯第一类错误的概率. 在 $\theta \in \Theta_1$ 时, $g(\theta) = 1 - \beta(\theta)$, $1 - g(\theta)$ 是检验犯第二类错误的概率.

在例 3.1 中, 检验的势函数为

$$g(\lambda) = \sum_{k=c}^{\infty} \frac{(n\lambda)^k}{k!} \exp\{-n\lambda\}, \quad (3.4)$$

它是 λ 的严增函数. 在 $n=10$ 时, 临界值分别取为 $c=16$ 和 $c=14$ 的两个检验的势函数的具体曲线形式如图 3.1 所示.

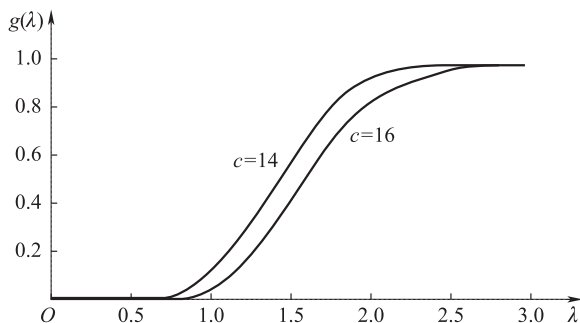


图 3.1 Poisson 分布均值检验的势函数的图像

§ 3.1.5 检验的水平

在例 3.1 中, 从式 (3.2) 和式 (3.3) 中可以看出, 在样本容量 n 固定时, 要减少犯第一类错误的概率, 必须增加 c , 从而导致犯第二类错误的概率增加; 反之, 若要减少犯第二类错误的概率, 必须减少 c , 从而使犯第一类错误的概率增加. 换句话说, 当样本容量 n 固定时, 不可能使得犯这两类错误的概率都减少. 这一现象在一般的检验问题中都会出现.

基于这种情况, 需要采取某种妥协方案. Neyman (内曼) 和 Pearson 的假设检验理论的基本思想, 就是使犯第一类错误的概率限制在某一个范围内, 然后寻找使犯第二类错误的概率尽可能小的检验. 在这种思想的指导下, 寻找一个好的检验法, 就是对选定的一个较小的数 $\alpha (0 < \alpha < 1)$, 在满足

$$g(\theta) = P_{\theta}(X \in W) \leq \alpha, \quad \theta \in \Theta_0$$

的检验中, 寻找这样的检验, 使得在 $\theta \in \Theta_1$ 时, $g(\theta)$ 尽可能大.

在 $\theta \in \Theta_0$ 时, $g(\theta) \leq \alpha$ 的检验称为显著性水平为 α 的检验, 记为 $(\alpha, \Theta_0, \Theta_1)$ 检验. α 是事先选定的, 由于 α 的大小反映了检验犯第一类错误的概率的大小, 所以常取一个较小的数. 例如取 0.1, 0.05, 0.01 等作为 α 的值. α 的选取也依赖于我们关于假设的先验信息, 依赖于我们对犯第二类错误的要求. 比如, 根据以往的经验, 非常相信原假设是真的, 而犯第二类错误又不会造成大的影响或后果, 此时 α 就可以取的小一些. 又比如, 犯第二类错误带来的影响较大, 需要严格控制犯第二类错误的概率, 此时 α 可以选的适当大一些.

在例 3.1 中, 由于检验的势函数 $g(\lambda)$ (见式 (3.4)) 是 λ 的严增函数, 给定 α 后, 为使犯第一类错误的概率不超过 α , 即在 $\lambda \leq 1$ 时, $g(\lambda) \leq \alpha$, 只需

$$g(1) = \sum_{k=c}^{\infty} \frac{n^k}{k!} \exp\{-n\} \leq \alpha, \quad (3.5)$$

$n=10$, c 取不同的值时, 式 (3.5) 左边的值如表 3.1 所示.

表 3.1 Poisson 分布均值检验犯第一类错误的概率

c	$g(1)$	c	$g(1)$
14	0.135 536	17	0.027 042
15	0.083 458	18	0.014 278
16	0.048 740	19	0.007 187

在 $n=10$ 时, 若取 $\alpha=0.05$, 由于

$$\sum_{k=16}^{\infty} \frac{10^k}{k!} \exp\{-10\} = 0.048 740 < 0.05, \quad (3.6)$$

$$\sum_{k=15}^{\infty} \frac{10^k}{k!} \exp\{-10\} = 0.083 458 > 0.05. \quad (3.7)$$

根据 Neyman 和 Pearson 的假设检验理论的基本思想, 取 $c=16$. 所以在 $X_1, X_2, \dots, X_{10}$ 是来自 Poisson 分布总体 $\mathcal{P} = \{P(\lambda), \lambda > 0\}$ 的一个样本时, 关于检验问题 (3.1) 的显著性水平 $\alpha=0.05$ 的拒绝域为

$$W = \left\{ x: \sum_{i=1}^{10} x_i \geq 16 \right\}.$$

在上述的叙述中, 对拒绝原假设, 给予特别的注意. 在原假设为真时, 样本观测值落入拒绝域是一个小概率事件. 根据实际推断原理, 在一次观测中, 小概率事件是几乎不可能发生的. 所以拒绝原假设的理由是充足的. 由于检验犯第二类错误的概率并没有受到限制, 仅要求它尽可能小, 所以在备择假设为真时, 样本观测值没有落入拒绝域可能不是一个小概率事件. 因此接受原假设的理由可能是不充足的. 较为科学的说法是不拒绝原假设. 把“不拒绝原假设”说成“接受原假设”, 这是为了防止在诸如“除了拒绝与接受之外, 还有第三种可能”之类的哲理性的讨论, 从而使实际工作者能较快地熟悉和使用假设检验方法.

对给定的显著性水平 α , 人们总是力图构造这样的检验, 使得检验犯第一类错误的概率

$$g(\theta) = P_{\theta}(X \in W) \leq \alpha, \quad \theta \in \Theta_0,$$

并且至少存在一个 $\theta \in \Theta_0$, 使得

$$g(\theta) = P_\theta(X \in W) = \alpha.$$

这意味着显著性水平被“足量”地使用了, 以满足尽可能减少检验犯第二类错误的概率的要求. 在例 3.1 中, $n=10$ 时, 所求得的显著性水平 $\alpha=0.05$ 的检验, 显著性水平并没有被“足量”地使用. 在原假设 $H_0: \lambda \leq 1$ 成立时,

$$\begin{aligned} \alpha(\lambda) &= \sum_{k=16}^{\infty} \frac{(10\lambda)^k}{k!} \exp\{-10\lambda\} \\ &\leq \sum_{k=16}^{\infty} \frac{10^k}{k!} \exp\{-10\} = 0.048\ 740. \end{aligned}$$

为了能构造出显著性水平 $\alpha=0.05$ 被“足量”地使用的检验, 需要引进随机化检验.

§ 3.1.6 检验函数和随机化检验

定义函数

$$\phi(x) = \begin{cases} 1, & x \in W, \\ 0, & x \notin W, \end{cases}$$

它是拒绝域 W 的示性函数, 仅取 0 和 1 两个值. 反之, 如果一个函数 $\phi(x)$ 仅取 0 和 1 两个值, 则 $W = \{x: \phi(x) = 1\}$ 可作为拒绝域. 于是, 给出了这样的函数, 就等于给了一个检验法. 这样的函数 $\phi(x)$ 称为检验函数, 简称检验. 它是仅取 0 和 1 两个值的统计量. 检验的势函数可表示为

$$g(\theta) = P_\theta(X \in W) = E_\theta(\phi(X)), \quad (3.8)$$

为了能构造出显著性水平被“足量”地使用的检验, 需要拓广检验函数的概念.

定义 3.4 设 $\phi(x)$ 为定义在 $\mathcal{X}$ 上的可测函数, 满足条件 $0 \leq \phi(x) \leq 1$, 则称 $\phi(x)$ 为检验函数, 简称检验. 在 $\phi(x)$ 仅取 0, 1 两个值时, 为非随机化检验. 否则, 为随机化检验. 其势函数为 $g(\theta) = E_\theta(\phi(X))$.

非随机化检验的拒绝域 $W = \{x: \phi(x) = 1\}$. 在随机化检验时, 假设样本观测值为 x , 则我们以概率 $\phi(x)$ 拒绝原假设; 以概率 $1-\phi(x)$ 接受原假设. 在 $\theta \in \Theta_0$ 时, $g(\theta)$ 是检验犯第一类错误的概率, 而在 $\theta \in \Theta_1$ 时, $1-g(\theta)$ 是检验犯第二类错误的概率.

在例 3.1 中, $n=10, \alpha=0.05$ 时, 考虑到检验的直观和可信, 取随机化检验函数

$$\phi(x) = \begin{cases} 1, & T \geq 16, \\ r, & T = 15, \\ 0, & T \leq 14, \end{cases}$$

其中 $T = \sum_{i=1}^{10} x_i$. 在原假设 $H_0: \lambda \leq 1$ 成立时, 该检验犯第一类错误的概率为

$$\begin{aligned} \alpha(\lambda) &= E_{\lambda}(\phi(X)) = \sum_{k=16}^{\infty} \frac{(10\lambda)^k}{k!} \exp\{-10\lambda\} + r \frac{(10\lambda)^{15}}{15!} \exp\{-10\lambda\} \\ &\leq \sum_{k=16}^{\infty} \frac{10^k}{k!} \exp\{-10\} + r \frac{10^{15}}{15!} \exp\{-10\} \\ &= 0.048\,740 + r \times 0.034\,718. \end{aligned}$$

为使得显著性水平 $\alpha=0.05$ 被“足量”地使用, 应有

$$\begin{aligned} 0.048\,740 + r \times 0.034\,718 &= 0.05, \\ r &= \frac{0.05 - 0.048\,740}{0.034\,718} = 0.036. \end{aligned}$$

这个随机化检验的实施过程如下: 若样本观测值 $x_1, x_2, \dots, x_{10}$ 满足条件 $\sum_{i=1}^{10} x_i \geq 16$, 则拒绝原假设; 若 $\sum_{i=1}^{10} x_i \leq 14$, 则接受原假设; 若 $\sum_{i=1}^{10} x_i = 15$, 先做一个成功概率为 0.036 的 Bernoulli 试验, 如果试验结果为成功, 则拒绝原假设, 反之则接受原假设.

在下面两节, 这个随机化检验的优良性将被说明.

§ 3.1.7 充分性原则

定义 3.5 设 $\phi_1(x)$ 和 $\phi_2(x)$ 都是某检验问题 (H_0, H_1) 的检验函数, 如果它们的势函数相同, 即

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

则称检验函数 $\phi_1(x)$ 和 $\phi_2(x)$ 等价.

这个定义意味着, 检验函数的统计性质完全取决于其势函数. 势函数相同, 就认为这两个检验函数等价.

定理 3.1 设 $X = (X_1, X_2, \dots, X_n)$ 是来自分布族 $\mathcal{P} = \{P_{\theta}: \theta \in \Theta\}$ 的样本, $T(X)$ 是关于 θ 的充分统计量. 对任意一个检验函数 $\phi(x)$, 存在另一个只依赖于 $T(x)$ 的检验函数, 它与 $\phi(x)$ 相互等价.

证明: 因为 $T(X)$ 是关于 θ 的充分统计量, 所以条件期望 $\varphi(t) = E[\phi(X)]$

$|T=t]$ 与 θ 无关, 又因为 $0 \leq \phi(x) \leq 1$, 所以 $0 \leq \varphi(t) \leq 1$. 因而 $\varphi(t)$ 也是检验函数. 由于

$$E_{\theta}(\varphi(T(X))) = E_{\theta}(E(\phi(X) | T)) = E_{\theta}(\phi(X)),$$

所以 $\varphi(t)$ 和 $\phi(x)$ 具有相同的势函数, 相互等价. 证毕.

这个定理告诉我们, 当 θ 的充分统计量存在时, 关于 θ 的任何假设检验问题, 其优良检验, 只要在充分统计量构成的检验函数中去寻找就可以了. 这就是假设检验中所谓的“充分性原则”.

§ 3.2 Neyman-Pearson 基本引理

如果一个假设只含有一个元素, 则称该假设为简单假设, 否则称为复合假设. 在例 3.1 中, 原假设 $H_0: \lambda \leq 1$ 和备择假设 $H_1: \lambda > 1$ 都是复合假设. 如果我们考虑的检验问题是, 原假设 $H_0: \lambda = 1$ 对备择假设 $H_1: \lambda > 1$, 那么这个原假设就是简单假设.

现在我们来考虑简单原假设对简单备择假设的检验问题. 这不仅是因为这样的检验问题特别简单, 已经圆满地解决, 而且也是我们解决许多更一般的检验问题的基础.

设 $(\mathcal{X}, \mathcal{B}, \{P_{\theta}: \theta \in \Theta\})$ 为一参数结构, 考虑检验问题:

$$\text{简单原假设 } H_0: \theta = \theta_0 \text{ 对简单备择假设 } H_1: \theta = \theta_1 (\theta_0 \neq \theta_1). \quad (3.9)$$

这时, 比较任意两个显著性水平为 α 的检验 $\phi_1(x)$ 和 $\phi_2(x)$ 的好坏是容易的. 在 $E_{\theta_1}(\phi_1(X)) \geq E_{\theta_1}(\phi_2(X))$ 时, 我们说 $\phi_1(x)$ 不比 $\phi_2(x)$ 差. 由之, 很自然地有下面的定义.

定义 3.6 在检验问题 (Θ_0, Θ_1) 中, 设 $\phi(x)$ 是显著性水平为 α 的检验. 如果对任意一个显著性水平为 α 的检验 $\phi_1(x)$, 都有

$$E_{\theta_1}(\phi(X)) \geq E_{\theta_1}(\phi_1(X)),$$

则称检验 $\phi(x)$ 是显著性水平为 α 的最优势检验, 记为 MPT (most powerful test).

下面的著名引理证明了, 在简单原假设对简单备择假设的检验问题中, MPT 一定存在, 并且可以具体构造出 MPT 的检验函数.

定理 3.2 (Neyman-Pearson 基本引理, 简称 N-P 基本引理) 设 P_{θ_0} 和 P_{θ_1} 是可测空间 $(\mathcal{X}, \mathcal{B})$ 上的两个不同的概率测度, 关于某个 σ 有限的测度 μ , 有

$$p(x; \theta_0) = \frac{dP_{\theta_0}}{d\mu}, \quad p(x; \theta_1) = \frac{dP_{\theta_1}}{d\mu},$$

则在检验问题 (3.9) 中,

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

$$E_{\theta_0}(\phi(X)) = \alpha, \quad (3.10)$$

$$\phi(x) = \begin{cases} 1, & p(x; \theta_1) > kp(x; \theta_0), \\ 0, & p(x; \theta_1) < kp(x; \theta_0); \end{cases} \quad (3.11)$$

(2) 由式(3.10)和式(3.11)确定的检验函数 $\phi(x)$ 是显著性水平为 α 的 MPT. 反之, 如果 $\phi(x)$ 是显著性水平为 α 的 MPT, 则一定存在常数 $k \geq 0$, 使得 $\phi(x)$ 满足式(3.11) (a. e. $[\mu]$).

证明: 1° 首先证明(1), 存在形如式(3.11)的检验函数, 且使式(3.10)成立.

对于任一实数 λ , 令

$$G(\lambda) = P_{\theta_0} \{ p(X; \theta_1) > \lambda p(X; \theta_0) \},$$

由于这个概率是在 P_{θ_0} 下计算的, 所以只要在集合 $\{x: p(x; \theta_0) > 0\}$ 上考虑不等式 $p(x; \theta_1) > \lambda p(x; \theta_0)$.

$G(\lambda)$ 是非负随机变量 $\frac{p(X; \theta_1)}{p(X; \theta_0)} > \lambda$ 的概率, 故 $1 - G(\lambda)$ 是随机变量 $\frac{p(X; \theta_1)}{p(X; \theta_0)}$ 的分布函数. 所以 $G(\lambda)$ 是一个非增的右连续函数, 且

$$G(+\infty) = 0, \quad G(0-0) = 1,$$

$$G(\lambda - 0) - G(\lambda) = P_{\theta_0} \left\{ \frac{p(X; \theta_1)}{p(X; \theta_0)} = \lambda \right\}.$$

给定 $\alpha \in (0, 1)$ 后, 有且只有下列两种情况:

I. 存在 $\lambda_0 \geq 0$, 使得 $G(\lambda_0) = \alpha$. 定义

$$\phi(x) = \begin{cases} 1, & p(x; \theta_1) > \lambda_0 p(x; \theta_0), \\ 0, & p(x; \theta_1) \leq \lambda_0 p(x; \theta_0), \end{cases}$$

则

$$E_{\theta_0}(\phi(X)) = P_{\theta_0} \{ p(X; \theta_1) > \lambda_0 p(X; \theta_0) \} = G(\lambda_0) = \alpha.$$

II. 存在 $\lambda_0 \geq 0$, 使得 $G(\lambda_0) < \alpha \leq G(\lambda_0 - 0)$. 定义

$$\phi(x) = \begin{cases} 1, & p(x; \theta_1) > \lambda_0 p(x; \theta_0), \\ \frac{\alpha - G(\lambda_0)}{G(\lambda_0 - 0) - G(\lambda_0)}, & p(x; \theta_1) = \lambda_0 p(x; \theta_0), \\ 0, & p(x; \theta_1) < \lambda_0 p(x; \theta_0), \end{cases}$$

则

$$\begin{aligned} E_{\theta_0}(\phi(X)) &= P_{\theta_0}\{p(X; \theta_1) > \lambda_0 p(X; \theta_0)\} + \\ &\quad \frac{\alpha - G(\lambda_0)}{G(\lambda_0 - 0) - G(\lambda_0)} P_{\theta_0}\{p(X; \theta_1) = \lambda_0 p(X; \theta_0)\} \\ &= G(\lambda_0) + [\alpha - G(\lambda_0)] = \alpha. \end{aligned}$$

在这两种情况下, λ_0 就可以被取作式 (3.11) 中的非负常数 k . $\phi(x)$ 使得式 (3.10) 成立, 这说明 $\phi(x)$ 是显著性水平为 α 的检验函数. 它是显著性水平 α 被“足量”地使用的检验函数.

2° 其次证明, 由式 (3.10) 和式 (3.11) 得到的检验函数 $\phi(x)$ 是 MPT.

设 $\phi^*(x)$ 是其他任意一个显著性水平为 α 的检验函数, 即 $E_{\theta_0}(\phi^*(X)) \leq \alpha$. 由于 $\phi(x)$ 满足式 (3.11), 所以

$$(\phi(x) - \phi^*(x))(p(x; \theta_1) - kp(x; \theta_0)) \geq 0,$$

因此

$$\int (\phi(x) - \phi^*(x))(p(x; \theta_1) - kp(x; \theta_0)) d\mu(x) \geq 0,$$

从而

$$\begin{aligned} E_{\theta_1}(\phi(X)) - E_{\theta_1}(\phi^*(X)) &\geq k[E_{\theta_0}(\phi(X)) - E_{\theta_0}(\phi^*(X))] \\ &= k[\alpha - E_{\theta_0}(\phi^*(X))] \geq 0, \end{aligned}$$

所以 $\phi(x)$ 是 MPT.

3° 最后证明: 若 $\phi^*(x)$ 是显著性水平为 α 的 MPT, 则一定存在非负常数 k , 使得 $\phi^*(x)$ 满足式 (3.11) (a. e. $[\mu]$).

设 $\phi(x)$ 是满足式 (3.10) 和式 (3.11) 的检验函数. 由 2° 知, $\phi(x)$ 是显著性水平为 α 的 MPT. 由于 $\phi^*(x)$ 和 $\phi(x)$ 都是显著性水平为 α 的 MPT, 则 $E_{\theta_1}(\phi^*(X)) = E_{\theta_1}(\phi(X))$. 由于 $\phi(x)$ 满足式 (3.11), 所以

$$(\phi(x) - \phi^*(x))(p(x; \theta_1) - kp(x; \theta_0)) \geq 0, \quad (3.12)$$

因为 $E_{\theta_1}(\phi^*(X)) = E_{\theta_1}(\phi(X))$, $E_{\theta_0}(\phi^*(X)) \leq \alpha = E_{\theta_0}(\phi(X))$, 所以

$$\begin{aligned} &\int [\phi(x) - \phi^*(x)][p(x; \theta_1) - kp(x; \theta_0)] d\mu(x) \\ &= E_{\theta_1}(\phi(X)) - E_{\theta_1}(\phi^*(X)) - k[E_{\theta_0}(\phi(X)) - E_{\theta_0}(\phi^*(X))] \leq 0. \end{aligned} \quad (3.13)$$

比较 (3.12) 和 (3.13) 两个不等式, 可得

$$(\phi(x) - \phi^*(x))(p(x; \theta_1) - kp(x; \theta_0)) = 0 \quad (\text{a. s. } [\mu]),$$

故在集合 $\{x: p(x; \theta_1) - kp(x; \theta_0) \neq 0\}$ 上, 有 $\phi^*(x) = \phi(x) \text{ (a. s. } [\mu])$. 这证明了 $\phi^*(x)$ 满足式 (3.11) (a. e. $[\mu]$). 证毕.

注 1. 满足式 (3.11) 的检验函数 $\phi(x)$ 通常称为似然比检验函数 (或称为概率比检验函数). 在集合 $\{x: p(x; \theta_0) > 0 \text{ 或 } p(x; \theta_1) > 0\}$ 上, 定义似然比检验函数

$$\lambda(x) = \frac{p(x; \theta_1)}{p(x; \theta_0)}.$$

在 $\lambda(x)$ 比较大的时候, $p(x; \theta_1)$ 比较大, 所以原假设 $H_0: \theta = \theta_0$ 为真时观察到样本点 x 的可能性比备择假设 $H_1: \theta = \theta_1$ 为真时观察到样本点 x 的可能性小. 而在 $\lambda(x)$ 比较小的时候, $p(x; \theta_0)$ 比较大, 所以备择假设 $H_1: \theta = \theta_1$ 为真时观察到样本点 x 的可能性比原假设 $H_0: \theta = \theta_0$ 为真时观察到样本点 x 的可能性小. 由此很自然地, 我们在 $\lambda(x)$ 比较大的时候, 拒绝原假设, 认为 $\theta = \theta_1$ 成立; 而在 $\lambda(x)$ 比较小的时候, 不拒绝原假设, 认为 $\theta = \theta_0$ 成立. N-P 基本引理告诉我们, 在简单原假设对简单备择假设的检验问题中, MPT 是似然比检验; 反之, 似然比检验也是 MPT.

注 2. 在似然比 $\lambda(x)$ 具有连续分布函数时, MPT 检验函数可取为非随机化的形式

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

其中 k 由式 (3.10), 即 $E_{\theta_0}(\phi(X)) = P_{\theta_0}\{\lambda(X) \geq k\} = \alpha$ 确定. 在 $\lambda(X)$ 为离散型随机变量时, 至多在集合 $\{x: \lambda(x) = k\}$ 上实施随机化. MPT 检验函数可取为

$$\phi(x) = \begin{cases} 1, & \lambda(x) > k, \\ r, & \lambda(x) = k, \\ 0, & \lambda(x) < k, \end{cases}$$

其中 k 与 r 由式 (3.10), 即由以下两式确定:

$$P_{\theta_0}\{\lambda(X) \geq k\} \geq \alpha > P_{\theta_0}\{\lambda(X) > k\} = \alpha_1,$$

$$r = \frac{\alpha - \alpha_1}{P_{\theta_0}\{\lambda(X) = k\}}.$$

有时, 在集合 $\{x: \lambda(x) = k\}$ 上可以实施非随机化 (见下面的例 3.3).

注 3. 在非参数统计结构 $(\mathcal{X}, \mathcal{B}, \mathcal{P})$ 中, 考虑简单原假设 $H_0: P = P_0$ 对简单备择假设 $H_1: P = P_1$ 的检验问题. 设 P_0 和 P_1 是可测空间 $(\mathcal{X}, \mathcal{B})$ 上的两个不同的概率测度, 且关于某个 σ 有限的测度 μ , 有

$$p_0(x) = \frac{dP_0}{d\mu}, \quad p_1(x) = \frac{dP_1}{d\mu},$$

定义似然比 $\lambda(x) = \frac{p_1(x)}{p_0(x)}$. N-P 基本引理仍然适用.

例 3.2 设 $X = (X_1, X_2, \dots, X_n)$ 是来自正态分布族 $\{N(\mu, 1) : -\infty < \mu < \infty\}$ 的样本, 考虑如下的检验问题: 原假设 $H_0: \mu = 0$ 对备择假设 $H_1: \mu = \mu_1 (\mu_1 > 0)$, 取显著性水平为 $\alpha (0 < \alpha < 1)$. 构造似然比统计量

$$\lambda(x) = \frac{\prod_{i=1}^n p(x_i; \mu_1)}{\prod_{i=1}^n p(x_i; 0)} = \exp\left\{n\mu_1 \bar{x} - \frac{1}{2}n\mu_1^2\right\},$$

MPT 的拒绝域具有形式

$$W = \{x; \lambda(x) \geq k\} = \{x; \bar{x} \geq c\}.$$

当 H_0 成立时, $\bar{X} \sim N\left(0, \frac{1}{n}\right)$. 所以对于给定的显著性水平 α , $c = \frac{U_{1-\alpha}}{\sqrt{n}}$. MPT 检验

仅与显著性水平 α 有关, 而与 μ_1 的具体数值无关, 只要求 $\mu_1 > 0$ 就行了.

如果把此例中的备择假设改为 $H_1: \mu = \mu_1 (\mu_1 < 0)$, 则 MPT 检验的拒绝域为

$$W = \left\{x; \bar{x} \leq \frac{U_\alpha}{\sqrt{n}}\right\}.$$

同样, 该检验仅与显著性水平 α 有关, 而与 μ_1 的具体数值无关, 只要求 $\mu_1 < 0$ 就行了.

若正态分布族为 $\{N(\mu, \sigma_0^2) : -\infty < \mu < \infty\}$, $H_0: \mu = \mu_0$, $H_1: \mu = \mu_1 (\mu_1 > \mu_0)$, 则可以通过变换 $y_i = \frac{x_i - \mu_0}{\sigma_0}$, 将这个假设检验问题归结为上面所述的情况, 其拒绝域为

$$W = \left\{\frac{\bar{x} - \mu_0}{\sigma_0} \geq \frac{U_{1-\alpha}}{\sqrt{n}}\right\}.$$

例 3.3 设 $X = (X_1, X_2, \dots, X_n)$ 是来自均匀分布族 $\{R(0, \theta) : \theta > 0\}$ 的样本, 考虑如下的检验问题:

$$H_0: \theta = 1 \quad \text{对} \quad H_1: \theta = \theta_1 (\theta_1 > 1),$$

取显著性水平为 $\alpha (0 < \alpha < 1)$. 构造似然比统计量

$$\lambda(x) = \frac{\prod_{i=1}^n p(x_i; \theta_1)}{\prod_{i=1}^n p(x_i; 1)} = \begin{cases} \theta_1^{-n}, & 0 < x_{(n)} < 1, \\ \infty, & 1 \leq x_{(n)} < \theta_1, \end{cases}$$

其中 $x_{(n)} = \max\{x_1, x_2, \dots, x_n\}$. 在原假设 H_0 成立时, $\lambda(x)$ 的分布是退化分布. 令 $G(\lambda) = P(\lambda(X) > \lambda \mid \theta = 1)$, 则

$$0 = G(\theta_1^{-n}) < \alpha < G(\theta_1^{-n} - 0) = 1.$$

由 N-P 基本引理知, 取 $k = \theta_1^{-n}$, 则 MPT 为

$$\phi(x) = 1, \quad 1 \leq x_{(n)} < \theta_1,$$

而在 $0 < x_{(n)} < 1$ 时, $\phi(x)$ 的值由条件 $E(\phi(X) \mid \theta = 1) = \alpha$ 而定. 例如, 取随机化检验

$$\phi(x) = \alpha, \quad 0 < x_{(n)} < 1,$$

也可取非随机化检验

$$\phi(x) = \begin{cases} 0, & 0 < x_{(n)} < c, \\ 1, & c \leq x_{(n)} < \theta_1. \end{cases}$$

由于在原假设 $H_0: \theta = 1$ 成立时, $T = X_{(n)}$ 的密度函数为 $p(t) = nt^{n-1} (0 < t < 1)$, 故由 $E(\phi(X) \mid \theta = 1) = \alpha$ 推得 $c = \sqrt[n]{1-\alpha}$. 显然, 后一个非随机化检验

$$\phi(x) = \begin{cases} 0, & 0 < x_{(n)} < \sqrt[n]{1-\alpha}, \\ 1, & \sqrt[n]{1-\alpha} \leq x_{(n)} < \theta_1 \end{cases} \quad (3.14)$$

较为合理.

如果把此例中的备择假设改为 $H_1: \theta = \theta_1 (\theta_1 < 1)$, 其 MPT 如何? 习题 3.3 将讨论这个问题.

如果原假设和备择假设分别为 $H_0: \theta = \theta_0$ 和 $H_1: \theta = \theta_1 (\theta_1 > \theta_0)$, 则可以通过变换 $y_i = \frac{x_i}{\theta_0}$, 将这个假设检验问题归结为上面所述的情况, 其非随机化检验为

$$\phi(x) = \begin{cases} 0, & 0 < y_{(n)} < \sqrt[n]{1-\alpha}, \\ 1, & \sqrt[n]{1-\alpha} \leq y_{(n)} < \theta_1/\theta_0, \end{cases}$$

其中 $y_{(n)} = \max(y_1, y_2, \dots, y_n) = \frac{x_{(n)}}{\theta_0}$.

例 3.4 设 $X = (X_1, X_2, \dots, X_n)$ 是来自 Poisson 分布族 $\{P(\lambda): \lambda > 0\}$ 的样本,

考虑如下的检验问题:

$$H_0: \lambda = 1 \quad \text{对} \quad H_1: \lambda = \lambda_1 (\lambda_1 > 1),$$

取显著性水平为 $\alpha (0 < \alpha < 1)$. 构造似然比统计量

$$\lambda(x) = \frac{\prod_{i=1}^n p(x_i; \lambda_1)}{\prod_{i=1}^n p(x_i; 1)} = (\lambda_1)^{\sum_{i=1}^n x_i} \exp \{ -n(\lambda_1 - 1) \},$$

$\lambda(x)$ 关于 $T = \sum_{i=1}^n x_i$ 严格单调上升. 根据 N-P 基本引理, 显著性水平 $\alpha = 0.05$ 的 MPT 检验函数满足条件

$$\phi(x) = \begin{cases} 1, & T > k, \\ 0, & T < k. \end{cases}$$

在 $n=10$ 时, 若原假设 $H_0: \lambda = 1$ 成立, $T \sim P(10)$. 由式 (3.6) 和式 (3.7) 知, 取 $k=15$, 则检验函数为

$$\phi(x) = \begin{cases} 1, & T > 15, \\ r, & T = 15, \\ 0, & T < 15. \end{cases}$$

同例 3.1, 由显著性水平 $\alpha = 0.05$ 算得 $r = 0.036$. 这个检验与 λ_1 的具体数值无关, 只要求 $\lambda_1 > 1$ 就行了.

上述三例的 MPT 检验函数都仅依赖于充分统计量, 而这正如定理 3.1 的充分性原则告诉我们的. 事实上, 可以基于充分统计量, 而不是样本构造似然比统计量. 这样求得的检验函数仍如上面所述.

关于 MPT 还有一个性质, 它在下一节的定理的证明中将要用到.

定理 3.3 设 $\phi(x)$ 是定理 3.2 中显著性水平为 α 的最优势检验函数, 则 $E_{\theta_1}(\phi(X)) \geq \alpha$. 又设 $0 < \alpha < 1$, 则除了 $p(x; \theta_1) = p(x; \theta_0)$, a. e. $[\mu]$ 外, 必有 $E_{\theta_1}(\phi(X)) > \alpha$.

证明: 令 $\phi_1(x) \equiv \alpha$. 显然, $\phi_1(x)$ 是显著性水平为 α 的检验. 因为 $\phi(x)$ 是显著性水平为 α 的最优势检验函数, 故 $E_{\theta_1}(\phi(X)) \geq E_{\theta_1}(\phi_1(X)) = \alpha$. 在 $0 < \alpha < 1$ 时, 倘若 $E_{\theta_1}(\phi(X)) = \alpha$, 则 $\phi_1(x) \equiv \alpha$ 也是显著性水平为 α 的 MPT. 由定理 3.2 知, 一定存在常数 $k \geq 0$, 使得 $\phi_1(x)$ 满足式 (3.11) (a. e. $[\mu]$). 因 $0 < \alpha < 1$, 故在集合 $\{x: p(x; \theta_1) > 0 \text{ 或 } p(x; \theta_0) > 0\}$ 上,

$$p(x; \theta_1) = kp(x; \theta_0), \quad \text{a. e. } [\mu].$$

由于 p 是密度函数, 因此 $k=1$, 即 $p(x; \theta_1) = p(x; \theta_0)$, a. e. $[\mu]$. 证毕.

§ 3.3 一致最优势检验

§ 3.3.1 一致最优势检验

关于简单原假设对简单备择假设的检验问题, N-P 基本引理给出了令人满意的回答. 但是, 在实际问题中, 往往出现的是复合假设的情况, 这时最优的标准是什么?

定义 3.7 设 $(\mathcal{X}, \mathcal{B}, \mathcal{P} = \{P_\theta: \theta \in \Theta\})$ 为一参数结构, 考虑检验问题 (Θ_0, Θ_1) . 设 $\phi(x)$ 是显著性水平为 α 的检验. 如果对任意一个显著性水平为 α 的检验 $\phi_1(x)$, 都有

$$E_\theta(\phi(X)) \geq E_\theta(\phi_1(X)), \quad \forall \theta \in \Theta_1,$$

则称检验 $\phi(x)$ 是显著性水平为 α 的一致最优势检验, 记为 UMPT (uniformly most powerful test).

在某些特殊的情况下, UMPT 可以直接从 N-P 基本引理推出, 但要利用下述的结论.

引理 3.4 设 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验, Θ_{01} 是 Θ_0 的子集. 如果 $\phi(x)$ 是 $(\alpha, \Theta_{01}, \Theta_1)$ 检验的 UMPT, 则 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验的 UMPT.

证明: 如果 $\phi_1(x)$ 是任意一个 $(\alpha, \Theta_0, \Theta_1)$ 检验, 则有 $E_\theta(\phi_1(X)) \leq \alpha, \theta \in \Theta_0$. 所以 $\phi_1(x)$ 也是 $(\alpha, \Theta_{01}, \Theta_1)$ 检验. 由于 $\phi(x)$ 是 $(\alpha, \Theta_{01}, \Theta_1)$ 检验的 UMPT, 则应有

$$E_\theta(\phi(X)) \geq E_\theta(\phi_1(X)), \quad \forall \theta \in \Theta_1,$$

这表明 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验的 UMPT. 证毕.

引理 3.5 设 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验, 则 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验的 UMPT 的充要条件是, 对每一个 $\theta_1 \in \Theta_1$, $\phi(x)$ 是 $(\alpha, \Theta_0, \{\theta_1\})$ 检验的 MPT.

必要性和充分性可从 UMPT 的定义看出. 证明略.

定理 3.6 设 $\phi(x)$ 是 $(\alpha, \Theta_0, \Theta_1)$ 检验. 假如对某个 $\theta_0 \in \Theta_0$ 和对每一个 $\theta_1 \in \Theta_1$, $\phi(x)$ 都是 $(\alpha, \{\theta_0\}, \{\theta_1\})$ 检验的 MPT, 则 $\phi(x)$ 也是 $(\alpha, \Theta_0, \Theta_1)$ 检验的 UMPT.

这个定理是引理 3.4 和 3.5 的推论, 证明略.

在例 3.2 中, 由 N-P 基本引理已求得简单原假设 $H_0: \mu=0$ 对简单备择假设

$H_1: \mu=\mu_1 (\mu_1 > 0)$ 的显著性水平为 α 的 MPT $\phi(x)$, 其拒绝域为 $W = \left\{x: \bar{x} \geq \frac{U_{1-\alpha}}{\sqrt{n}}\right\}$.

这个检验与 μ_1 的具体数值无关, 只要求 $\mu_1 > 0$. 由于原假设 $H_0: \mu = 0$ 成立时, $E_0(\phi(X)) = \alpha$, 并且由于 $\phi(x)$ 的势函数 $g(\mu) = E_\mu(\phi(X))$ 是 μ 的严增函数, 则在 $\mu \leq 0$ 时, $E_\mu(\phi(X)) \leq E_0(\phi(X)) = \alpha$. 所以 $\phi(x)$ 是原假设 $H_0: \mu \leq 0$ 对备择假设 $H_1: \mu > 0$ 的显著性水平为 α 的检验. 根据定理 3.6, 这个检验 $\phi(x)$ 是 $H_0: \mu \leq 0$ 对 $H_1: \mu > 0$ 的 UMPT.

这个例子告诉我们, 如果简单原假设对简单备择假设的检验问题的 MPT 不依赖于备择假设的具体数值, 则可适当扩大备择假设, 而当势函数是单调函数时, 也可适当扩大原假设, 这样就可由 MPT 获得 UMPT. 这就扩大了 N-P 基本引理的应用范围. 例 3.4 也有类似的结果, 它是例 3.1 考虑的检验问题的 UMPT. 例 3.3 的检验函数 $\phi(x)$ (见式 (3.14)) 似依赖于 θ_1 , 事实上这个检验函数可等价地表示为

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

它与 θ_1 的具体数值无关, 只要求 $\theta_1 > 1$ 就行了. 所以, 与例 3.2、例 3.4 一样, 例 3.3 也有类似的结果. 它是原假设 $H_0: \theta \leq 1$ 对备择假设 $H_1: \theta > 1$ 的 UMPT.

一般来说, 对于复合假设检验问题, MPT 的 $\phi(x)$ 依赖于备择假设中的 θ 值, 则 UMPT 不一定存在. 那么在什么情况下存在呢? 在什么情况下不存在呢? 下面将引进单调似然比以及单参数指数型分布族的概念, 然后就下列几种检验问题进行讨论:

- (I) 原假设 $H_0: \theta \leq \theta_0$ 对备择假设 $H_1: \theta > \theta_0$;
- (II) 原假设 $H_0: \theta \geq \theta_0$ 对备择假设 $H_1: \theta < \theta_0$;
- (III) 原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$;
- (IV) 原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对备择假设 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$;
- (V) 原假设 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 对备择假设 $H_1: \theta_1 < \theta < \theta_2$.

检验问题 (I) 和 (II) 称为单边假设检验问题, 而检验问题 (III), (IV) 和 (V) 称为双边假设检验问题. 首先寻求单边假设检验问题的 UMPT. 这需要引进单调似然比的概念.

§ 3.3.2 单调似然比

定义 3.8 设 $\{p(x; \theta): \theta \in \Theta\}$ 是含有实参数 θ 的概率密度族, 其中 Θ 是实直线上的一个区间. 如果存在实值统计量 $T(X)$, 使得对任意的 $\theta_1 < \theta_2$, 都有

(1) 概率分布 P_{θ_1} 与 P_{θ_2} 是不同的;

(2) 似然比 $\lambda(x) = \frac{p(x; \theta_2)}{p(x; \theta_1)}$ 是 $T(x)$ 的非降函数 (或非增函数), 则称概率密

度族 $\{p(x; \theta) : \theta \in \Theta\}$ 关于 $T(x)$ 具有非降 (或非增) 单调似然比. 单调似然比简称为 MLR (monotone likelihood ratio).

单参数指数型分布族

$$p(x; \theta) = c(\theta) \exp\{Q(\theta)T(x)\}h(x), \quad (3.15)$$

其中 $c(\theta) > 0$. 假设 $Q(\theta)$ 是 θ 的严增函数 (或严减函数), 则在 $\theta_1 < \theta_2$ 时, 其似然比

$$\lambda(x) = \frac{p(x; \theta_2)}{p(x; \theta_1)} = \frac{c(\theta_2)}{c(\theta_1)} \exp\{[Q(\theta_2) - Q(\theta_1)]T(x)\}$$

是 $T(x)$ 的严增函数 (或严减函数). 故单参数指数型分布族 (3.15) 关于 $T(x)$ 具有非降 (或非增) MLR. 这里的 $T(X)$ 是充分统计量. 二项分布族、负二项分布族、Poisson 分布族、正态分布族 (均值已知, 方差未知或均值未知, 方差已知的情况) 和指数分布族, 以及来自这些分布族的样本分布族都是单参数指数型分布族, 它们关于其充分统计量都具有 MLR. 虽然均匀分布族以及来自它的样本分布族不是单参数指数型分布族, 但如果定义 $\frac{0}{0} = \infty$, 则均匀分布族以及来自它的样本分布族关于其充分统计量也都具有 MLR.

下面的定理揭示了 MLR 分布族的一个基本性质.

定理 3.7 设概率密度族 $\{p(x; \theta) : \theta \in \Theta \subseteq \mathbf{R}\}$ 关于 $T(x)$ 具有非降 MLR. 若 $\psi(t)$ 是 t 的一个非降函数, 则 $E_\theta(\psi(T(X)))$ 是 θ 的一个非降函数.

证明: 设 $\theta_1 < \theta_2$, 令

$$A = \{x : p(x; \theta_1) < p(x; \theta_2)\}, \quad B = \{x : p(x; \theta_1) > p(x; \theta_2)\}.$$

对任意的 $x_1 \in A$ 和 $x_2 \in B$, 有

$$\lambda(x_1) = \frac{p(x_1; \theta_2)}{p(x_1; \theta_1)} > 1, \quad \lambda(x_2) = \frac{p(x_2; \theta_2)}{p(x_2; \theta_1)} < 1.$$

因为似然比 $\lambda(x)$ 是 $T(x)$ 的非降函数, 则由 $\lambda(x_1) > \lambda(x_2)$ 可以推出 $T(x_1) \geq T(x_2)$. 由于 $\psi(t)$ 是 t 的一个非降函数, 所以 $\psi(T(x_1)) \geq \psi(T(x_2))$. 令

$$a = \inf\{\psi(T(x)) : x \in A\}, \quad b = \sup\{\psi(T(x)) : x \in B\},$$

则有 $a \geq b$. 从而

$$\begin{aligned} & E_{\theta_2}(\psi(T(X))) - E_{\theta_1}(\psi(T(X))) \\ &= \int \psi(T(x)) [p(x; \theta_2) - p(x; \theta_1)] d\mu(x) \end{aligned}$$

$$\geq a \int_A [p(x; \theta_2) - p(x; \theta_1)] d\mu(x) + b \int_B [p(x; \theta_2) - p(x; \theta_1)] d\mu(x), \quad (3.16)$$

由于

$$\int_{A \cup B} [p(x; \theta_2) - p(x; \theta_1)] d\mu(x) = 0,$$

所以

$$\int_B [p(x; \theta_2) - p(x; \theta_1)] d\mu(x) = - \int_A [p(x; \theta_2) - p(x; \theta_1)] d\mu(x),$$

从而

$$\begin{aligned} & E_{\theta_2}(\psi(T(X))) - E_{\theta_1}(\psi(T(X))) \\ & \geq (a - b) \int_A [p(x; \theta_2) - p(x; \theta_1)] d\mu(x) \geq 0. \end{aligned}$$

证毕.

推论 1° 若 $\psi(t)$ 是 t 的一个非增函数, 则 $E_{\theta}(\psi(T(X)))$ 是 θ 的一个非增函数.

推论 2° 如果概率密度族 $\{p(x; \theta) : \theta \in \Theta \subseteq \mathbf{R}\}$ 关于 $T(x)$ 具有非增 MLR, 那么它关于 $(-T(x))$ 具有非降 MLR. 由于在 $\psi(t)$ 是 t 的一个非降函数时, $\psi(t)$ 是 $(-t)$ 的一个非增函数, 所以 $E_{\theta}(\psi(T(X)))$ 是 θ 的一个非增函数; 而在 $\psi(t)$ 是 t 的一个非增函数时, $E_{\theta}(\psi(T(X)))$ 是 θ 的一个非降函数.

在概率密度族 $\{p(x; \theta) : \theta \in \Theta \subseteq \mathbf{R}\}$ 关于 $T(x)$ 具有非降 MLR 时, 取 $\psi(t) = t$, 则在 $\theta_1 < \theta_2$ 时

$$E_{\theta_1}(T(X)) \leq E_{\theta_2}(T(X)), \quad (3.17)$$

所以, 在概率密度族关于 $T(x)$ 具有非降 MLR 时, $T(X)$ 的均值 $E_{\theta}(T(X))$ 是 θ 的一个非降函数. 对任意给定的 t_0 , 分别在 $t > t_0$ 和 $t \leq t_0$ 时, 取 $\psi(t) = 0$ 和 1. 则在 $\theta_1 < \theta_2$ 时, 有

$$P_{\theta_1}(T(X) \leq t_0) \geq P_{\theta_2}(T(X) \leq t_0), \quad (3.18)$$

所以, 在概率密度族关于 $T(x)$ 具有非降 MLR 时, $T(X)$ 的分布函数 $P_{\theta}(T(X) \leq t)$ 是 θ 的一个非增函数.

譬如, 随机变量 $X \sim b(n, p)$, 其中 $0 < p < 1$ 为未知参数. 其概率密度为

$$p(x; p) = (1 - p)^n \exp\left\{\ln\left(\frac{p}{1 - p}\right) x\right\} \binom{n}{x},$$

这是单参数指数型分布族. 由于 $Q(p) = \ln\left(\frac{p}{1-p}\right)$ 是 p 的严增函数, 所以它关于 x 具有非降 MLR. 二项分布的均值 $E_p(X) = np$ 和分布函数 $P_p(X \leq x) = \sum_{k \leq x} \binom{n}{k} p^k (1-p)^{n-k}$ 分别是 p 的非降和非增函数. 事实上, 它们分别是 p 的严增和严降函数.

又设 $X = (X_1, X_2, \dots, X_n)$ 是来自 Poisson 分布总体 $\mathcal{P} = \{P(\lambda), \lambda > 0\}$ 的一个样本. 样本联合概率密度为

$$p(x; \lambda) = e^{-n\lambda} \exp\left\{\ln \lambda \sum_{i=1}^n x_i\right\} \left(\prod_{i=1}^n (x_i!)\right)^{-1},$$

这是单参数指数型分布族. 由于 $Q(\lambda) = \ln \lambda$ 是 λ 的严增函数, 所以它关于 $T(x) = \sum_{i=1}^n x_i$ 具有非降 MLR, $T(X) \sim P(n\lambda)$. $T(X)$ 的均值 $E_\lambda(T(X)) = n\lambda$ 和分布函数 $P_\lambda(T(X) \leq t) = \sum_{k \leq t} \frac{(n\lambda)^k}{k!} e^{-n\lambda}$ 分别是 λ 的非降和非增函数. 事实上, 它们分别是 λ 的严增和严降函数.

在概率密度族关于 $T(x)$ 具有非降 MLR 时, 如果考虑单边假设检验问题 (I), 则式 (3.17) 和式 (3.18) 皆表示, 应在 $T(x)$ 的值较大时拒绝原假设. 这个检验方法是符合人们的直观感觉的, 也是可信的. 下面我们将证明它是 UMPT.

§ 3.3.3 单边假设检验

定理 3.8 设单参数概率密度族 $\{p(x; \theta) : \theta \in \Theta \subseteq \mathbf{R}\}$ 关于实值统计量 $T(X)$ 具有非降 MLR, 则关于单边假设检验问题 (I),

(1) 存在显著性水平为 α 的 UMPT 的检验函数

$$\phi(T(x)) = \begin{cases} 1, & T(x) > c, \\ r, & T(x) = c, \\ 0, & T(x) < c, \end{cases} \quad (3.19)$$

其中常数 $r(0 \leq r \leq 1)$ 和 c 由下式确定:

$$E_{\theta_0}(\phi(T(X))) = \alpha. \quad (3.20)$$

(2) 这个检验的势函数

$$g(\theta) = E_\theta(\phi(T(X)))$$

是非降的,且在集合 $\{\theta: 0 < g(\theta) < 1\}$ 上是严增.

(3) 在一切使得 $E_{\theta_0}(\phi(X)) = \alpha$ 的检验函数中,由式(3.19)和式(3.20)所确定的检验函数 $\phi(T(x))$,使得对任意的 $\theta < \theta_0$, $E_{\theta}(\phi(X))$ 都达到最小.

证明: 先证明(1). 考虑简单原假设 $H_0: \theta = \theta_0$ 对简单备择假设 $H_1: \theta = \theta_1$ ($\theta_1 > \theta_0$) 的检验问题. 由 N-P 基本引理知,应在似然比统计量 $\lambda(x) = \frac{p(x; \theta_1)}{p(x; \theta_0)}$ 比较大的时候,拒绝原假设. 由于 $\lambda(x)$ 是 $T(x)$ 的非降函数,所以在 $T(x)$ 比较大的时候,拒绝原假设. 为此,我们采用形如式(3.19)的检验函数 $\phi(T(x))$. 用类似于定理 3.2 的证明方法可以知道,满足式(3.19)和式(3.20)的检验 $\phi(T(x))$ 是存在的. 具体地说,在 $T(X)$ 为连续型随机变量的时候, $\phi(T(X))$ 是非随机化检验,其拒绝域为 $W = \{T(X) \geq c\}$, c 由 $P_{\theta_0}\{T(X) \geq c\} = \alpha$ 确定. 在 $T(X)$ 为离散型随机变量的时候, c 由 $P_{\theta_0}\{T(X) > c\} \leq \alpha \leq P_{\theta_0}\{T(X) \geq c\}$ 确定,而

$$r = \frac{\alpha - P_{\theta_0}\{T(X) > c\}}{P_{\theta_0}\{T(X) = c\}}.$$

在 $T(x) = c$ 时,假设 $\lambda(x) = k$. 由于 $\lambda(x)$ 是 $T(x)$ 的非降函数,所以

$$\{x: \lambda(x) < k\} \subseteq \{x: T(x) < c\},$$

$$\{x: \lambda(x) > k\} \subseteq \{x: T(x) > c\},$$

因而满足式(3.19)的检验 $\phi(T(x))$ 必满足式(3.11). 检验 $\phi(T(x))$ 满足式(3.20),也就是满足式(3.10). 从而由 N-P 基本引理知. 满足式(3.19)和式(3.20)的检验 $\phi(T(x))$ 是该简单假设检验问题, $H_0: \theta = \theta_0$ 对 $H_1: \theta = \theta_1$ ($\theta_1 > \theta_0$) 的显著性水平为 α 的 MPT. 由于检验 $\phi(T(x))$ 与 θ_1 无关,只要求 $\theta_1 > \theta_0$, 所以由引理 3.5 知, $\phi(T(x))$ 也是 $H_0: \theta = \theta_0$ 对 $H_1: \theta > \theta_0$ 的检验问题的显著性水平为 α 的 UMPT.

由于 $\phi(T(x))$ 关于 $T(x)$ 非降,则据定理 3.7 可知,势函数 $g(\theta) = E_{\theta}(\phi(T(X)))$ 是 θ 的非降函数. 故在 $\theta \leq \theta_0$ 时, $E_{\theta}(\phi(T(X))) \leq E_{\theta_0}(\phi(T(X))) = \alpha$. 所以对于原假设 $H_0: \theta \leq \theta_0$, $\phi(T(x))$ 是显著性水平为 α 的检验. 则据引理 3.4 可知,由式(3.19)和式(3.20)确定的检验函数 $\phi(T(x))$ 是单边假设检验问题 $H_0: \theta \leq \theta_0$ 对 $H_1: \theta > \theta_0$ 的显著性水平为 α 的 UMPT. 结论(1)成立.

(2) 的证明. 我们已经证明了,势函数 $g(\theta) = E_{\theta}(\phi(T(X)))$ 是 θ 的非降函数. 接下来证明,它在集合 $\{\theta: 0 < g(\theta) < 1\}$ 上是严增的. 由 N-P 基本引理知, $\phi(T(x))$ 也是 $H_0: \theta = \theta_1$ 对 $H_1: \theta = \theta_2$ ($\theta_2 > \theta_1$) 的显著性水平为 $\alpha_1 = E_{\theta_1}(\phi(T(X)))$ 的 MPT,则在 $0 < \alpha_1 < 1$ 时,由定理 3.3 知, $E_{\theta_2}(\phi(T(X))) > \alpha_1$. 所以在集合 $\{\theta: 0 < g(\theta) < 1\}$ 上,势函数 $g(\theta) = E_{\theta}(\phi(T(X)))$ 是严增的. 结论(2)成立.

(3) 的证明. 考虑简单原假设 $H_0: \theta = \theta_0$ 对简单备择假设 $H_1: \theta = \theta_2 (\theta_2 < \theta_0)$ 的显著性水平为 $\alpha^* = 1 - \alpha$ 的检验问题. 由 N-P 基本引理知, 我们在似然比统计量 $\lambda(x) = \frac{p(x; \theta_2)}{p(x; \theta_0)}$ 比较大的时候, 拒绝原假设. 由于 $\theta_2 < \theta_0$, 所以 $\lambda(x)$ 是 $T(x)$ 的非增函数. 因此我们在 $T(x)$ 比较小的时候, 拒绝原假设. 用类似于 (1) 的证明方法可以知道, $\phi^*(T(x)) = 1 - \phi(T(x))$ 是 $H_0: \theta = \theta_0$ 对 $H_1: \theta = \theta_2 (\theta_2 < \theta_0)$ 的显著性水平为 $\alpha^* = 1 - \alpha$ 的 MPT, 其中 $\phi(T(x))$ 由式 (3.19) 和式 (3.20) 给出. 因此 $\phi^*(T(x))$ 在 $E_{\theta_0}(\phi(X)) \leq \alpha^*$ 的条件下, 使 $E_{\theta_2}(\phi(X))$ 达到最大. 从而 $\phi(T(x))$ 在 $E_{\theta_0}(\phi(X)) \geq \alpha$ 的条件下, 使 $E_{\theta_2}(\phi(X))$ 达到最小. 由于 θ_2 是小于 θ_0 的任意一个数, 所以对任意的 $\theta < \theta_0$, $\phi(T(x))$ 是在 $E_{\theta_0}(\phi(X)) \geq \alpha$, 当然也是在 $E_{\theta_0}(\phi(X)) = \alpha$ 的条件下, 使 $E_{\theta}(\phi(X))$ 达到最小的检验法. 结论 (3) 成立. 证毕.

关于这个定理有以下一些推论.

推论 1° 在定理 3.8 的条件下, 考虑

(1) 检验问题 $H_0: \theta = \theta_0$ 对 $H_1: \theta > \theta_0$, 则定理 3.8 的结论仍全部成立;

(2) 单边假设检验问题 (II), 则类似于定理 3.8 的结论全部成立, 只需要将式 (3.19) 中的不等号改变方向;

(3) 检验问题 $H_0: \theta = \theta_0$ 对 $H_1: \theta < \theta_0$, 则推论 1° (2) 的结论仍全部成立.

推论 2° 设单参数概率密度族 $\{p(x; \theta): \theta \in \Theta \subseteq \mathbf{R}\}$ 关于实值统计量 $T(X)$ 具有非增 MLR, 从而它关于 $-T(X)$ 具有非降 MLR, 定理 3.8 及其推论 1° 的结论仍全部成立, 只需要将 $T(X)$ 换为 $-T(X)$.

推论 3° 单参数指数型分布族 (3.15) 在 $Q(\theta)$ 是 θ 的严格单调函数时, 关于充分统计量 $T(X)$ 具有 MLR. 由定理 3.8 及其推论 1° 和 2° 可以知道, 对于单参数指数型分布族 (3.15) 的未知参数 θ 的单边假设检验问题, 都可以找到 UMPT.

推论 4° 设单参数概率密度族 $\{p(x; \theta): \theta \in \Theta \subseteq \mathbf{R}\}$ 关于实值统计量 $T(X)$ 具有非降 MLR, 式 (3.18) 告诉我们 $T(X)$ 的分布函数 $P_{\theta}(T(X) \leq t) = 1 - P_{\theta}(T(X) > t)$ 是 θ 的一个非增函数. 现由定理 3.8 (2) 可进一步推得, 在集合 $\{\theta: 0 < P_{\theta}(T(X) \leq t) < 1\}$ 上, $T(X)$ 的分布函数是 θ 的一个严降函数. 如果单参数概率密度族关于实值统计量 $T(X)$ 具有非增 MLR, 与上述相类似的结论仍然成立, 只需要将关于 $T(X)$ 的分布函数的结论中的非增换为非降, 严降换为严增. 这个结论将被用于第四章的置信区间 (限) 的问题.

例 3.5 若有 N 件产品, 其中含有 m 件不合格品. 今从中不放回地随机抽取 n 件进行检验. 设 n 件中含有 x 件不合格品. x 的可能取值为 $0, 1, 2, \dots, \min(m, n)$. X 服从超几何分布. 当 m 未知时, 考察超几何分布概率密度族 $\{p(x; m): m =$

$0, 1, 2, \dots, N\}$, 其中

$$p(x; m) = \frac{\binom{m}{x} \binom{N-m}{n-x}}{\binom{N}{n}}.$$

显然, 此概率密度族不是单参数指数型分布族. 因为

$$\begin{aligned} \frac{p(x; m+1)}{p(x; m)} &= \frac{\binom{m+1}{x} \binom{N-m-1}{n-x}}{\binom{m}{x} \binom{N-m}{n-x}} \\ &= \frac{(m+1)(N-m-n+x)}{(N-m)(m+1-x)} \end{aligned}$$

是 x 的严增函数, 所以此概率密度族关于 $T(x) = x$ 具有非降 MLR.

考察如下的单边假设检验问题:

$$H_0: 0 \leq m \leq m_0 \quad \text{对} \quad H_1: m_0 < m \leq N.$$

由定理 3.8 知, 此检验问题存在显著性水平为 α 的 UMPT, 其检验函数为

$$\phi(x) = \begin{cases} 1, & x > c, \\ \frac{(\alpha - \alpha_1) \binom{N}{n}}{\binom{m_0}{c} \binom{N-m_0}{n-c}}, & x = c, \\ 0, & x < c, \end{cases}$$

其中 c 满足条件

$$\sum_{x \geq c} \frac{\binom{m_0}{x} \binom{N-m_0}{n-x}}{\binom{N}{n}} > \alpha \geq \sum_{x \geq c+1} \frac{\binom{m_0}{x} \binom{N-m_0}{n-x}}{\binom{N}{n}} = \alpha_1.$$

设有 $N=1\,000$ 件产品, 其中含有 m 件不合格品. 今从中不放回地随机抽取 $n=100$ 件进行检验. 在 $m_0=10$ 时, 考察上述的单边假设检验问题. 取 $\alpha=0.05$, 则有

$$\sum_{x \geq 3} \frac{\binom{10}{x} \binom{990}{100-x}}{\binom{1000}{100}} = 1 - \sum_{x \leq 2} \frac{\binom{10}{x} \binom{990}{100-x}}{\binom{1000}{100}} = 0.0691 > 0.05,$$

$$\sum_{x \geq 4} \frac{\binom{10}{x} \binom{990}{100-x}}{\binom{1000}{100}} = 1 - \sum_{x \leq 3} \frac{\binom{10}{x} \binom{990}{100-x}}{\binom{1000}{100}} = 0.0122 < 0.05.$$

所以此检验问题的显著性水平为 $\alpha=0.05$ 的 UMPT 为

$$\phi(x) = \begin{cases} 1, & x \geq 4, \\ 0.6643, & x = 3, \\ 0, & x \leq 2. \end{cases}$$

例 3.6 设 $X=(X_1, X_2, \dots, X_n)$ 是来自正态分布族 $\{N(0, \sigma^2): \sigma^2 > 0\}$ 的样本. 求 $H_0: \sigma^2 \leq \sigma_0^2$ 对 $H_1: \sigma^2 > \sigma_0^2$ 的单边假设检验问题的显著性水平为 α 的 UMPT. 样本的联合密度为

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

其中 $(-2\sigma^2)^{-1} = Q(\sigma^2)$ 是 σ^2 的严增函数. 由定理 3.8 及其推论 3° 知, 其 UMPT 存在, 且检验函数仅依赖于充分统计量 $T(X) = \sum_{i=1}^n X_i^2$, 拒绝域为 $W = \{x: T(x) \geq c\}$, 其中 c 由 $E_{\sigma_0^2}(\phi(T(X))) = \alpha$ 确定. 在 $\sigma^2 = \sigma_0^2$ 时, $T(X) \sim \sigma_0^2 \chi^2(n)$. 故 $c = \sigma_0^2 \chi_{1-\alpha}^2(n)$.

取 $\sigma_0^2 = 1, n = 10$ 和 $\alpha = 0.1$. 该 UMPT 的拒绝域为

$$W = \left\{x: \sum_{i=1}^{10} x_i^2 \geq 15.99\right\},$$

其势函数 $g(\sigma^2)$ 的图形如图 3.2 上的粗曲线所示. 若令 $T_1(x) = \sum_{i=1}^{10} (x_i - \bar{x})^2$, 取另一检验的拒绝域为 $W_1 = \{x: T_1(x) \geq \sigma_0^2 \chi_{1-\alpha}^2(n-1)\}$. 可以证明, 这个检验也是该检验问题的显著性水平为 α 的一个检验. 在 $\sigma_0^2 = 1, n = 10$ 和 $\alpha = 0.1$ 时, 其势函数 $g_1(\sigma^2)$ 的图形如图 3.2 上的细曲线所示. 在 $\sigma^2 > 1$ 时, 粗曲线一致地在细曲线的上面, 这正是定理 3.8(1) 所说的. 在 $\sigma^2 = 1$ 时, 这两个检验的势函数值都是 0.1, 而在 $\sigma^2 < 1$ 时, 粗曲线一致地在细曲线的下面, 这正是定理 3.8(3) 所说的.

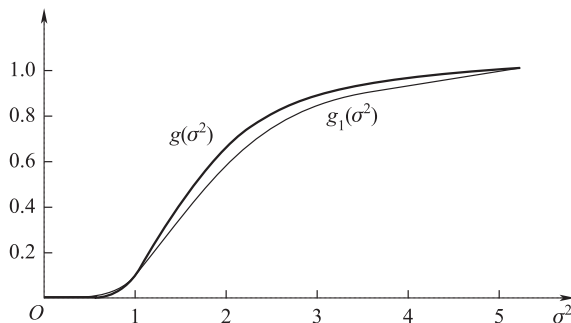


图 3.2 例 3.6 的两个检验的势函数图形的比较

§ 3.3.4 双边假设检验

对单参数指数型分布族, 双边假设检验问题 (V) 的 UMPT 是存在的. 但一般来说, 双边假设检验问题 (III) 和 (IV) 的 UMPT 是不存在的. 看下面的例子.

例 3.7 设 $X = (X_1, X_2, \dots, X_n)$ 为来自正态总体 $\{N(\mu, 1) : -\infty < \mu < \infty\}$ 的一个样本. 考虑原假设 $H_0: \mu = 0$ 对备择假设 $H_1: \mu \neq 0$ 的双边假设检验问题. 倘若 $\phi(x)$ 是它的显著性水平为 α 的 UMPT. 由于 $\phi^*(x) \equiv \alpha$ 显然是该检验问题的显著性水平为 α 的一个检验, 则有

$$E_\mu(\phi(X)) \geq E_\mu(\phi^*(X)) = \alpha, \quad \mu \neq 0. \quad (3.21)$$

此外, 据引理 3.5, $\phi(x)$ 也是 $H_0: \mu = 0$ 对 $H_1: \mu = \mu_1 (\mu_1 > 0)$ 的显著性水平为 α 的

MPT. 则由例 3.2 知, $\phi(x)$ 是拒绝域为 $W = \left\{x: \bar{x} \geq \frac{U_{1-\alpha}}{\sqrt{n}}\right\}$ 的一个检验, 其势函数

$g(\mu) = E_\mu(\phi(X)) = \Phi(\sqrt{n}\mu - U_{1-\alpha})$ 是 μ 的严增函数, 故在 $\mu < 0$ 时, $E_\mu(\phi(X)) < E_0(\phi(X)) = \alpha$. 这与式 (3.21) 相矛盾. 所以, 该双边假设检验问题不存在 UMPT.

同样, 原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对备择假设 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$ 的双边假设检验问题也不存在 UMPT. 这里的“不存在”是对所有的检验函数而言的. 如果我们对检验函数加上某种要求, 例如无偏性, 那么检验函数类就缩小了. 下一节, 我们将说明, 在缩小的检验函数类中存在 UMPT.

§ 3.3.5 N-P 基本引理的推广 (一)

要解决双边假设检验问题 (V) 的 UMPT 的存在性问题, 先要对 N-P 基本引理进行推广.

定理 3.9 设 μ 为可测空间 $(\mathcal{X}, \mathcal{B})$ 上的测度. $p_0(x), p_1(x), \dots, p_m(x)$ 为 $(\mathcal{X}, \mathcal{B})$ 上 $m+1$ 个实可测且 μ 可积函数. $\alpha_1, \alpha_2, \dots, \alpha_m$ 为给定的 m 个实常数. 如

果存在检验函数 $\phi(x)$ 满足以下两个条件:

(1) 若有 m 个非负常数 $k_1, k_2, \dots, k_m$, 使得

$$\phi(x) = \begin{cases} 1, & p_0(x) > \sum_{i=1}^m k_i p_i(x), \\ 0, & p_0(x) < \sum_{i=1}^m k_i p_i(x); \end{cases}$$

(2) $\int \phi(x) p_i(x) d\mu(x) = \alpha_i, i=1, 2, \dots, m.$

对任一满足下式的检验函数 $\phi^*(x)$,

$$\int \phi^*(x) p_i(x) d\mu(x) \leq \alpha_i, \quad i=1, 2, \dots, m, \quad (3.22)$$

都有

$$\int \phi(x) p_0(x) d\mu(x) \geq \int \phi^*(x) p_0(x) d\mu(x).$$

证明: 定理的证明类似于定理 3.2 的(2)的证明. 显然有

$$\left[p_0(x) - \sum_{i=1}^m k_i p_i(x) \right] [\phi(x) - \phi^*(x)] \geq 0,$$

于是

$$\begin{aligned} & \int p_0(x) [\phi(x) - \phi^*(x)] d\mu(x) \\ & \geq \sum_{i=1}^m k_i \int p_i(x) [\phi(x) - \phi^*(x)] d\mu(x), \end{aligned}$$

从而

$$\begin{aligned} & \int \phi(x) p_0(x) d\mu(x) - \int \phi^*(x) p_0(x) d\mu(x) \\ & \geq \sum_{i=1}^m k_i \left[\alpha_i - \int \phi^*(x) p_i(x) d\mu(x) \right] \geq 0. \end{aligned} \quad (3.23)$$

证毕.

在下一节,将进一步推广 N-P 基本引理. 这些定理中的结论实际上是一个充分条件. 类似于定理 3.2 的(1)和(3), 它们的存在性和必要性也是可以证明的^[1].

§ 3.3.6 单参数指数型分布族的双边假设检验问题(一)

定理 3.10 设样本 $X = (X_1, X_2, \dots, X_n)$ 服从单参数指数型分布(3.15), 其

中 θ 为实参数, $Q(\theta)$ 是 θ 的严增函数. 则关于双边假设检验问题 (V), 存在显著性水平为 $\alpha (0 < \alpha < 1)$ 的 UMPT, 其检验函数仅依赖于充分统计量 $T(X)$, 形如

$$\phi(T(x)) = \begin{cases} 1, & c_1 < T(x) < c_2, \\ r_i, & T(x) = c_i, \quad i = 1, 2, \\ 0, & T(x) < c_1, \text{ 或 } T(x) > c_2, \end{cases} \quad (3.24)$$

其中常数 $r_i (0 \leq r_i \leq 1)$ 和 c_i 由下式确定

$$E_{\theta_1}(\phi(T(X))) = E_{\theta_2}(\phi(T(X))) = \alpha. \quad (3.25)$$

证明: 从 $T(X)$ 是 θ 的充分统计量和定理 3.1 知, 最优检验要在由 $T(X)$ 所构成的检验函数类中寻找. 对于单参数指数型分布族, $T(X)$ 的导出分布族仍然是单参数指数型分布族. 考虑到 $Q(\theta)$ 是 θ 的严增函数, 为简化起见, 设 $T(x)$ 的密度函数为

$$p(t; \theta) = c(\theta) \exp \{ \theta t \} h(t), \quad (3.26)$$

其中 $c(\theta) > 0$.

先考虑原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta = \theta_3 (\theta_1 < \theta_3 < \theta_2)$ 的检验问题. 定理 3.9 启示我们, 取检验函数形如

$$\phi(t) = \begin{cases} 1, & p(t; \theta_3) > \sum_{i=1}^2 k_i p(t; \theta_i), \\ 0, & p(t; \theta_3) < \sum_{i=1}^2 k_i p(t; \theta_i), \end{cases} \quad (3.27)$$

并满足条件

$$\int \phi(x) p(t; \theta_1) dx = \int \phi(x) p(t; \theta_2) dx = \alpha, \quad (3.28)$$

式(3.27)化简后得到

$$\phi(t) = \begin{cases} 1, & a_1 e^{b_1 t} + a_2 e^{b_2 t} < 1, \\ 0, & a_1 e^{b_1 t} + a_2 e^{b_2 t} > 1, \end{cases}$$

其中 $a_i = \frac{k_i c(\theta_i)}{c(\theta_3)}$, $b_i = \theta_i - \theta_3, i = 1, 2$, 则 $b_1 < 0 < b_2$.

倘若 a_1 和 a_2 都是负的, 则检验总是拒绝的, 其势函数恒等于 1, 这与显著性水平 $\alpha (0 < \alpha < 1)$ 的条件相矛盾. 倘若 a_1 和 a_2 中, 一个是负的, 而另一个是正的, 或 a_1 和 a_2 中有一个为 0, 则 $a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 是严格单调的, 所得到的检验如式

(3.19)所示(见定理3.8,单边假设检验),其势函数是单调函数,它不可能满足式(3.28).所以, a_1 和 a_2 必定都是正的.从而 k_1 和 k_2 也都是正的.故定理3.9的一个条件:“ m 个非负常数 $k_1, k_2, \dots, k_m$ ”得到满足.

由于 $b_1 < 0 < b_2$,并且 a_1 和 a_2 都是正的,所以在 $t \rightarrow \pm\infty$ 时, $y = a_1 e^{b_1 t} + a_2 e^{b_2 t} \rightarrow +\infty$,并且 $y'' > 0$.故曲线 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 是严格下凸的(见图3.3).故存在常数 c_1 和 c_2 ,使得式(3.24)成立.因此,由式(3.27)和式(3.28)确定的检验函数就是由式(3.24)和式(3.25)确定的检验函数.

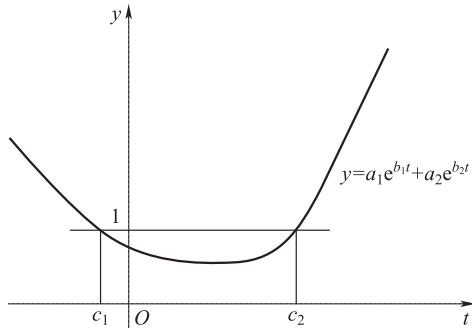


图 3.3 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t}$ ($b_1 < 0 < b_2, a_1, a_2 > 0$) 的图形

因为 k_1 和 k_2 都是正的,则据定理3.9知,由式(3.24)和式(3.25)确定的检验 $\phi(T(x))$ (也就是 $\phi(t)$)是原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta = \theta_3$ ($\theta_1 < \theta_3 < \theta_2$)的检验问题的显著性水平为 α 的MPT.因为 $\phi(t)$ 与 θ_3 无关,只要求 $\theta_1 < \theta_3 < \theta_2$ 就行了,所以它也是原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta_1 < \theta < \theta_2$ 的检验问题的UMPT.

为了应用引理3.4,完成定理的证明,即证明 $\phi(t)$ 是原假设 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 对备择假设 $H_1: \theta_1 < \theta < \theta_2$ 的检验问题的UMPT,尚需证明,对于原假设 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 而言, $\phi(t)$ 是显著性水平为 α 的一个检验.我们将把这个证明留待下一节.

§ 3.4 一致最优势无偏检验

§ 3.4.1 无偏检验

在例3.7,我们结合正态分布族 $\{N(\mu, 1): -\infty < \mu < \infty\}$ 指出,原假设 $H_0: \mu = 0$ 对备择假设 $H_1: \mu \neq 0$ 的双边假设检验问题(Ⅲ)不存在UMPT.下面对这个例子作进一步的分析.任意选取一个检验函数 $\phi(x)$,其对应的势函数

$g(\mu) = E_\mu(\phi(X))$ 是参数 μ 的函数. 当然, 我们希望在备择假设 $H_1: \mu \neq 0$ 成立时, $E_\mu(\phi(X))$ 越大越好. 因此, 很自然地要求选取的检验函数 $\phi(x)$ 的势函数 $E_\mu(\phi(X))$ 除了具有性质: $E_0(\phi(X)) \leq \alpha$ 外, 还应具有性质: $E_\mu(\phi(X)) \geq \alpha$, $\mu \neq 0$. 这样的 $\phi(x)$ 就称为无偏检验函数. 如果在所有的检验函数中无法找到 UMPT, 则可考虑在无偏检验函数类中去找 UMPT.

定义 3.9 设 $\phi(x)$ 是原假设 $H_0: \theta \in \Theta_0$ 对备择假设 $H_1: \theta \in \Theta_1$ 的检验函数, 若其势函数 $g(\theta) = E_\theta(\phi(X))$ 满足条件:

$$g(\theta) \leq \alpha, \quad \forall \theta \in \Theta_0,$$

$$g(\theta) \geq \alpha, \quad \forall \theta \in \Theta_1,$$

则称 $\phi(x)$ 为显著性水平为 α 的无偏检验.

显然, 显著性水平为 α 的 UMPT 一定是显著性水平为 α 的无偏检验.

定义 3.10 在检验问题 (Θ_0, Θ_1) 中, 如果存在一个显著性水平为 α 的无偏检验 $\phi(x)$, 对于任何显著性水平为 α 的无偏检验 $\phi_1(x)$, 下列条件均满足

$$E_\theta(\phi(X)) \geq E_\theta(\phi_1(X)), \quad \theta \in \Theta_1,$$

则称检验 $\phi(x)$ 是显著性水平为 α 的一致最优势无偏检验, 记为 UMPUT (uniformly most powerful unbiased test).

§ 3.4.2 相似检验

如果无偏检验 $\phi(x)$ 的势函数 $E_\theta(\phi(X))$ 是 θ 的连续函数, 则一定有

$$E_\theta(\phi(X)) = \alpha, \quad \theta \in \Gamma, \quad (3.29)$$

其中 Γ 是 Θ_0 和 Θ_1 的公共边界.

定义 3.11 称满足式 (3.29) 的检验 $\phi(x)$ 为边界相似检验 (border similar test).

令 $\Phi^{(\alpha)}$, $\Phi_u^{(\alpha)}$ 和 $\Phi_s^{(\alpha)}$ 分别表示所有显著性水平为 α 的检验, 所有显著性水平为 α 的无偏检验和所有的满足式 (3.29) 的边界相似检验构成的类. 显然, $\Phi_u^{(\alpha)} \subseteq \Phi^{(\alpha)}$, 并且在势函数是连续函数时, $\Phi_u^{(\alpha)} \subseteq \Phi_s^{(\alpha)}$. 它们之间的关系如图 3.4 所示, 其中 Φ 表示所有的检验函数构成的类.

一般来说, 构造边界相似检验比构造无偏检验更为方便. 下面的引理在构造 UMPUT 方面是有意义的.

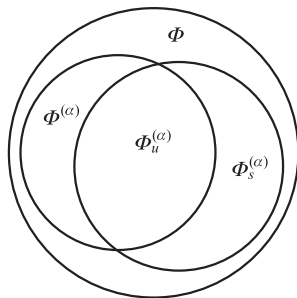


图 3.4 各种检验函数类之间的关系

引理 3.11 考虑原假设 $H_0: \theta \in \Theta_0$ 对备择假设 $H_1: \theta \in \Theta_1$ 的检验问题. 设每一个检验函数的势函数皆为 θ 的连续函数. 如果 $\phi(x)$ 是显著性水平为 α 的检验函数, 则在 $\phi(x)$ 是所有满足式 (3.29) 的边界相似检验函数类中的 UMPT 时, $\phi(x)$ 是显著性水平为 α 的 UMPUT.

证明: 显然, $\phi_1(x) \equiv \alpha$ 是边界相似的检验, 故

$$E_\theta(\phi(X)) \geq E_\theta(\phi_1(X)) = \alpha, \quad \forall \theta \in \Theta_1,$$

所以 $\phi(x)$ 是无偏检验. 由于满足式 (3.29) 的检验函数类 $\Phi_s^{(\alpha)}$ 包含了显著性水平为 α 的无偏检验函数类 $\Phi_u^{(\alpha)}$, 故 $\phi(x)$ 是显著性水平为 α 的 UMPUT. 证毕.

由定理 3.8 的证明过程知, MLR 分布族的单边假设检验问题的 UMPT 就是一致最优势边界相似检验.

§ 3.4.3 N-P 基本引理的推广(二)

要解决双边假设检验 (III) 和 (IV) (见 § 3.3.1) 的 UMPUT 问题, 需要对 N-P 基本引理作进一步的推广.

定理 3.12 定理的条件和 $\phi(x)$ 同定理 3.9, 只是常数 $k_1, k_2, \dots, k_m$ 可以取负值, 则对任一满足

$$\int \phi^*(x) p_i(x) d\mu(x) = \alpha_i, \quad i = 1, 2, \dots, m$$

的检验函数 $\phi^*(x)$, 都有

$$\int \phi(x) p_0(x) d\mu(x) \geq \int \phi^*(x) p_0(x) d\mu(x).$$

定理的证明完全类似于定理 3.9, 只是式 (3.23) 中的最后一个不等号要改为等号. 证毕.

同定理 3.9 后的说明, 这些定理的结论实际上是一个充分条件. 类似于定理 3.2 的 (1) 和 (3), 它们的存在性和必要性也是可以证明的^[1].

现回到 § 3.3.6 中定理 3.10 的证明. 我们已经证明了, 存在满足式 (3.24) 和式 (3.25) 的检验函数 $\phi(t)$, 它是原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta_1 < \theta < \theta_2$ 的检验问题的 UMPT. 为了应用引理 3.4, 完成定理的证明, 即证明 $\phi(t)$ 是 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 对 $H_1: \theta_1 < \theta < \theta_2$ 的检验问题的 UMPT, 尚需证明, 对于原假设 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 而言, $\phi(t)$ 是显著性水平为 α 的一个检验. 为此任取 $\theta' < \theta_1$ (若取 $\theta' > \theta_2$, 可同样处理), 考虑原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta = \theta'$ 的检验问题. 如果存在 k'_1 和 k'_2 , 使得

$$1 - \phi(t) = \begin{cases} 1, & p(t; \theta') > \sum_{i=1}^2 k'_i p(t; \theta_i), \\ 0, & p(t; \theta') < \sum_{i=1}^2 k'_i p(t; \theta_i), \end{cases} \quad (3.30)$$

则由 $E_{\theta_1}(1 - \phi(T)) = E_{\theta_2}(1 - \phi(T)) = 1 - \alpha$, 以及定理 3.12 知

$$E_{\theta'}(1 - \phi(T)) \geq E_{\theta'}(\phi^*(T)) = 1 - \alpha,$$

其中 $\phi^*(t) \equiv 1 - \alpha$. 从而 $E_{\theta'}(\phi(T)) \leq \alpha$. 这说明, 对于原假设 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 而言, $\phi(t)$ 是显著性水平为 α 的一个检验. 若找到使得式 (3.30) 成立的 k'_1 和 k'_2 , 则定理 3.10 证毕.

下面寻找使得式 (3.30) 成立的 k'_1 和 k'_2 . 由式 (3.24) 知

$$1 - \phi(t) = \begin{cases} 1, & t < c_1 \text{ 或 } t > c_2, \\ 0, & c_1 < t < c_2, \end{cases}$$

把它与式 (3.30) 进行比较. 由此可见, k'_1 和 k'_2 应满足条件:

$$p(t; \theta') > \sum_{i=1}^2 k'_i p(t; \theta_i) \Leftrightarrow t < c_1 \text{ 或 } t > c_2, \quad (3.31)$$

$$p(t; \theta') < \sum_{i=1}^2 k'_i p(t; \theta_i) \Leftrightarrow c_1 < t < c_2, \quad (3.32)$$

所以在 $t = c_1$ 和 $t = c_2$ 时, $p(t; \theta') = \sum_{i=1}^2 k'_i p(t; \theta_i)$. 从而有下列方程组:

$$\begin{cases} p(c_1; \theta') = k'_1 p(c_1; \theta_1) + k'_2 p(c_1; \theta_2), \\ p(c_2; \theta') = k'_1 p(c_2; \theta_1) + k'_2 p(c_2; \theta_2), \end{cases}$$

其中 c_1 和 c_2 已知, k'_1 和 k'_2 未知. 解此方程组, 得到 k'_1 和 k'_2 的值. 它们分别是

$$\begin{cases} k'_1 = \frac{c(\theta')}{c(\theta_1)} \frac{\exp\{\theta'c_1 + \theta_2c_2\} - \exp\{\theta'c_2 + \theta_2c_1\}}{\exp\{\theta_1c_1 + \theta_2c_2\} - \exp\{\theta_1c_2 + \theta_2c_1\}}, \\ k'_2 = \frac{c(\theta')}{c(\theta_2)} \frac{\exp\{\theta_1c_1 + \theta'c_2\} - \exp\{\theta_1c_2 + \theta'c_1\}}{\exp\{\theta_1c_1 + \theta_2c_2\} - \exp\{\theta_1c_2 + \theta_2c_1\}}, \end{cases} \quad (3.33)$$

由于 $\theta' < \theta_1 < \theta_2, c_1 < c_2$, 所以

$$(\theta_1c_1 + \theta_2c_2) - (\theta_1c_2 + \theta_2c_1) = (\theta_1 - \theta_2)(c_1 - c_2) > 0,$$

$$(\theta'c_1 + \theta_2c_2) - (\theta'c_2 + \theta_2c_1) = (\theta' - \theta_2)(c_1 - c_2) > 0,$$

$$(\theta_1c_1 + \theta'c_2) - (\theta_1c_2 + \theta'c_1) = (\theta_1 - \theta')(c_1 - c_2) < 0.$$

因此 $k'_1 > 0, k'_2 < 0$. 令

$$y = \frac{k'_1 p(t; \theta_1) + k'_2 p(t; \theta_2)}{p(t; \theta')} = a_1 e^{b_1 t} + a_2 e^{b_2 t},$$

其中 $a_1 = k'_1 \frac{c(\theta_1)}{c(\theta')} > 0, a_2 = k'_2 \frac{c(\theta_2)}{c(\theta')} < 0, b_1 = \theta_1 - \theta' > 0, b_2 = \theta_2 - \theta' > 0$, 且 $b_1 < b_2$. 在 $t = c_1$ 或 c_2 时, $y = a_1 e^{b_1 t} + a_2 e^{b_2 t} = 1$.

曲线 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 是单峰的. 自峰顶起向右 (即当 $t \rightarrow \infty$ 时), $y \rightarrow -\infty$; 向左 (即当 $t \rightarrow -\infty$ 时), $y \rightarrow 0$. $y = a_1 e^{b_1 t} + a_2 e^{b_2 t} (0 < b_1 < b_2, a_1 > 0, a_2 < 0)$ 的图形由图 3.5 给出. 由此可见,

$$p(t; \theta') > \sum_{i=1}^2 k'_i p(t; \theta_i) \Leftrightarrow 1 > a_1 e^{b_1 t} + a_2 e^{b_2 t} \Leftrightarrow t < c_1 \text{ 或 } t > c_2,$$

$$p(t; \theta') < \sum_{i=1}^2 k'_i p(t; \theta_i) \Leftrightarrow 1 < a_1 e^{b_1 t} + a_2 e^{b_2 t} \Leftrightarrow c_1 < t < c_2,$$

所以, 使得式 (3.31) 和式 (3.32), 即使得式 (3.30) 成立的 k'_1 和 k'_2 找到了, 如式 (3.33) 所示. 定理 3.10 证毕.

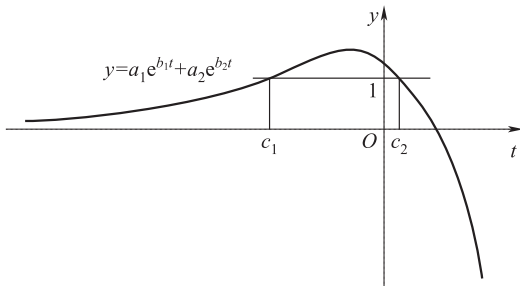


图 3.5 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t} (0 < b_1 < b_2, a_1 > 0, a_2 < 0)$ 的图形

§ 3.4.4 单参数指数型分布族的双边假设检验问题 (二)

定理 3.13 设样本 $X = (X_1, X_2, \dots, X_n)$ 服从单参数指数型分布 (3.15), 其中 θ 为实参数, $Q(\theta)$ 是 θ 的严增函数. 则关于原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对备择假设 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$ 的双边假设检验问题 (IV), 存在显著性水平为 $\alpha (0 < \alpha < 1)$ 的 UMPUT, 其检验函数为

$$\phi(T(x)) = \begin{cases} 1, & T(x) < c_1 \text{ 或 } T(x) > c_2, \\ r_i, & T(x) = c_i, \quad i = 1, 2, \\ 0, & c_1 < T(x) < c_2, \end{cases} \quad (3.34)$$

其中常数 $r_i (0 \leq r_i \leq 1)$ 和 c_i 由下式确定:

$$E_{\theta_1}(\phi(T(X))) = E_{\theta_2}(\phi(T(X))) = \alpha. \quad (3.35)$$

证明: 同定理 3.10, 我们仅在由 $T(x)$ 所构造的检验函数类中去找最优检验. $T(x)$ 的密度函数是式 (3.26). 先考虑原假设 $H_0: \theta = \theta_1$ 或 $\theta = \theta_2$ 对备择假设 $H_1: \theta = \theta_3$, 其中 $\theta_3 < \theta_1 < \theta_2$ (若 $\theta_1 < \theta_2 < \theta_3$, 可同样处理). 受定理 3.9 和定理 3.12 的启示, 取检验函数形如

$$\phi(t) = \begin{cases} 1, & p(t; \theta_3) > \sum_{i=1}^2 k_i p(t; \theta_i), \\ 0, & p(t; \theta_3) < \sum_{i=1}^2 k_i p(t; \theta_i), \end{cases} \quad (3.36)$$

并满足条件式 (3.35). 化简后得

$$\phi(t) = \begin{cases} 1, & a_1 e^{b_1 t} + a_2 e^{b_2 t} < 1, \\ 0, & a_1 e^{b_1 t} + a_2 e^{b_2 t} > 1, \end{cases}$$

其中 $a_i = \frac{k_i c(\theta_i)}{c(\theta_3)}$, $b_i = \theta_i - \theta_3$, $i=1, 2$, 则 $0 < b_1 < b_2$.

倘若 a_1 和 a_2 都是负的, 则检验总是拒绝的, 其势函数恒等于 1, 这与显著性水平为 $\alpha (0 < \alpha < 1)$ 的条件相矛盾. 倘若 a_1 和 a_2 都是正的, 或 a_1 和 a_2 中有一个为 0, 则 $a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 是严格单调的, 所得到的检验如式 (3.19) 所示 (见定理 3.8, 单边假设检验), 其势函数是单调函数, 它不可能满足式 (3.35). 所以在 a_1 和 a_2 中, 一个是负的, 而另一个是正的. 倘若 $a_1 < 0, a_2 > 0$, 那么在 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t} > 0$ 时, 由于 $a_1 e^{b_1 t} > -a_2 e^{b_2 t}$, 且 $0 < b_1 < b_2$, 则

$$\begin{aligned} y' &= a_1 b_1 e^{b_1 t} + a_2 b_2 e^{b_2 t} \\ &> b_1 (-a_2 e^{b_2 t}) + a_2 b_2 e^{b_2 t} \\ &= a_2 e^{b_2 t} (-b_1 + b_2) > 0. \end{aligned}$$

这就是说, 在 t 轴的上方, 曲线 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 是上升的. 所得到的检验也如式 (3.19) 所示, 其势函数是单调函数, 该检验不可能满足式 (3.35). 所以, a_1 必定是正的, 而 a_2 必定是负的. 从而 $k_1 > 0, k_2 < 0$. 这就使得我们无法使用定理 3.9 得到 UMPT, 而只能使用定理 3.12 得到 UMPUT.

在 $0 < b_1 < b_2, a_1 > 0, a_2 < 0$ 时, 根据函数 $y = a_1 e^{b_1 t} + a_2 e^{b_2 t}$ 的图形可以知道, 存在常数 c_1 和 c_2 , 使得式 (3.34) 成立. 因此, 由式 (3.34) 和式 (3.35) 确定的检验函

数就是由式(3.36)和式(3.35)确定的检验函数. 据定理 3.12, 由式(3.34)和式(3.35)确定的检验函数 $\phi(T(x))$ (也就是 $\phi(t)$) 具有这样一个性质: 对任一满足 $E_{\theta_1}(\phi^*(T)) = E_{\theta_2}(\phi^*(T)) = \alpha$ 的检验函数 $\phi^*(t)$, 都有 $E_{\theta_3}(\phi(T)) \geq E_{\theta_3}(\phi^*(T))$. 由于检验 $\phi(t)$ 与 θ_3 无关, 只要求 $\theta_3 < \theta_1 < \theta_2$, 或 $\theta_1 < \theta_2 < \theta_3$ 就行了, 所以由式(3.34)和式(3.35)确定的检验 $\phi(t)$ 是原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对备择假设 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$ 的双边假设检验问题的一致最优势边界相似检验. 为了应用引理 3.11 完成定理的证明, 只需证明: 对于原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 而言, $\phi(t)$ 是显著性水平为 α 的一个检验. 运用与证明定理 3.10 相应结论完全类似的方法与步骤, 可以得到其证明(证明过程从略). 定理证毕.

定理 3.14 设样本 $X = (X_1, X_2, \dots, X_n)$ 服从单参数指数型分布(3.15), 其中 θ 为实值参数, $Q(\theta)$ 是关于 θ 的严增函数. 则关于原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的双边假设检验问题(III), 存在显著性水平为 α ($0 < \alpha < 1$) 的 UMPUT, 其检验函数由式(3.34)给出, 其中的常数 r_i ($0 \leq r_i \leq 1$) 和 c_i 由以下两式确定

$$E_{\theta_0}(\phi(T(X))) = \alpha, \quad (3.37)$$

$$E_{\theta_0}(T(X)\phi(T(X))) = \alpha E_{\theta_0}(T(X)). \quad (3.38)$$

证明: 同定理 3.13, 我们仅在由 $T(x)$ 所构造的检验函数类中去找最优检验. $T(x)$ 的密度函数是式(3.26).

首先证明, 原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的任意一个无偏检验 $\phi(t)$ 必满足式(3.37)和式(3.38). 显然, 由检验函数 $\phi(t)$ 的无偏性可知, 其势函数 $g(\theta) = E_{\theta}(\phi(T))$ 在 $\theta = \theta_0$ 处达到最小值 α , 故 $g(\theta_0) = \alpha, g'(\theta_0) = 0$. 由 $g(\theta_0) = \alpha$ 得式(3.37). 由于 $g(\theta) = E_{\theta}(\phi(T)) = \int \phi(t)c(\theta)\exp\{\theta t\}h(t)d\mu(t)$, 所以

$$\begin{aligned} g'(\theta) &= \int \phi(t)c'(\theta)\exp\{\theta t\}h(t)d\mu(t) + \int \phi(t)c(\theta)t\exp\{\theta t\}h(t)d\mu(t) \\ &= \frac{c'(\theta)}{c(\theta)}E_{\theta}(\phi(T)) + E_{\theta}(T\phi(T)). \end{aligned} \quad (3.39)$$

由于 $\frac{c'(\theta)}{c(\theta)} = -E_{\theta}(T)$ (见习题 3.10), 则有

$$g'(\theta) = -E_{\theta}(T)E_{\theta}(\phi(T)) + E_{\theta}(T\phi(T)),$$

从而由 $g'(\theta_0) = 0$, 以及式(3.37), 得 $0 = -E_{\theta_0}(T) \cdot \alpha + E_{\theta_0}(T\phi(T))$, 则有式(3.38).

接下来考虑原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta = \theta_1 (\theta_1 \neq \theta_0)$ 的检验问题, 并要求检验函数 $\phi(t)$ 满足式 (3.37) 和式 (3.38), 即以下两式:

$$\begin{aligned}\int \phi(t) p(t; \theta_0) d\mu(t) &= \alpha, \\ \int \phi(t) t p(t; \theta_0) d\mu(t) &= \alpha E_{\theta_0}(T).\end{aligned}$$

定理 3.12 启示我们, 取检验函数形如

$$\phi(t) = \begin{cases} 1, & p(t; \theta_1) > k_1 p(t; \theta_0) + k_2 t p(t; \theta_0), \\ 0, & p(t; \theta_1) < k_1 p(t; \theta_0) + k_2 t p(t; \theta_0), \end{cases} \quad (3.40)$$

并满足条件式 (3.37) 和式 (3.38). 化简后得

$$\phi(t) = \begin{cases} 1, & a_1 + a_2 t < e^{bt}, \\ 0, & a_1 + a_2 t > e^{bt}, \end{cases}$$

其中 $a_i = \frac{k_i c(\theta_0)}{c(\theta_1)}, i=1, 2, b = \theta_1 - \theta_0$.

上述不等式两边, 右边是指数函数, 左边是线性函数. 倘若指数曲线 $y = e^{bt}$ 与直线 $y = a_1 + a_2 t$ 没有交点, 则指数曲线在直线的上方. 故检验总是拒绝的, 其势函数恒等于 1, 这与显著性水平 $\alpha (0 < \alpha < 1)$ 的条件相矛盾. 倘若指数曲线与直线相交一点, 则所得到的检验如式 (3.19) 所示 (见定理 3.8, 单边假设检验), 其势函数是严格单调函数, 它不可能满足式 (3.38) (为什么? 见习题 3.11). 所以指数曲线与直线相交两点, 即存在常数 c_1 和 c_2 , 使得式 (3.34) 成立. 因此, 由式 (3.40)、式 (3.37) 和式 (3.38) 确定的检验函数就是由式 (3.34)、式 (3.37) 和式 (3.38) 确定的检验函数.

由式 (3.34)、式 (3.37) 和式 (3.38) 确定的检验函数 $\phi(T(x))$ (也就是 $\phi(t)$) 与 θ_1 无关, 只要求 $\theta_1 \neq \theta_0$ 就行了. 据定理 3.12, 该检验函数 $\phi(t)$ 具有这样一个性质: 对任一满足式 (3.37) 和式 (3.38) 的检验函数 $\phi^*(t)$, 都有 $E_{\theta_1}(\phi(T)) \geq E_{\theta_1}(\phi^*(T))$. 若取 $\phi^*(t) \equiv \alpha$, 它满足式 (3.37) 和式 (3.38), 则由 $\phi(t)$ 具有的这样一个性质, 有 $E_{\theta_1}(\phi(T)) \geq E_{\theta_1}(\phi^*(T)) = \alpha, \forall \theta_1 \neq \theta_0$. 所以该检验函数 $\phi(t)$ 是原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的一个无偏检验, 其势函数 $g(\theta) = E_{\theta}(\phi(T))$ 在 $\theta = \theta_0$ 处达到最小值 α . 又取 $\phi^*(t)$ 为原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的任意一个无偏检验, 它满足式 (3.37) 和式 (3.38), 则由 $\phi(t)$ 具有这样一个性质, 有 $E_{\theta_1}(\phi(T)) \geq E_{\theta_1}(\phi^*(T)), \forall \theta_1 \neq \theta_0$. 所以该检验函数 $\phi(t)$ 是原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的显著性水平为 α 的 UMPUT.

定理证毕.

关于这个定理有三点说明.

1° 由定理 3.10 和定理 3.13 的证明过程知,单参数指数型分布族的双边假设检验问题(V)的 UMPT 就是一致最优势边界相似检验;双边假设检验问题(IV)的 UMPUT 就是一致最优势边界相似检验. 而关于双边假设检验问题(III),由定理 3.14 的证明过程知,其 UMPUT 就是所有在边界上满足条件(3.37)和(3.38)的检验函数类中的一致最优势检验. 此外,如本节前面所述的,单参数指数型分布族的单边假设检验问题的 UMPT 就是一致最优势边界相似检验. 这些结论将被用于多参数指数型分布族的假设检验问题.

2° 如果在 $\theta = \theta_0$ 时, $T(x)$ 的分布关于某个数 r_0 是对称的,则双边假设检验问题(III)的 UMPUT 可以简化为如下的形式:

$$\phi(T(x)) = \begin{cases} 1, & T(x) < r_0 - \delta \text{ 或 } T(x) > r_0 + \delta, \\ r, & T(x) = r_0 \pm \delta, \\ 0, & r_0 - \delta < T(x) < r_0 + \delta. \end{cases} \quad (3.41)$$

当其中的常数 δ 与 $r(0 \leq r \leq 1)$ 由

$$P_{\theta_0}(T < r_0 - \delta) + rP_{\theta_0}(T = r_0 - \delta) = \frac{\alpha}{2}$$

确定时,式(3.37)与式(3.38)一定成立. 其证明如下:由 $T(x)$ 的分布关于 r_0 是对称的,可以推得

$$\begin{aligned} E_{\theta_0}(\phi(T)) &= P_{\theta_0}(T < r_0 - \delta) + P_{\theta_0}(T > r_0 + \delta) + rP_{\theta_0}(T = r_0 \pm \delta) \\ &= 2(P_{\theta_0}(T < r_0 - \delta) + rP_{\theta_0}(T = r_0 - \delta)) = \alpha, \end{aligned}$$

式(3.37)成立. 从而由 $E_{\theta_0}(T) = r_0$ 以及 $E_{\theta_0}((T - r_0)\phi(T)) = 0$ 可进一步推得

$$\begin{aligned} E_{\theta_0}(T\phi(T)) &= E_{\theta_0}((T - r_0)\phi(T)) + r_0 E_{\theta_0}(\phi(T)) \\ &= r_0 \alpha = \alpha E_{\theta_0}(T), \end{aligned}$$

式(3.38)也成立. 所以在 $\theta = \theta_0$ 时,如果 $T(x)$ 的分布关于某个数 r_0 是对称的,那么关于双边假设检验问题(III)的 UMPUT 由式(3.41)给出.

3° 设 $T(x)$ 的密度函数形如 $p(t - \theta)$, 其中 $p(t)$ 关于原点对称. 这说明 $T(x)$ 的分布关于 θ 是对称的,即对任意的 r 都有, $P_{\theta}(T < \theta - r) = P_{\theta}(T > \theta + r)$. 由 2° 知,关于双边假设检验问题(III),只要取 $r_0 = \theta_0$ 就可得到形如式(3.41)的 UMPUT. 此外,关于双边假设检验问题(IV)与(V),也有类似的结论. 对前者,取检验函数形如

$$\phi(T(x)) = \begin{cases} 1, & T(x) < \theta_1 - \delta \text{ 或 } T(x) > \theta_2 + \delta, \\ r, & T(x) = \theta_1 - \delta \text{ 或 } T(x) = \theta_2 + \delta, \\ 0, & \theta_1 - \delta < T(x) < \theta_2 + \delta. \end{cases} \quad (3.42)$$

在 $T \sim p(t - \theta_1)$ 时, 有 $T - \theta_1 + \theta_2 \sim p(t - \theta_2)$, 并且对任意的 a , 有 $P_\theta(T < \theta - a) = P_\theta(T > \theta + a)$. 所以

$$\begin{aligned} E_{\theta_1}(\phi(T)) &= P_{\theta_1}(T < \theta_1 - \delta) + P_{\theta_1}(T > \theta_2 + \delta) + \\ &\quad r[P_{\theta_1}(T = \theta_1 - \delta) + P_{\theta_1}(T = \theta_2 + \delta)] \\ &= P_{\theta_2}(T < \theta_2 - \delta) + P_{\theta_2}(T > \theta_2 + \theta_2 - \theta_1 + \delta) + \\ &\quad r[P_{\theta_2}(T = \theta_2 - \delta) + P_{\theta_2}(T = \theta_2 + \theta_2 - \theta_1 + \delta)] \\ &= P_{\theta_2}(T > \theta_2 + \delta) + P_{\theta_2}(T < \theta_1 - \delta) + \\ &\quad r[P_{\theta_2}(T = \theta_2 + \delta) + P_{\theta_2}(T = \theta_1 - \delta)] \\ &= E_{\theta_2}(\phi(T)). \end{aligned}$$

据定理 3.13 可知, 如果式 (3.42) 中的常数 r 和 δ 由下式确定:

$$E_{\theta_1}(\phi(T)) = \alpha \text{ (或 } E_{\theta_2}(\phi(T)) = \alpha), \quad (3.43)$$

则由式 (3.42) 给出的检验函数是双边假设检验问题 (IV) 的 UMPUT. 类似地, 取检验函数形如

$$\phi(T(x)) = \begin{cases} 1, & \theta_1 - \delta < T(x) < \theta_2 + \delta, \\ r, & T(x) = \theta_1 - \delta \text{ 或 } T(x) = \theta_2 + \delta, \\ 0, & T(x) < \theta_1 - \delta \text{ 或 } T(x) > \theta_2 + \delta, \end{cases}$$

其中的常数 r 和 δ 由式 (3.43) 确定, 则据定理 3.10 知, 该检验函数是双边假设检验问题 (V) 的 UMPT.

例 3.8 样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态分布族 $\{N(\mu, 1) : -\infty < \mu < \infty\}$. 考虑双边假设检验问题 (1) $H_0: \mu = 0$ 对 $H_1: \mu \neq 0$; (2) $H_0: -1 \leq \mu \leq 1$ 对 $H_1: \mu < -1$ 或 $\mu > 1$; (3) $H_0: \mu \leq -1$ 或 $\mu \geq 1$ 对 $H_1: -1 < \mu < 1$.

样本 X 的联合密度函数

$$p(x; \sigma^2) = c(\mu) \exp\{Q(\mu)T(x)\}h(x),$$

其中 $Q(\mu) = n\mu$ 是 μ 的严增函数, 充分统计量 $T(X) = \bar{X} \sim N(\mu, 1/n)$.

关于检验问题 (1), 因为在 $\mu = 0$ 时, $\bar{X} \sim N(0, 1/n)$, 其分布关于原点对称. 由

上述 2° 的结论知, 检验函数 $\phi(t)$ 的拒绝域取为 $\{x: \bar{x} \leq -\delta \text{ 或 } \bar{x} \geq \delta\}$, δ 由下式确定:

$$E(\phi(T) | \mu = 0) = 1 - P(-\delta < \bar{X} < \delta | \mu = 0) = \alpha,$$

则据定理 3.14 知, 检验问题(1)的显著性水平为 α 的 UMPUT 的拒绝域为 $\{x:$

$|\bar{x}| \geq \frac{1}{\sqrt{n}} U_{1-\alpha/2}\}$. 当 $n=10$ 且 $\alpha=0.05$ 时, 拒绝域为 $\{x: |\bar{x}| \geq 0.6198\}$, 其势函数的图形如图 3.6(1) 所示.

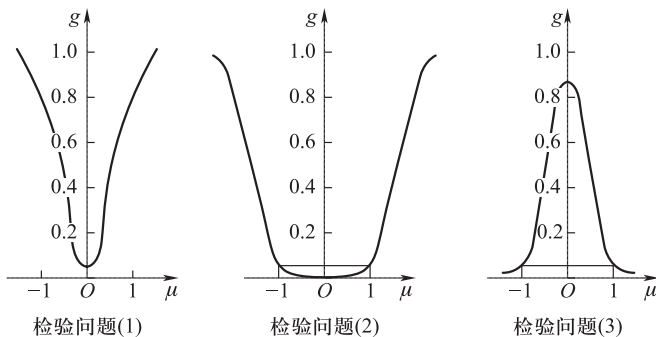


图 3.6 例 3.8 的势函数的图形

关于检验问题(2), 因为充分统计量 $\bar{X} \sim N(\mu, 1/n)$, 其分布关于 μ 对称, 则由上述 3° 的结论知, 检验函数 $\phi(t)$ 的拒绝域取为 $\{x: \bar{x} \leq -1-\delta \text{ 或 } \bar{x} \geq 1+\delta\}$, δ 由下式确定:

$$E(\phi(T) | \mu = 1) = 1 - P(-1 - \delta < \bar{X} < 1 + \delta | \mu = 1) = \alpha.$$

据定理 3.13 知, 这是检验问题(2)的显著性水平为 α 的 UMPUT. 当 $n=10$ 且 $\alpha=0.05$ 时, δ 由下式确定:

$$\begin{aligned} & P(-1 - \delta < \bar{X} < 1 + \delta | \mu = 1) \\ &= \Phi(\sqrt{10}\delta) - \Phi(\sqrt{10}(-2 - \delta)) = 0.95. \end{aligned}$$

取 $\sqrt{10}\delta = U_{0.95} = 1.64$, $\delta = 0.5186$. 则有

$$\begin{aligned} & P(-1 - \delta < \bar{X} < 1 + \delta | \mu = 1) \\ &= \Phi(1.64) - \Phi(-7.9645) \approx 0.95. \end{aligned}$$

检验的拒绝域为 $\{x: |\bar{x}| \geq 1.5186\}$. 其势函数的图形如图 3.6(2) 所示.

类似地, 关于检验问题(3), 取检验函数 $\phi(t)$ 的拒绝域形如 $\{x: -1-\delta \leq \bar{x} \leq 1+\delta\}$. 若取 δ 满足条件

$$E(\phi(T) | \mu = 1) = P(-1 - \delta \leq \bar{X} \leq 1 + \delta) = \alpha,$$

则据定理 3.10 知,该检验为检验问题(3)的显著性水平为 α 的 UMPT. 当 $n=10$ 且 $\alpha=0.05$ 时, δ 由下式确定:

$$\begin{aligned} & P(-1 - \delta < \bar{X} < 1 + \delta | \mu = 1) \\ &= \Phi(\sqrt{10}\delta) - \Phi(\sqrt{10}(-2 - \delta)) = 0.05. \end{aligned}$$

取 $\sqrt{10}\delta = U_{0.05} = -1.64$, $\delta = -0.5186$. 则有

$$\begin{aligned} & P(-1 - \delta < \bar{X} < 1 + \delta | \mu = 1) \\ &= \Phi(-1.64) - \Phi(-4.6846) \approx 0.05, \end{aligned}$$

检验的拒绝域为 $\{x: |\bar{x}| \leq 0.4814\}$. 其势函数的图形如图 3.6(3) 所示.

例 3.9 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态分布族 $\{N(0, \sigma^2): \sigma^2 > 0\}$. 考虑检验问题

$$(1) H_0: \sigma^2 \leq \sigma_1^2 \text{ 或 } \sigma^2 \geq \sigma_2^2, H_1: \sigma_1^2 < \sigma^2 < \sigma_2^2;$$

$$(2) H_0: \sigma_1^2 \leq \sigma^2 \leq \sigma_2^2, H_1: \sigma^2 < \sigma_1^2 \text{ 或 } \sigma^2 > \sigma_2^2;$$

$$(3) H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2.$$

求问题(1)的显著性水平为 α ($0 < \alpha < 1$) 的 UMPT; 问题(2)和(3)的显著性水平为 α ($0 < \alpha < 1$) 的 UMPUT.

解: 样本 X 的联合密度函数

$$p(x; \sigma^2) = c(\sigma^2) \exp\{Q(\sigma^2)T(x)\},$$

其中 $Q(\sigma^2) = (-2\sigma^2)^{-1}$ 是 σ^2 的严增函数, 充分统计量 $T(X) = \sum_{i=1}^n X_i^2 \sim \sigma^2 \chi^2(n)$.

这说明 $T(X)/\sigma^2 \sim \chi^2(n)$.

根据定理 3.10, 问题(1)的显著性水平为 α 的 UMPT 的拒绝域为 $W = \{x: c_1 \leq T(X) \leq c_2\}$, 其中的 c_1 和 c_2 由下式确定:

$$P_{\sigma_1^2}(c_1 \leq T(X) \leq c_2) = P_{\sigma_2^2}(c_1 \leq T(X) \leq c_2) = \alpha.$$

这就是说, c_1 和 c_2 由下式确定:

$$P_{\sigma_1^2}\left(\frac{c_1}{\sigma_1^2} \leq \frac{T(X)}{\sigma_1^2} \leq \frac{c_2}{\sigma_1^2}\right) = P_{\sigma_2^2}\left(\frac{c_1}{\sigma_2^2} \leq \frac{T(X)}{\sigma_2^2} \leq \frac{c_2}{\sigma_2^2}\right) = \alpha,$$

因此 c_1 和 c_2 由以下两式确定:

$$\int_{c_1/\sigma_1^2}^{c_2/\sigma_1^2} \chi^2(x | n) dx = \alpha, \quad \int_{c_1/\sigma_2^2}^{c_2/\sigma_2^2} \chi^2(x | n) dx = \alpha,$$

其中 $\chi^2(x|n)$ 表示自由度为 n 的 χ^2 分布的密度函数.

根据定理 3.13, 问题(2)的显著性水平为 α 的 UMPUT 的拒绝域为 $W = \{x: T(x) \leq c_1 \text{ 或 } T(x) \geq c_2\}$. 与上述相类似地, 其中的 c_1 和 c_2 由以下两式确定:

$$\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2} \chi^2(x|n) dx = 1 - \alpha, \quad \int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2} \chi^2(x|n) dx = 1 - \alpha.$$

根据定理 3.14, 问题(3)的显著性水平为 α 的 UMPUT $\phi(t)$ 的拒绝域为 $W = \{x: T(x) \leq c_1 \text{ 或 } T(x) \geq c_2\}$. 由式(3.37), 则有

$$E_{\sigma_0^2}(1 - \phi(T)) = P_{\sigma_0^2}\left(\frac{c_1}{\sigma_0^2} \leq \frac{T(X)}{\sigma_0^2} \leq \frac{c_2}{\sigma_0^2}\right) = 1 - \alpha,$$

即

$$\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2} \chi^2(x|n) dx = 1 - \alpha, \quad (3.44)$$

由式(3.38)得 $E_{\sigma_0^2}(T\phi(T)) = \alpha E_{\sigma_0^2}(T)$. 由于 $E_{\sigma_0^2}(T) = n\sigma_0^2$, 从而有

$$E_{\sigma_0^2}(T(1 - \phi(T))) = n\sigma_0^2(1 - \alpha) = n\sigma_0^2 E_{\sigma_0^2}(1 - \phi(T)),$$

$$E_{\sigma_0^2}\left(\left(\frac{T}{\sigma_0^2} - n\right)(1 - \phi(T))\right) = 0,$$

这里 $\frac{T}{\sigma_0^2} \sim \chi^2(n)$, 则有

$$\int_{c_1/\sigma_0^2}^{c_2/\sigma_0^2} (x - n) \chi^2(x|n) dx = 0, \quad (3.45)$$

故问题(3)的 UMPUT 中的 c_1 和 c_2 可由式(3.44)和式(3.45)确定. 另外, 由势

函数 $g(\sigma^2) = 1 - \int_{c_1/\sigma^2}^{c_2/\sigma^2} \chi^2(x|n) dx$ 在 $\sigma^2 = \sigma_0^2$ 时达到最小值, 推得等式

$$c_2 \chi^2\left(\frac{c_2}{\sigma_0^2} \middle| n\right) - c_1 \chi^2\left(\frac{c_1}{\sigma_0^2} \middle| n\right) = 0, \quad (3.46)$$

则问题(3)的 UMPUT 中, c_1 和 c_2 也可由式(3.44)和式(3.46)确定.

由式(3.44)和式(3.45)确定, 或由式(3.44)和式(3.46)确定 c_1 和 c_2 过于复杂、不实用. 常简单地取

$$c_1 = \chi_{\alpha/2}^2(n) \sigma_0^2, \quad c_2 = \chi_{1-\alpha/2}^2(n) \sigma_0^2,$$

这能使得式(3.44)成立, 但不能使得式(3.45), 或式(3.46)成立.

§ 3.5 多参数指数型分布族的假设检验

§ 3.5.1 多参数指数型分布族

单参数指数型分布族的单边和双边假设检验问题,基本上都被解决了.但是有些总体包含两个或两个以上的未知参数,当仅对其中一个未知参数做假设检验时,其他的未知参数就成了多余参数.一般来说,即使对单边假设检验问题而言,其 UMPT 也是不存在的,但是有例外的情况.例如,样本 $X = (X_1, X_2, \dots, X_n)$ 来自均值和方差皆未知的正态分布族 $\{N(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 > 0\}$. 考虑关于方差 σ^2 的单边假设检验问题, (1) 原假设 $H_0: \sigma^2 \leq \sigma_0^2$ 对备择假设 $H_1: \sigma^2 > \sigma_0^2$; (2) 原假设 $H_0: \sigma^2 \geq \sigma_0^2$ 对备择假设 $H_1: \sigma^2 < \sigma_0^2$. 这时均值 μ 就是多余参数. 常用的检验是 χ^2 检验. 它们的拒绝域分别为 $\left\{x: \sum_{i=1}^n (x_i - \bar{x})^2 \geq \sigma_0^2 \chi_{1-\alpha}^2(n-1)\right\}$ 和 $\left\{x: \sum_{i=1}^n (x_i - \bar{x})^2 \leq \sigma_0^2 \chi_{\alpha}^2(n-1)\right\}$. Lehmann(莱曼)和 Stein(斯坦)在 1948 年证明了,问题(1)的 χ^2 检验是 UMPT,但问题(2)的 χ^2 检验不是 UMPT,问题(2)的 UMPT 是不存在的. 本节将证明问题(2)的 χ^2 检验是 UMPUT.

一般来说,多参数指数型分布族的单边和双边假设检验问题,UMPT 是不存在的,但存在 UMPUT. 对这类问题寻找 UMPUT 需要用到指数型分布族的有关性质.

设样本 X 服从多参数指数型分布,其密度函数为

$$p(x; \theta, \mathbf{r}) = c(\theta, \mathbf{r}) \exp \left\{ \theta u(x) + \sum_{i=1}^k r_i t_i(x) \right\} h(x), \quad (3.47)$$

其中 θ 为被检验的参数, $\mathbf{r} = (r_1, r_2, \dots, r_k)$ 为多余参数, $(\theta, \mathbf{r}) \in \Omega \subset \mathbf{R}^{k+1}$. 若参数空间 Ω 有内点,则 $(U(X), T(X)) = (U(X), T_1(X), \dots, T_k(X))$ 是参数 $(\theta, \mathbf{r})$ 的充分、完备的统计量. 据定理 3.1 的充分性原则,我们仅需考虑基于 $(U(X), T(X))$ 的检验. $(U(X), T(X))$ 的分布仍为多参数指数型分布,其密度函数为

$$p(u, t; \theta, \mathbf{r}) = c_1(\theta, \mathbf{r}) \exp \left\{ \theta u + \sum_{i=1}^k r_i t_i \right\} h_1(u, t).$$

在 $T(X) = t$ 的条件下, $U(X)$ 的条件分布与多余参数 $\mathbf{r}$ 无关,且为单参数指数型分布

$$p(u | T = t) = c_1(\theta) \exp \{ \theta u \} h_1(u). \quad (3.48)$$

由此,多参数指数型分布族中的假设检验问题转化为单参数指数型分布族中的

假设检验问题,即在 $\{x; T(x) = t\}$ 的子空间上,寻找 UMPT 或 UMPUT. 这些在 $T(X) = t$ 的条件下最优的检验,是不是在整个样本空间上一致最优的检验,这是需要我们讨论的问题.

$T(X)$ 的边际分布为

$$p(t) = c_{\theta}(\mathbf{r}) \exp \left\{ \sum_{i=1}^k r_i t_i \right\} h_{\theta}(t),$$

在 θ 给定时, $T(X)$ 的边际分布为指数型分布. 令 $G_{\theta} = \{\mathbf{r}; (\theta, \mathbf{r}) \in \Omega\}$. 在 θ 给定时,若空间 G_{θ} 有内点,则 $T(X)$ 是参数 $\mathbf{r}$ 的充分、完备的统计量. 由此我们可以证明,在 $T(X) = t$ 的条件下最优的检验,在整个样本空间上也是一致最优的检验.

§ 3.5.2 多参数指数型分布族的假设检验

考虑单边或双边假设检验问题 (I), ..., (V) (见 § 3.3.1). 设原假设 $H_0: \theta \in \Theta_0$, 备择假设 $H_1: \theta \in \Theta_1$. 令 Γ 是 Θ_0 和 Θ_1 的公共边界, 则 Γ 仅含有一或两个点. 在 $\theta \in \Gamma$ 时, 设 G_{θ} 有内点. 在 $T(X) = t$ 的条件下, 基于条件单参数指数型分布 (3.48), 若 $\phi(u, t)$ 是该检验问题的显著性水平为 α 的 (条件) UMPT, 或 (条件) UMPUT, 则可证明: 在整个样本空间上, $\phi(u, t)$ 是该检验问题的显著性水平为 α 的 UMPUT.

由于在 $T(X) = t$ 的条件下, $\phi(u, t)$ 是检验问题的显著性水平为 α 的条件检验, 所以

$$E_{\theta}(\phi(U, T) \mid T = t) \leq \alpha, \quad \theta \in \Theta_0,$$

从而

$$E_{\theta, \mathbf{r}}(\phi(U, T)) = E_{\theta, \mathbf{r}}(E_{\theta}(\phi(U, T) \mid T)) \leq \alpha, \quad \theta \in \Theta_0.$$

这说明 $\phi(u, t)$ 在整个样本空间上, 也是该检验问题的显著性水平为 α 的一个检验. 设 $\phi^*(u, t)$ 在整个样本空间上, 是检验问题的显著性水平为 α 的任意一个无偏检验. 对于指数族, 任意一个检验的势函数都是连续的. 所以 $E_{\theta, \mathbf{r}}(\phi^*(U, T)) = \alpha$, $\theta \in \Gamma$. 从而在 $\theta \in \Gamma$ 给定时, 对所有的 $\mathbf{r} \in G_{\theta}$, 都有 $E_{\theta, \mathbf{r}}(E_{\theta}(\phi^*(U, T) \mid T)) = \alpha$. 由于在 θ 给定时, $T(X)$ 是完备的统计量, 因此在 $\theta \in \Gamma$ 给定时

$$E_{\theta}(\phi^*(U, T) \mid T = t) = \alpha.$$

这里可能有一个零测集 B , 在 $t \in B$ 时, 上述等式不成立. 但这并不影响后面的证明. 我们不妨假定上述等式总是成立的. 这说明在 $T(X) = t$ 的条件下, 基于条件单参数指数型分布 (3.48), $\phi^*(u, t)$ 是该检验问题的显著性水平为 α 的边界相似检验.

首先考虑单边假设检验问题(I), (II)和双边假设检验问题(IV), (V). 我们知道, 单参数指数型分布族的这些检验问题的 UMPT, 或 UMPUT 与一致最优势边界相似检验是等价的. 由于在 $T(X)=t$ 的条件下, 基于条件单参数指数型分布(3.48), $\phi(u, t)$ 是该检验问题的显著性水平为 α 的条件 UMPT, 或条件 UMPUT. 因此

$$E_{\theta}(\phi(U, T) | T=t) \geq E_{\theta}(\phi^*(U, T) | T=t), \quad \theta \in \Theta_1,$$

从而

$$E_{\theta, r}(E_{\theta}(\phi(U, T) | T)) \geq E_{\theta, r}(E_{\theta}(\phi^*(U, T) | T)), \quad \theta \in \Theta_1,$$

$$E_{\theta, r}(\phi(U, T)) \geq E_{\theta, r}(\phi^*(U, T)), \quad \theta \in \Theta_1.$$

由于 $\phi^*(u, t)$ 在整个样本空间上, 是显著性水平为 α 的任意一个无偏检验, 所以在整个样本空间上 $\phi(u, t)$ 是单边假设检验问题(I), (II)和双边假设检验问题(IV), (V)的显著性水平为 α 的 UMPUT. 至于在整个样本空间上, $\phi(u, t)$ 是双边假设检验问题(III)的显著性水平为 α 的 UMPUT 的证明, 我们将它留作习题 3.12.

由上述结果易得下述定理.

定理 3.15 设样本 X 服从多参数指数型分布(3.47). 则

(1) 原假设 $H_0: \theta \leq \theta_0$ 对备择假设 $H_1: \theta > \theta_0$ 的单边假设检验问题(I)的显著性水平为 α 的 UMPUT 为

$$\phi(u, t) = \begin{cases} 1, & u > c(t), \\ r(t), & u = c(t), \\ 0, & u < c(t), \end{cases} \quad (3.49)$$

其中 $r(t)$ ($0 \leq r(t) \leq 1$) 和 $c(t)$ 由下式确定

$$E_{\theta_0}(\phi(U, T) | T=t) = \alpha. \quad (3.50)$$

如果考虑 $H_0: \theta \geq \theta_0$ 对 $H_1: \theta < \theta_0$ 的单边假设检验问题(II), 类似的结论仍全部成立, 但是式(3.49)中的不等号要改变方向.

(2) 原假设 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对备择假设 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$ 的双边假设检验问题(IV)的显著性水平为 α 的 UMPUT 为

$$\phi(u, t) = \begin{cases} 0, & c_1(t) < u < c_2(t), \\ r_i(t), & u = c_i(t), \quad i = 1, 2, \\ 1, & u < c_1(t) \text{ 或 } u > c_2(t), \end{cases} \quad (3.51)$$

其中 $r_i(t)$ ($0 \leq r_i(t) \leq 1$) 和 $c_i(t)$ 由下式确定

$$E_{\theta_1}(\phi(U, T) \mid T = t) = E_{\theta_2}(\phi(U, T) \mid T = t) = \alpha. \quad (3.52)$$

如果考虑 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 对 $H_1: \theta_1 < \theta < \theta_2$ 的双边假设检验问题(V), 类似的结论仍全部成立, 但是式(3.51)中的“ $c_1(t) < u < c_2(t)$ ”与“ $u < c_1(t)$ 或 $u > c_2(t)$ ”要互相交换.

(3) 原假设 $H_0: \theta = \theta_0$ 对备择假设 $H_1: \theta \neq \theta_0$ 的双边假设检验问题(III)的显著性水平为 α 的 UMPUT 由式(3.51)给出, 其中 $r_i(t)$ ($0 \leq r_i(t) \leq 1$) 和 $c_i(t)$ 由式(3.50)和式(3.53)确定:

$$E_{\theta_0}(U\phi(U, T) \mid T = t) = \alpha E_{\theta_0}(U \mid T = t). \quad (3.53)$$

如果所要检验的参数是 $\theta^* = a\theta + \sum_{i=1}^k b_i r_i$ ($a \neq 0$), 只需将式(3.47)改写成下面的形式:

$$p(x; \theta, \mathbf{r}) = c(\theta, \mathbf{r}) \exp \left\{ \theta^* u^*(x) + \sum_{i=1}^k r_i t_i^*(x) \right\} h(x),$$

其中 $u^*(x) = \frac{u(x)}{a}$, $t_i^*(x) = t_i(x) - \frac{b_i}{a} u(x)$, 利用定理 3.15 就可解决关于 θ^* 的各种单边和双边假设检验问题.

§ 3.5.3 两个 Poisson 总体的比较

设相互独立的两个总体 x 和 y 的分布分别是参数为 λ 和 μ 的 Poisson 分布. 要比较这两个总体的差异, 就是要比较参数 λ 和 μ 的差异. 为了应用定理 3.15, 首先把它们的联合分布表示成指数族分布(3.47)的形式

$$P(x, y; \lambda, \mu) = \frac{e^{-(\lambda+\mu)} \exp\{\theta u + r t\}}{x! y!},$$

其中 $\theta = \ln(\mu/\lambda)$, $r = \ln(\lambda)$, $u = y$, $t = x + y$. 比较参数 λ 和 μ 的检验问题就变为关于参数 θ 的检验问题. 例如原假设 $H_0: \lambda \leq \mu$ 对备择假设 $H_1: \lambda > \mu$ 的检验问题等价于原假设 $H_0: \theta \geq 0$ 对备择假设 $H_1: \theta < 0$ 的检验问题. 按定理 3.15, 只要在直线 $x + y = t$ 的整数点上寻找最优势的条件检验. 在给定 $X + Y = t$ 后, Y 的条件分布为

$$P(y \mid x + y = t) = \binom{t}{y} p^y (1-p)^{t-y}, \quad y = 0, 1, \dots, t,$$

这是二项分布 $b(t, p)$, 其中 $p = \frac{\mu}{\lambda + \mu} = \frac{e^\theta}{1 + e^\theta}$. 所以原来的假设检验问题就化为关

于二项分布的参数 p 的假设检验问题. 例如原假设 $H_0: \lambda \leq \mu$ 对备择假设 $H_1: \lambda > \mu$ 的检验问题等价于原假设 $H_0: p \geq 0.5$ 对备择假设 $H_1: p < 0.5$ 的假设检验问题.

上述方法也可用来检验两个样本 $X_1, X_2, \dots, X_m$ 和 $Y_1, Y_2, \dots, Y_n$ 是否来自同一个 Poisson 总体. 这是因为 $\sum_{i=1}^m X_i$ 和 $\sum_{i=1}^n Y_i$ 分别是参数为 $m\lambda$ 和 $n\mu$ 的 Poisson 分布. 检验这两个样本是否来自同一个 Poisson 总体, 等价于检验 $m\lambda$ 和 $n\mu$ 的比值是否等于 m/n .

§ 3.5.4 两个二项总体的比较

设相互独立的两个总体 x 和 y 皆服从二项分布. 它们分别是 $b(m, p_1)$ 和 $b(n, p_2)$. 它们的联合分布可表示成指数族分布 (3.47) 的形式

$$P(x, y; p_1, p_2) = (1 - p_1)^m (1 - p_2)^n \exp \{ \theta u + rt \} \binom{m}{x} \binom{n}{y},$$

其中 $\theta = \ln \frac{p_2(1-p_1)}{p_1(1-p_2)}$, $r = \ln \frac{p_1}{1-p_1}$, $u = y$, $t = x + y$. 那么比较这两个总体的差异, 即比较参数 p_1 和 p_2 的差异的检验问题就变为关于参数 θ 的检验问题. 例如原假设 $H_0: p_1 = p_2$ 对备择假设 $H_1: p_1 \neq p_2$ 的检验问题等价于原假设 $H_0: \theta = 0$ 对备择假设 $H_1: \theta \neq 0$ 的检验问题. 按定理 3.15, 只要在直线 $x + y = t$ 的整数点上寻找最优势的条件检验. 在给定 $x + y = t$ 后, y 的条件分布为

$$P(y | x + y = t) = c_t(\theta) \binom{m}{t-y} \binom{n}{y} e^{\theta y}, \quad y = 0, 1, \dots, t,$$

其中 $c_t(\theta) = \left[\sum_{y=0}^t \binom{m}{t-y} \binom{n}{y} e^{\theta y} \right]^{-1}$. 在 $\theta = 0$ 时, 这个条件分布是超几何分布

$$P(y | x + y = t) = \frac{\binom{m}{t-y} \binom{n}{y}}{\binom{m+n}{t}}, \quad y = 0, 1, \dots, t,$$

用它确定原假设 $H_0: \theta = 0$ 对备择假设 $H_1: \theta \neq 0$ 的检验问题的临界值.

§ 3.5.5 正态总体参数的检验问题

为了能用更简单的形式给出正态总体参数的检验问题的检验函数, 首先证明下述定理.

定理 3.16 设样本 X 服从多参数指数型分布 (3.47).

(1) 如果存在一个统计量 $V=V(U, T)$, 给定 $T=t$ 后, V 是 U 的单调增函数, 且当 $\theta=\theta_0$ 时, V 与 T 互相独立, 则 $H_0: \theta \leq \theta_0$ 对 $H_1: \theta > \theta_0$ 的单边假设检验问题 (I) 的显著性水平为 α 的 UMPUT 为

$$\phi(v) = \begin{cases} 1, & v > c, \\ r, & v = c, \\ 0, & v < c, \end{cases} \quad (3.54)$$

其中 $r(0 \leq r \leq 1)$ 和 c 不依赖于 t , 它们由下式确定:

$$E_{\theta_0}(\phi(V)) = \alpha, \quad (3.55)$$

如果考虑 $H_0: \theta \geq \theta_0$ 对 $H_1: \theta < \theta_0$ 的单边假设检验问题 (II), 类似的结论仍全部成立, 但是式 (3.54) 中的不等号要改变方向.

(2) 如果存在一个统计量 $V=V(U, T)$, 给定 $T=t$ 后, V 是 U 的单调函数, 且当 $\theta=\theta_1$ 或 θ_2 时, V 与 T 互相独立, 则 $H_0: \theta_1 \leq \theta \leq \theta_2$ 对 $H_1: \theta < \theta_1$ 或 $\theta > \theta_2$ 的双边假设检验问题 (IV) 的显著性水平为 α 的 UMPUT 为

$$\phi(v) = \begin{cases} 0, & c_1 < v < c_2, \\ r_i, & v = c_i, \quad i = 1, 2, \\ 1, & v < c_1 \text{ 或 } v > c_2, \end{cases} \quad (3.56)$$

其中 $r_i(0 \leq r_i \leq 1)$ 和 c_i 不依赖于 t , 它们由下式确定:

$$E_{\theta_1}(\phi(V)) = E_{\theta_2}(\phi(V)) = \alpha. \quad (3.57)$$

如果考虑 $H_0: \theta \leq \theta_1$ 或 $\theta \geq \theta_2$ 对 $H_1: \theta_1 < \theta < \theta_2$ 的双边假设检验问题 (V), 类似的结论仍全部成立, 但是式 (3.56) 中的“ $c_1 < v < c_2$ ”与“ $v < c_1$ 或 $v > c_2$ ”要互相交换.

(3) 如果存在一个统计量 $V=a(T)U+b(T)$ ($a(T) \neq 0$), 即给定 $T=t$ 后, V 是 U 的线性函数, 且当 $\theta=\theta_0$ 时, V 与 T 互相独立, 则 $H_0: \theta=\theta_0$ 对 $H_1: \theta \neq \theta_0$ 的双边假设检验问题 (III) 的显著性水平为 α 的 UMPUT 由式 (3.56) 给出, 其中的 r_i ($0 \leq r_i \leq 1$) 和 c_i 不依赖于 t . 它们由式 (3.55) 和以下的式子确定:

$$E_{\theta_0}(V\phi(V)) = \alpha E_{\theta_0}(V). \quad (3.58)$$

证明: (1) 据定理 3.15(1), 单边假设检验问题 (I) 的显著性水平为 α 的 UMPUT 由式 (3.49) 和式 (3.50) 给出. 因为对给定的 $t, v=v(u, t)$ 是 u 的单调增函数, 所以检验函数等价地由下式给出:

$$\phi(v, t) = \begin{cases} 1, & v > c(t), \\ r(t), & v = c(t), \\ 0, & v < c(t), \end{cases} \quad (3.59)$$

其中 $r(t)$ ($0 \leq r(t) \leq 1$) 和 $c(t)$ 由 $E_{\theta_0}(\phi(V, T) \mid T=t) = \alpha$ 确定. 又因为在 $\theta = \theta_0$ 时, V 与 T 互相独立, 所以 $r(t)$ 和 $c(t)$ 不依赖于 t , 而由式 (3.55) 给出. 因此, 由式 (3.59) 给出的检验就是由式 (3.54) 给出的检验. 关于单边假设检验问题 (II) 有类似的结论. (1) 得证.

(2) 据定理 3.15(2), 双边假设检验问题 (IV) 的显著性水平为 α 的 UMPUT 由式 (3.51) 和式 (3.52) 给出. 因为对给定的 $t, v = v(u, t)$ 是 u 的单调函数, 又因为在 $\theta = \theta_1$ 或 θ_2 时, V 与 T 互相独立, 所以检验函数等价地由式 (3.56) 给出, 且 r_i ($0 \leq r_i \leq 1$) 和 c_i 不依赖于 t , 而由式 (3.57) 给出. 关于双边假设检验问题 (V) 有类似的结论. (2) 得证.

(3) 据定理 3.15(3), 双边假设检验问题 (III) 的显著性水平为 α 的 UMPUT 由式 (3.51)、式 (3.50) 和式 (3.53) 给出. 因为对给定的 $t, v = a(t)u + b(t)$, v 是 u 的单调函数, 所以检验函数等价地由下式给出:

$$\phi(v, t) = \begin{cases} 0, & c_1(t) < v < c_2(t), \\ r_i(t), & v = c_i(t), \quad i = 1, 2, \\ 1, & v < c_1(t) \text{ 或 } v > c_2(t), \end{cases} \quad (3.60)$$

其中 $r_i(t)$ ($0 \leq r_i(t) \leq 1$) 和 $c_i(t)$ 由以下两式确定:

$$E_{\theta_0}(\phi(V, T) \mid T = t) = \alpha,$$

$$E_{\theta_0}\left(\frac{V - b(T)}{a(T)} \phi(V, T) \mid T = t\right) = \alpha E_{\theta_0}\left(\frac{V - b(T)}{a(T)} \mid T = t\right).$$

由此可以推得, 它们由以下两式确定:

$$E_{\theta_0}(\phi(V, T) \mid T = t) = \alpha,$$

$$E_{\theta_0}(V \phi(V, T) \mid T = t) = \alpha E_{\theta_0}(V \mid T = t).$$

因为在 $\theta = \theta_0$ 时, V 与 T 互相独立, 所以 $r_i(t)$ ($0 \leq r_i(t) \leq 1$) 和 $c_i(t)$ 不依赖于 t , 而由式 (3.55) 和式 (3.58) 确定, 则由式 (3.60) 给出的检验就是由式 (3.56) 给出的检验. (3) 得证. 定理证毕.

为了能用更简单的形式给出正态总体参数的检验问题的检验函数, 尚需下

述引理.

引理 3.17 设 $(\mathcal{X}, \mathcal{B}, \mathcal{P} = \{P_\theta : \theta \in \Theta\})$ 为一参数结构, $T(X)$ 为 θ 的充分且完备的统计量. 如果可测统计量 $S(X)$ 的分布与 θ 无关, 则对所有的 $\theta \in \Theta$, $T(X)$ 与 $S(X)$ 相互独立.

证明见参考文献[2].

例 3.10 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态总体 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty, \sigma^2 > 0$. 考虑关于 μ 的假设检验问题, (1) $H_0: \mu \geq 0$ 对 $H_1: \mu < 0$; (2) $H_0: \mu \leq 0$ 对 $H_1: \mu > 0$; (3) $H_0: \mu = 0$ 对 $H_1: \mu \neq 0$. 若考虑的假设检验问题是 $H_0: \mu \geq \mu_0$ 对 $H_1: \mu < \mu_0$ 等, 则可以通过变换 $Y_i = X_i - \mu_0$, 归结为上述情况之一.

样本的联合密度为

$$p(x; \mu, \sigma^2) = (\sqrt{2\pi}\sigma)^{-n} \exp\left\{-\frac{n\mu^2}{2\sigma^2}\right\} \exp\{\theta u + rt\},$$

其中 $\theta = n\mu/\sigma^2, r = (-2\sigma^2)^{-1}, u = \bar{x}, t = \sum_{i=1}^n x_i^2$. 这时, 关于 μ 的检验问题就变为关于 θ 的检验问题. 例如, 检验问题(1)等价于 $H_0: \theta \geq 0$ 对 $H_1: \theta < 0$ 的检验问题.

为了能用定理 3.16(1), 以较为简单的形式给出问题(1)和(2)的检验函数, 首先寻找统计量 $V = V(U, T)$, 使得对给定的 T, V 是 U 的单增函数, 且在 $\theta = 0$, 即在 $\mu = 0$ 时, V 与 T 互相独立. 为此令

$$V = \sqrt{n(n-1)} \frac{U}{\sqrt{T - nU^2}} = \sqrt{n(n-1)} \frac{\bar{X}}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}},$$

下面证明它满足上述要求. 对给定的 T, V 是 U 的单增函数. 在 $\mu = 0$ 时, 样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态总体 $N(0, \sigma^2)$. 这时, $T = \sum_{i=1}^n X_i^2$ 是 σ^2 的充分、完备的统计量. 在 $\mu = 0$ 时, 由于

$$V = \sqrt{n(n-1)} \frac{\bar{X}}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}} = \sqrt{n(n-1)} \frac{\bar{Z}}{\sqrt{\sum_{i=1}^n (Z_i - \bar{Z})^2}},$$

其中 $Z_i = X_i/\sigma \sim N(0, 1)$, 所以 V 的分布与参数 σ^2 无关. 因为在 $\mu = 0$ 时, T 是 σ^2 的充分、完备的统计量, 而 V 的分布又与参数 σ^2 无关, 则据引理 3.17 知, V 与 T 互相独立. 在 $\mu = 0$ 时, $V \sim t(n-1)$. 根据定理 3.16(1), 前两个检验问题的 UMPUT 的拒绝域分别为 $\{x: v \leq t_\alpha(n-1)\}$ 和 $\{x: v \geq t_{1-\alpha}(n-1)\}$. 这就是常用的 t 检验法.

为了能用定理 3.16(3), 以较为简单的形式给出问题(3)的检验函数, 首先寻找统计量 W . 它形如 $W = a(T)U + b(T)$ ($a(T) \neq 0$), 且在 $\theta = 0$ 时, W 与 T 互相独立. 为此令

$$W = \frac{U}{\sqrt{T}} = \frac{\bar{X}}{\sqrt{\sum_{i=1}^n X_i^2}},$$

则对固定的 T , W 是 U 的线性函数; 且在 $\mu = 0$ 时, 由于 $T = \sum_{i=1}^n X_i^2$ 是 σ^2 的充分、完备的统计量, 且由于 W 的分布与参数 σ^2 无关(习题 3.13), 所以据引理 3.17 知, W 与 T 互相独立. 根据定理 3.16(3), 并由于在 $\mu = 0$ 时, W 的分布关于原点对称(习题 3.13), 从而根据定理 3.14 的说明, 最后一个检验问题的 UMPUT 的拒绝域为 $\{x: |w| \geq c\}$. 由于

$$V = \sqrt{n(n-1)} \frac{W}{\sqrt{1 - nW^2}},$$

故 $|V|$ 是 $|W|$ 的单增函数. 从而拒绝域可等价地表示为 $\{x: |v| \geq c\}$. 由于在 $\mu = 0$ 时, $V \sim t(n-1)$. 从而该检验问题的 UMPUT 的拒绝域为 $\{x: |v| \geq t_{1-\alpha/2}(n-1)\}$. 这也是常用的 t 检验法.

例 3.11 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态总体 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty$, $\sigma^2 > 0$. 考虑关于 σ^2 的假设检验问题 (1) $H_0: \sigma^2 \geq \sigma_0^2$ 对 $H_1: \sigma^2 < \sigma_0^2$; (2) $H_0: \sigma^2 \leq \sigma_0^2$ 对 $H_1: \sigma^2 > \sigma_0^2$; (3) $H_0: \sigma^2 = \sigma_0^2$ 对 $H_1: \sigma^2 \neq \sigma_0^2$. 样本的联合密度为

$$p(x; \mu, \sigma^2) = (\sqrt{2\pi}\sigma)^{-n} \exp\left\{-\frac{n\mu^2}{2\sigma^2}\right\} \exp\{\theta u + rt\},$$

其中 $\theta = (-2\sigma^2)^{-1}$, $r = n\mu/\sigma^2$, $u = \sum_{i=1}^n x_i^2$, $t = \bar{x}$. 这时, 关于 σ^2 的检验问题就变为关于 θ 的检验问题, 例如检验问题(1)就变为 $H_0: \theta \geq (-2\sigma_0^2)^{-1}$ 对 $H_1: \theta < (-2\sigma_0^2)^{-1}$.

令 $V = U - nT^2 = \sum_{i=1}^n (X_i - \bar{X})^2$, 则对固定的 $T = \bar{X}$, V 是 U 的线性增函数. 显然, V 与 T 相互独立. 在 $\theta = (-2\sigma_0^2)^{-1}$, 即 $\sigma^2 = \sigma_0^2$ 时, $V \sim \sigma_0^2 \cdot \chi^2(n-1)$. 根据定理 3.16(1), 前两个检验问题的 UMPUT 的拒绝域分别为 $\{x: v \leq \sigma_0^2 \cdot \chi_{\alpha}^2(n-1)\}$ 和 $\{x: v \geq \sigma_0^2 \cdot \chi_{1-\alpha}^2(n-1)\}$. 这就是常用的 χ^2 检验法. 最后一个检验问题的 UMPUT 的拒绝域为 $\{x: v \leq c_1\}$ 或 $\{x: v \geq c_2\}$, 其中 c_1 与 c_2 由式(3.44)与式(3.45), 或式(3.44)与式(3.46)确定, 但需把其中的自由度 n 换为 $n-1$. 一般常取 $c_1 =$

$\chi^2_{\alpha/2}(n-1), c_2 = \chi^2_{1-\alpha/2}(n-1)$. 这也是常用的 χ^2 检验法.

例 3.12 设样本 $X = (X_1, X_2, \dots, X_m)$ 和样本 $Y = (Y_1, Y_2, \dots, Y_n)$ 互相独立. 它们分别来自正态总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$, 其中: $-\infty < \mu_1 < \infty, -\infty < \mu_2 < \infty, \sigma^2 > 0$. 由于这两个正态总体的方差相等, 所以比较这两个正态总体的差异的检验问题就是比较这两个正态总体的均值的检验问题, 例如 $H_0: \mu_1 = \mu_2$ 对 $H_1: \mu_1 \neq \mu_2$.

样本的联合密度为

$$p(x, y; \mu_1, \mu_2, \sigma^2) = (2\pi\sigma^2)^{-\frac{m+n}{2}} \exp\left\{-\frac{m\mu_1^2 + n\mu_2^2}{2\sigma^2}\right\} \exp\{\theta u + r_1 t_1 + r_2 t_2\},$$

其中

$$\theta = \frac{mn(\mu_2 - \mu_1)}{(m+n)\sigma^2}, \quad r_1 = \frac{m\mu_1 + n\mu_2}{(m+n)\sigma^2}, \quad r_2 = -\frac{1}{2\sigma^2},$$

$$u = \bar{y} - \bar{x}, \quad t_1 = m\bar{x} + n\bar{y}, \quad t_2 = \sum_{i=1}^m x_i^2 + \sum_{i=1}^n y_i^2.$$

这时, 所考虑的检验问题 $H_0: \mu_1 = \mu_2$ 对 $H_1: \mu_1 \neq \mu_2$ 就变为关于 θ 的检验问题, $H_0: \theta = 0$ 对 $H_1: \theta \neq 0$.

根据定理 3.16(3), 首先寻找统计量 V , 使得对固定的 T_1 与 T_2 , V 是 U 的线性函数, 且在 $\theta = 0$ 时, V 与 (T_1, T_2) 互相独立. 为此令

$$V = \frac{U}{\sqrt{T_2 - \frac{T_1^2}{m+n}}},$$

则对固定的 T_1 与 T_2 , V 是 U 的线性函数. 在 $\theta = 0$, 即在 $\mu_1 = \mu_2 = \mu$ 时, 合样本 $(X_1, \dots, X_m, Y_1, \dots, Y_n)$ 不仅互相独立, 而且同分布. 这时, σ^2 和 μ 的充分、完备的统计量为 T_1 和 T_2 . 由于在 $\theta = 0$ 时, V 的分布与参数 σ^2 和 μ 无关 (习题 3.14), 所以据引理 3.17 知, V 与 (T_1, T_2) 互相独立. 根据定理 3.16(3), 并由于在 $\theta = 0$ 时, V 的分布关于原点对称 (习题 3.14), 从而根据定理 3.14 的说明, 该检验问题的 UMPUT 的拒绝域为 $\{x, y: |v| \geq c\}$. 进一步令

$$\begin{aligned} W &= \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{V}{\sqrt{1 - \frac{mn}{m+n} V^2}} \\ &= \sqrt{\frac{mn(m+n-2)}{m+n}} \frac{\bar{Y} - \bar{X}}{\sqrt{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2}}, \end{aligned}$$

则 $|W|$ 是 $|V|$ 的单增函数, 并且在 $\theta=0$ 时, $W \sim t(m+n-2)$. 所以该检验问题的 UMPUT 的拒绝域为 $\{x, y: |w| \geq t_{1-\alpha/2}(m+n-2)\}$. 这就是常用的 t 检验法. 类似地, 检验问题 (1) $H_0: \mu_1 \leq \mu_2$ 对 $H_1: \mu_1 > \mu_2$ 和 (2) $H_0: \mu_1 \geq \mu_2$ 对 $H_1: \mu_1 < \mu_2$ 的 UMPUT 的拒绝域分别为 $\{x, y: w \leq t_{\alpha}(m+n-2)\}$ 和 $\{x, y: w \geq t_{1-\alpha}(m+n-2)\}$. 它们都是常用的 t 检验法.

例 3.13 设样本 $X = (X_1, X_2, \dots, X_m)$ 和样本 $Y = (Y_1, Y_2, \dots, Y_n)$ 互相独立. 它们分别来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$, 其中 $-\infty < \mu_1 < \infty, -\infty < \mu_2 < \infty, \sigma_1^2 > 0, \sigma_2^2 > 0$. 考虑比较这两个正态总体的方差的检验问题, $H_0: \sigma_1^2 = \sigma_2^2$ 对 $H_1: \sigma_1^2 \neq \sigma_2^2$.

样本的联合密度为

$$p(x, y; \mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = c(\mu_1, \mu_2, \sigma_1, \sigma_2) \exp\{\theta u + r_1 t_1 + r_2 t_2 + r_3 t_3\},$$

其中

$$\theta = \frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_2^2}, \quad r_1 = -\frac{1}{2\sigma_1^2}, \quad r_2 = \frac{m\mu_1}{\sigma_1^2}, \quad r_3 = \frac{n\mu_2}{\sigma_2^2},$$

$$u = \sum_{i=1}^n y_i^2, \quad t_1 = \sum_{i=1}^m x_i^2 + \sum_{i=1}^n y_i^2, \quad t_2 = \bar{x}, \quad t_3 = \bar{y}.$$

这时, 关于 σ^2 的检验问题 $H_0: \sigma_1^2 = \sigma_2^2$ 对 $H_1: \sigma_1^2 \neq \sigma_2^2$, 就变为关于 θ 的检验问题 $H_0: \theta=0$ 对 $H_1: \theta \neq 0$.

根据定理 3.16(3), 首先寻找统计量 V , 使得对固定的 T_1, T_2 与 T_3 , V 是 U 的线性函数, 且在 $\theta=0$ 时, V 与 (T_1, T_2, T_3) 互相独立. 为此令

$$V = \frac{U - nT_3^2}{T_1 - mT_2^2 - nT_3^2} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2},$$

则对给定的 T_1, T_2 和 T_3 , V 是 U 的线性增函数. 在 $\theta=0$, 即在 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时, μ_1, μ_2 和 σ^2 的充分、完备的统计量为

$$T_2 = \bar{X}, \quad T_3 = \bar{Y} \text{ 和 } T_1 - mT_2^2 - nT_3^2 = \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

由于在 $\theta=0$ 时, V 的分布与参数 μ_1, μ_2 和 σ^2 无关(习题 3.15), 所以据引理 3.17 知, V 与 $(T_2, T_3, T_1 - mT_2^2 - nT_3^2)$ 互相独立, 从而与 (T_1, T_2, T_3) 互相独立. 由于在 $\theta=0$ 时, $V \sim Be\left(\frac{n-1}{2}, \frac{m-1}{2}\right)$, 根据定理 3.16(3), 检验问题 $H_0: \sigma_1^2 = \sigma_2^2$ 对 $H_1: \sigma_1^2 \neq$

σ_2^2 的 UMPUT 的拒绝域为 $\{x, y: v \leq c_1 \text{ 或 } v \geq c_2\}$, 其中 c_1 和 c_2 由以下两式确定 (习题 3.15):

$$\int_{c_1}^{c_2} Be\left(x \mid \frac{n-1}{2}, \frac{m-1}{2}\right) dx = 1 - \alpha,$$

$$\int_{c_1}^{c_2} \left(x - \frac{n-1}{m+n-2}\right) Be\left(x \mid \frac{n-1}{2}, \frac{m-1}{2}\right) dx = 0,$$

这里 $Be\left(x \mid \frac{n-1}{2}, \frac{m-1}{2}\right)$ 表示 $Be\left(\frac{n-1}{2}, \frac{m-1}{2}\right)$ 分布的密度函数. 令

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

则在 $\theta=0$ 时, $F \sim F(n-1, m-1)$. 由于 F 是 V 的严增函数, 所以该检验问题的 UMPUT 的拒绝域可等价地写为 $\{x, y: F \leq c_1 \text{ 或 } F \geq c_2\}$, 其中 c_1 和 c_2 由以下两式确定 (习题 3.15):

$$\int_{c_1}^{c_2} F(x \mid n-1, m-1) dx = 1 - \alpha,$$

$$\int_{c_1}^{c_2} \frac{x-1}{(n-1)x + (m-1)} F(x \mid n-1, m-1) dx = 0,$$

这里 $F(x \mid n-1, m-1)$ 表示 $F(n-1, m-1)$ 分布的密度函数.

拒绝域为 $\{x, y: F \leq c_1 \text{ 或 } F \geq c_2\}$ 的检验的势函数与参数 μ_1 和 μ_2 无关, 仅依赖于比值 $\rho = \frac{\sigma_1^2}{\sigma_2^2}$. 它的势函数为

$$g(\rho) = 1 - \int_{\rho c_1}^{\rho c_2} F(x \mid n-1, m-1) dx,$$

与例 3.9 相类似地, 由势函数在 $\rho=1$ 时达到最小值推得, c_1 与 c_2 也可由以下两式确定:

$$\int_{c_1}^{c_2} F(x \mid n-1, m-1) dx = 1 - \alpha,$$

$$c_2 F(c_2 \mid n-1, m-1) = c_1 F(c_1 \mid n-1, m-1).$$

一般常取 $c_1 = F_{\alpha/2}(n-1, m-1)$, $c_2 = F_{1-\alpha/2}(n-1, m-1)$. 这是常用的 F 检验法.

由于 F 可改写为

$$F = \frac{U - nT_3^2}{T_1 - mT_2^2 - U} \frac{m-1}{n-1},$$

则对给定的 T_1, T_2 和 T_3 , F 是 U 的严增函数. 前已证明, 在 $\theta=0$ 时, V 与 (T_1, T_2, T_3) 互相独立, 而 F 仅与 V 有关, 所以 F 与 (T_1, T_2, T_3) 也互相独立. 根据定理 3.16(1), 检验问题 $H_0: \sigma_1^2 \leq \sigma_2^2$ 对 $H_1: \sigma_1^2 > \sigma_2^2$, 以及 $H_0: \sigma_1^2 \geq \sigma_2^2$ 对 $H_1: \sigma_1^2 < \sigma_2^2$ 的 UMPUT 的拒绝域分别为 $\{x, y: F \leq F_\alpha(n-1, m-1)\}$ 和 $\{x, y: F \geq F_{1-\alpha}(n-1, m-1)\}$. 这些都是常用的 F 检验法.

§ 3.6 似然比检验

§ 3.6.1 似然比检验

前面讨论了一致最优势检验和一致最优势无偏检验. 但是在许多情况下, 对于一个复合假设的检验问题, 上述这些最优势检验可能不存在. 这时, 我们当然可以采用类似于前面的处理方法, 对检验提出进一步的限制, 然后在这被限制的较小的检验类中寻找最优检验, 这样做, 在实际使用中往往是困难的, 也是不可取的. 另一个途径是给出一个构造检验法的一般方法. N-P 基本引理告诉我们, 在简单原假设对简单备择假设的检验问题中, 最优势检验由似然比检验给出. 事实上, 似然比检验方法也可用在复合假设的检验问题中构造检验. 这是构造检验的常用方法.

设 $X = (X_1, X_2, \dots, X_n)$ 的密度函数是 $p(x; \theta)$, 其中未知参数 $\theta \in \Theta$ 可以是向量. 我们知道, 简单原假设 $H_0: \theta = \theta_0$ 对简单备择假设 $H_1: \theta = \theta_1 (\theta_1 \neq \theta_0)$ 的检验问题的似然比为

$$\lambda(x) = \frac{p(x; \theta_1)}{p(x; \theta_0)}.$$

现考虑复合假设的检验问题: 原假设 $H_0: \theta \in \Theta_0$ 对备择假设 $H_1: \theta \in \Theta_1$. 在复合假设的检验问题中, 很自然地定义似然比 $\lambda(x)$ 为

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_1} p(x; \theta)}{\sup_{\theta \in \Theta_0} p(x; \theta)} = \frac{p(x; \hat{\theta}_1)}{p(x; \hat{\theta}_0)},$$

其中 $\hat{\theta}_0$ 和 $\hat{\theta}_1$ 分别是 H_0 和 H_1 成立时 θ 的 MLE. $p(x; \hat{\theta}_0)$ 是原假设 H_0 成立时, 观察到样本点 x 的可能性的一个度量, 而 $p(x; \hat{\theta}_1)$ 是备择假设 H_1 成立时, 观察到

样本点 x 的可能性的一个度量. 在 $\lambda(x)$ 比较大时, 备择假设成立观察到样本点 x 的可能性比较大, 因此很自然地, 在 $\lambda(x)$ 比较大时拒绝原假设. 故取检验的拒绝域为 $\{x: \lambda(x) \geq c\}$. 这个检验方法常用于区分样本来自这类分布, 还是那类分布的检验问题.

例 3.14 设有样本 $X = (X_1, X_2, \dots, X_n)$. 考虑检验问题, 原假设和备择假设分别是 H_0 : 样本来自正态分布 $N(\mu, \sigma^2)$, 其密度函数为

$$p_0(x; \mu, \sigma^2) = (\sqrt{2\pi}\sigma)^{-1} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\},$$

和 H_1 : 样本来自双参数指数型分布族 $Exp(\mu, \sigma)$, 其密度函数为

$$p_1(x; \mu, \sigma) = \frac{1}{\sigma} \exp\left\{-\frac{x-\mu}{\sigma}\right\}, x - \mu \geq 0 \text{ 或 } x - \mu < 0,$$

其中 $-\infty < \mu < \infty, \sigma > 0$. 原假设成立时, μ 和 σ 的 MLE 分别为

$$\hat{\mu}_0 = \bar{X}, \quad \hat{\sigma}_0 = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2},$$

而备择假设成立时, μ 和 σ 的 MLE 分别为

$$\hat{\mu}_1 = X_{(1)} = \min(X_1, X_2, \dots, X_n), \quad \hat{\sigma}_1 = \bar{X} - X_{(1)},$$

所以该检验问题的似然比

$$\lambda(x) = \frac{\prod_{i=1}^n p_1(x_i; \hat{\mu}_1, \hat{\sigma}_1)}{\prod_{i=1}^n p_0(x_i; \hat{\mu}_0, \hat{\sigma}_0^2)} = (\sqrt{2\pi}e^{-1})^n D^n,$$

其中

$$D = \frac{\sqrt{n \sum_{i=1}^n (x_i - \bar{x})^2}}{\sum_{i=1}^n (x_i - x_{(1)})}.$$

由于 $\lambda(x)$ 关于 D 严增, 所以拒绝域可取为 $\{x: D \geq c\}$. 由于

$$D = \frac{\sqrt{n \sum_{i=1}^n (x_i - \bar{x})^2}}{\sum_{i=1}^n (x_i - x_{(1)})} = \frac{\sqrt{n \sum_{i=1}^n (u_i - \bar{u})^2}}{\sum_{i=1}^n (u_i - u_{(1)})},$$

其中 $u_i = \frac{x_i - \mu}{\sigma}$, $i = 1, 2, \dots, n$, 所以不论原假设为真, 还是备择假设为真, D 的分布

皆与 μ 和 σ 无关. 原假设为真时, 检验犯第一类错误的概率为

$$P\{D \geq c \mid \text{样本来自标准正态分布 } N(0, 1)\}.$$

利用随机模拟法, 对不同的样本容量 n , 求得原假设为真时 D 的分位数值, 从而得到了检验的临界值 c . 由于备择假设为真时, 检验犯第二类错误的概率为

$$P\{D < c \mid \text{样本来自标准指数分布 } \text{Exp}(0, 1)\}.$$

利用随机模拟法, 可以求得检验犯第二类错误的概率.

简单原假设对简单备择假设的似然比 $\lambda(x)$ 的另一个很自然的推广为

$$\lambda(x) = \frac{\sup_{\theta \in \Theta} p(x; \theta)}{\sup_{\theta \in \Theta_0} p(x; \theta)} = \frac{p(x; \hat{\theta})}{p(x; \hat{\theta}_0)},$$

其中, $\hat{\theta}$ 是 $\theta \in \Theta$, 即 θ 没有受到限制时的 MLE; $\hat{\theta}_0$ 是原假设 $H_0: \theta \in \Theta_0$ 成立时, 即 θ 受到限制时的 MLE. 显然, $\lambda(x) \geq 1$. 检验的拒绝域为 $\{x: \lambda(x) \geq c\}$. 似然比的这种推广较前者普遍. 它常用于参数假设检验问题.

例 3.15 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态分布 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty$, $\sigma^2 > 0$. 考虑检验问题, 原假设 $H_0: \mu = 0$ 对备择假设 $H_1: \mu \neq 0$. 原假设 $H_0: \mu = 0$

成立时, σ^2 的 MLE 为 $\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$; 而备择假设 $H_1: \mu \neq 0$ 成立时, 即均值 μ 和方

差 σ^2 皆未知时, μ 和 σ^2 的 MLE 分别为 $\hat{\mu} = \bar{X}$ 和 $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$. 因而似然比

$$\begin{aligned} \lambda(x) &= \frac{\prod_{i=1}^n (\sqrt{2\pi} \hat{\sigma})^{-1} \exp\left\{-\frac{(x_i - \hat{\mu})^2}{2\hat{\sigma}^2}\right\}}{\prod_{i=1}^n (\sqrt{2\pi} \hat{\sigma}_0)^{-1} \exp\left\{-\frac{x_i^2}{2\hat{\sigma}_0^2}\right\}} \\ &= \left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2 + n \bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right]^{\frac{n}{2}} = \left(1 + \frac{t^2}{n-1} \right)^{\frac{n}{2}}, \end{aligned}$$

其中

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

显然, $\lambda(x)$ 是 $|t|$ 的严增函数. 故拒绝域可取为 $\{x: |t| \geq c\}$. 原假设 $H_0: \mu=0$ 成立时, $t \sim t(n-1)$. 故取 $c = t_{1-\alpha/2}(n-1)$. 这就是常用的 t 检验法. 不仅 t 检验如此, 前面几节给出的不少最优检验都是似然比检验. 由于似然比检验的适用面广, 且由它构造出来的检验常具有某种优良性, 所以似然比检验是假设检验中的一个重要方法.

一般来说, 难以求得似然比统计量的精确分布. 因而有必要寻求在样本容量趋于无穷时, 似然比统计量的极限分布. 1938 年, Wilks (威尔克斯) 证明了关于似然比统计量的极限分布的重要定理. 下面分简单原假设和复合原假设两种情况进行讨论.

§ 3.6.2 简单原假设的检验问题

定理 3.18 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自密度函数族 $\{p(x; \theta) : \theta \in \Theta\}$, 其中未知参数 $\theta = (\theta_1, \theta_2, \dots, \theta_k)$, 参数空间 Θ 是 k 维欧氏空间 $\mathbf{R}^k$ 中的一个含有内点的集合. 假设参数真值 θ_0 是 Θ 的一个内点.

假设该密度函数族满足下列四个条件:

$$(1) \int_x \frac{\partial p(x; \theta)}{\partial \theta_i} d\mu(x) = \int_x \frac{\partial^2 p(x; \theta)}{\partial \theta_i \partial \theta_j} d\mu(x) = 0, \quad i, j = 1, 2, \dots, k;$$

$$(2) I(\theta) = (I_{ij}(\theta))_{k \times k} > 0, \quad \forall \theta \in \Theta, \text{ 其中}$$

$$I_{ij}(\theta) = \int_x \frac{\partial \ln p(x; \theta)}{\partial \theta_i} \frac{\partial \ln p(x; \theta)}{\partial \theta_j} p(x; \theta) d\mu(x);$$

(3) 存在 $M(x)$, 使得 $\int_x M(x) p(x; \theta) d\mu(x) < K, \quad \forall \theta \in \Theta$, 其中 K 与 θ 无关, 且在含有参数真值 θ_0 的一个邻域内,

$$\left| \frac{\partial^3 \ln p(x; \theta)}{\partial \theta_i \partial \theta_j \partial \theta_l} \right| \leq M(x), \quad i, j, l = 1, 2, \dots, k;$$

(4) 不同的 θ 值, 对应着不同的概率分布.

除上述四个条件外, 再假设 θ 的 MLE $\hat{\theta}$ 是似然方程 $\frac{\partial l(\theta)}{\partial \theta} = 0$ 的解, 其中

$$l(\theta) = \ln \prod_{i=1}^n p(x_i; \theta) = \sum_{i=1}^n \ln p(x_i; \theta),$$

并且在 $n \rightarrow \infty$ 时, $\hat{\theta}$ 依概率收敛于参数真值 θ_0 .

考虑如下双边检验问题:

简单原假设 $H_0: \theta = \theta_0$ 对 备择假设 $H_1: \theta \neq \theta_0$,

该检验问题的似然比统计量为

$$\lambda(X) = \frac{\prod_{i=1}^n p(X_i; \hat{\theta})}{\prod_{i=1}^n p(X_i; \theta_0)}.$$

那么在原假设 H_0 成立时, $2 \ln \lambda(X)$ 随 n 增大而依分布收敛于 $\chi^2(k)$.

证明: 下面的讨论都是在原假设 H_0 成立的条件下进行的. 由于 $\hat{\theta}$ 是似然方程的相合解, 则据 MLE 的渐近正态性定理可知,

$$D(\theta_0) \xrightarrow{L} N(\mathbf{0}, I^{-1}(\theta_0)), \quad n \rightarrow \infty, \quad (3.61)$$

其中 $D(\theta_0) = \sqrt{n}(\hat{\theta} - \theta_0)$. 不难证明(习题 3.17)

$$I_{ij}(\theta) = E_{\theta} \left[\frac{\partial \ln p(X; \theta)}{\partial \theta_i} \frac{\partial \ln p(X; \theta)}{\partial \theta_j} \right] = - E_{\theta} \left[\frac{\partial^2 \ln p(X; \theta)}{\partial \theta_i \partial \theta_j} \right],$$

将 $l(\theta_0) = \sum_{i=1}^n \ln p(x_i; \theta_0)$ 在 $\hat{\theta}$ 处 Taylor 展开. 由于 $\hat{\theta}$ 是似然方程的解, 则有

$$\begin{aligned} l(\theta_0) &= l(\hat{\theta}) + (\theta_0 - \hat{\theta})^T l'(\hat{\theta}) + \frac{1}{2}(\theta_0 - \hat{\theta})^T l''(\theta^*)(\theta_0 - \hat{\theta}) \\ &= l(\hat{\theta}) - \frac{1}{2} D(\theta_0)^T C D(\theta_0), \end{aligned}$$

其中

$$C = (c_{ij}(\theta^*))_{k \times k}, \quad c_{ij}(\theta) = -\frac{1}{n} \sum_{r=1}^n \frac{\partial^2 \ln p(x_r; \theta)}{\partial \theta_i \partial \theta_j}.$$

这里 $\theta^* = \hat{\theta} - \eta(\hat{\theta} - \theta_0)$, $0 \leq \eta \leq 1$, 则有

$$\begin{aligned} 2 \ln \lambda(x) &= 2 \left(\sum_{i=1}^n \ln p(x_i; \hat{\theta}) - \sum_{i=1}^n \ln p(x_i; \theta_0) \right) \\ &= 2(l(\hat{\theta}) - l(\theta_0)) = D(\theta_0)^T C D(\theta_0). \end{aligned} \quad (3.62)$$

将 $c_{ij}(\theta^*)$ 在 θ_0 处 Taylor 展开, 并应用定理关于三阶偏导数的假设条件, 则有

$$c_{ij}(\boldsymbol{\theta}^*) = c_{ij}(\boldsymbol{\theta}_0) + (\boldsymbol{\theta}^* - \boldsymbol{\theta}_0)^T \left[\frac{\partial c_{ij}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]_{\boldsymbol{\theta} = \boldsymbol{\theta}_0 - v(\boldsymbol{\theta}_0 - \boldsymbol{\theta}^*)}$$

$$\leq c_{ij}(\boldsymbol{\theta}_0) + \left(\sum_{l=1}^k |\theta_l^* - \theta_{0l}| \right) \frac{\sum_{r=1}^n M(x_r)}{n},$$

其中 $0 \leq v \leq 1$, $\boldsymbol{\theta}^* = (\theta_1^*, \theta_2^*, \dots, \theta_k^*)$, $\boldsymbol{\theta}_0 = (\theta_{01}, \theta_{02}, \dots, \theta_{0k})$. 由于 $\hat{\boldsymbol{\theta}}$ 随 n 增大而依概率收敛于参数真值 $\boldsymbol{\theta}_0$, 从而 $\boldsymbol{\theta}^* \xrightarrow{P} \boldsymbol{\theta}_0$. 故对每一个 l , 都有 $\theta_l^* \xrightarrow{P} \theta_{0l}$. 又由大数定律知, $\frac{1}{n} \sum_{r=1}^n M(X_r) \xrightarrow{P} E_{\boldsymbol{\theta}_0}(M(X)) < K$, 所以 $c_{ij}(\boldsymbol{\theta}^*) \xrightarrow{P} c_{ij}(\boldsymbol{\theta}_0)$. 由大数定律知

$$c_{ij}(\boldsymbol{\theta}_0) \xrightarrow{P} -E_{\boldsymbol{\theta}_0} \left(\frac{\partial^2 \ln p(X; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right) = I_{ij}(\boldsymbol{\theta}_0),$$

所以 $c_{ij}(\boldsymbol{\theta}^*) \xrightarrow{P} I_{ij}(\boldsymbol{\theta}_0)$. 从而

$$\boldsymbol{C} \xrightarrow{P} \boldsymbol{I}(\boldsymbol{\theta}_0), \quad (3.63)$$

则由式(3.62)知, 在 $n \rightarrow \infty$ 时, $2 \ln \lambda(X)$ 与 $\boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{I}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0)$ 有相同的极限分布. 由式(3.61)知, 在 $n \rightarrow \infty$ 时, $\boldsymbol{I}^{\frac{1}{2}}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0) \xrightarrow{L} N_k(\mathbf{0}, \boldsymbol{I}_k)$. 故 $\boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{I}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0)$, 从而 $2 \ln \lambda(X)$ 随 n 增大而依分布收敛于 $\chi^2(k)$. 定理证毕.

由式(3.62)和式(3.63)知,

$$2 \ln \lambda(X) = 2 \left(\sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}) - \sum_{i=1}^n \ln p(X_i; \boldsymbol{\theta}_0) \right)$$

$$= \boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{C} \boldsymbol{D}(\boldsymbol{\theta}_0)$$

$$= \boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{I}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0) + \xi_n, \quad (3.64)$$

其中 $\xi_n \xrightarrow{P} 0$. 这说明, $2 \ln \lambda(X)$ 与 $\boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{I}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0)$ 仅相差一个依概率收敛于 0 的量. 在将定理 3.18 推广到复合原假设的情况时, 式(3.64)起着重要的作用.

例 3.16 分类数据的检验. 根据某项指标, 总体被分成 r 类: $A_1, A_2, \dots, A_r$. 由经验或某个理论提出了如下的检验问题:

原假设 H_0 : 随机抽取一个个体 x 进行观测, $P(x \in A_i)$ 已知, 等于 p_{0i} ($i = 1, 2, \dots, r$), 其中

$$\sum_{i=1}^r p_{0i} = 1.$$

备择假设 $H_1: P(x \in A_i) = p_{0i} (i=1, 2, \dots, r)$ 不全成立.

随机抽取 n 个个体进行观测. 设有 n_i 个个体属于类 A_i . 则有 $\sum_{i=1}^r n_i = n$. 样本 $\{n_i: i=1, 2, \dots, r\}$ 服从多项分布, 其概率密度为

$$p(n_1, n_2, \dots, n_r) = \frac{n!}{n_1! n_2! \dots n_r!} \prod_{i=1}^r p_i^{n_i},$$

这里 p_i 是类 A_i 所占的比例. 由于 $\sum_{i=1}^r p_i = 1$, 所以参数空间

$$\Theta = \{(p_1, p_2, \dots, p_r) : p_i \geq 0, i=1, 2, \dots, r; \sum_{i=1}^r p_i = 1\}$$

中独立参数只有 $r-1$ 个. 故 Θ 是 $r-1$ 维欧氏空间的一个含有内点的集合. p_i 的 MLE 为 $\hat{p}_i = n_i/n$. 该检验问题的似然比统计量为

$$\Lambda = \frac{\prod_{i=1}^r \left(\frac{n_i}{n}\right)^{n_i}}{\prod_{i=1}^r (p_{0i})^{n_i}}.$$

据定理 3.18 知, 在原假设 H_0 为真时, $2 \ln \Lambda = 2 \sum_{i=1}^r n_i \ln \frac{n_i}{np_{0i}} \xrightarrow{L} \chi^2(r-1)$. 故

在 $2 \sum_{i=1}^r n_i \ln \frac{n_i}{np_{0i}} \geq \chi_{1-\alpha}^2(r-1)$ 时, 拒绝原假设.

在原假设 H_0 为真时, 可以证明 (见习题 3.19)

$$\begin{aligned} 2 \ln \Lambda &= 2 \sum_{i=1}^r n_i \ln \left(1 + \frac{n_i - np_{0i}}{np_{0i}} \right) \\ &= 2 \sum_{i=1}^r n_i \left[\frac{n_i - np_{0i}}{np_{0i}} - \frac{1}{2} \left(\frac{n_i - np_{0i}}{np_{0i}} \right)^2 \right] + \text{依概率收敛于 0 的量} \\ &= 2 \sum_{i=1}^r [np_{0i} + (n_i - np_{0i})] \left[\frac{n_i - np_{0i}}{np_{0i}} - \frac{1}{2} \left(\frac{n_i - np_{0i}}{np_{0i}} \right)^2 \right] + \text{依概率收} \end{aligned}$$

敛于 0 的量

$$= \sum_{i=1}^r \frac{(n_i - np_{0i})^2}{np_{0i}} + \text{依概率收敛于 0 的量.}$$

右边就是分类数据检验问题的 Pearson χ^2 拟合优度检验统计量. 所以 χ^2 拟合优度检验可以近似看作分类数据的似然比检验.

§ 3.6.3 复合原假设的检验问题

设样本 $X=(X_1, X_2, \dots, X_n)$ 来自密度函数族 $\{p(x; \theta) : \theta \in \Theta\}$, 其中未知参数 $\theta=(\theta_1, \theta_2, \dots, \theta_k)$, 参数空间 Θ 是 k 维欧氏空间 $\mathbf{R}^k$ 中的一个含有内点的集合. 考虑复合原假设检验问题: 原假设 $H_0: \theta \in \Theta_0$ 对备择假设 $H_1: \theta \notin \Theta_0$, 其中 Θ_0 是 Θ 的一个非单点子集. 一般来说, 复合原假设大致有以下几种情况:

情况 1. 原假设 $H_0: \theta_{r+1} = \dots = \theta_k = 0$.

这时, $\Theta_0 = \{\theta: \theta = (\theta_1, \dots, \theta_r, 0, \dots, 0) \in \Theta\}$. 例如, 总体为 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty, \sigma^2 > 0$. 这时, $\theta = (\mu, \sigma^2) \in \mathbf{R} \times \mathbf{R}^+$. 复合原假设 $H_0: \mu = 0$ 就属于情况 1. 情况 1 最为简单. 下述情况 2 是情况 1 的推广.

情况 2. 原假设 $H_0: R_i(\theta_1, \theta_2, \dots, \theta_k) = 0, i=1, 2, \dots, s$, 其中 $R_i(i=1, 2, \dots, s)$ 具有一阶连续偏导数. 这时, 参数 θ 受到 s 个条件的限制, $\Theta_0 = \{\theta: \theta \in \Theta; R_i(\theta) = 0, i=1, 2, \dots, s\}$. 例如, 相互独立的两个总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$, 其中 $-\infty < \mu_1, \mu_2 < \infty, \sigma^2 > 0$. 这时 $\theta = (\mu_1, \mu_2, \sigma^2) \in \mathbf{R} \times \mathbf{R} \times \mathbf{R}^+$. 复合原假设 $H_0: \mu_1 = \mu_2$ 就属于情况 2. 事实上, 可适当地重新定义参数 $\omega = (\omega_1, \omega_2, \dots, \omega_k)$, 使得原假设变换为 $H_0: \omega_{k-s+1} = \dots = \omega_k = 0$, 则情况 2 就变成情况 1. 在上例中, 若用新参数 $(\mu_1, \mu_1 - \mu_2, \sigma^2)$ 取代原来的参数 (μ_1, μ_2, σ^2) , 就变成情况 1.

情况 3. 原假设 $H_0: \theta_i = g_i(\varphi_1, \varphi_2, \dots, \varphi_r), i=1, 2, \dots, k (r < k)$, 其中 $g_i(i=1, 2, \dots, k)$ 具有一阶连续偏导数. 这时, $\Theta_0 = \{\theta: \theta \in \Theta; \theta_i = g_i(\varphi), i=1, 2, \dots, k, \varphi = (\varphi_1, \varphi_2, \dots, \varphi_r) \in A\}$, A 是 r 维欧氏空间 $\mathbf{R}^r$ 中的一个含有内点的集合. 例如, 设有 X, Y 两个离散型随机变量. X 表示人的性别, “ $X=1$ ”表示男性, “ $X=2$ ”表示女性; Y 表示是否色盲, “ $Y=1$ ”表示正常, “ $Y=2$ ”表示色盲. 令 $p_{ij} = P(X=i, Y=j), i, j=1, 2$. 这时 $\theta = (p_{11}, p_{12}, p_{21})$, 其中 $0 \leq p_{11}, p_{12}, p_{21} \leq 1$, 且 $p_{11} + p_{12} + p_{21} \leq 1$. 所以参数空间 Θ 为 $\mathbf{R}^3$ 中的一个单纯形.

有 1 000 人按性别与色盲分类如下

	正常	色盲
男	442	38
女	514	6

按遗传学模型,数据应有下列相对应的概率

	正常	色盲
男	$\frac{p}{2}$	$\frac{1-p}{2}$
女	$\frac{p^2}{2} + p(1-p)$	$\frac{(1-p)^2}{2}$

其中 $0 \leq p \leq 1$. 问数据与模型是否相符? 这里所要检验的原假设为

$$H_0: p_{11} = \frac{p}{2}, \quad p_{12} = \frac{1-p}{2}, \quad p_{21} = \frac{p^2}{2} + p(1-p),$$

这属于情况 3.

取 $\varphi = (\theta_1, \theta_2, \dots, \theta_r)$, 则情况 1 变换成情况 3. 从而情况 2 也能变换成情况 3. 事实上, 情况 3 也可以变换成情况 2. 如果能将 $\varphi_1, \varphi_2, \dots, \varphi_r$ 表示成 $\theta_1, \theta_2, \dots, \theta_r$ 的函数, 然后把把这些函数代入 $\theta_j = g_j(\varphi_1, \varphi_2, \dots, \varphi_r)$, $j=r+1, \dots, k$, 由此可得到参数 θ 受到的 $k-r$ 个条件的限制, 则情况 3 变换成情况 2. 在上例中, 由于 $p_{11} = p/2$, 所以 $p = 2p_{11}$. 因此参数 $\theta = (p_{11}, p_{12}, p_{21})$ 受到两个条件的限制: $p_{12} = (1 - 2p_{11})/2$, 以及 $p_{21} = 2p_{11}^2 + 2p_{11}(1 - 2p_{11})$.

下面我们只讨论情况 3 的复合原假设的检验问题.

定理 3.19 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自密度函数族 $\{p(x; \theta) : \theta \in \Theta\}$, 其中未知参数 $\theta = (\theta_1, \theta_2, \dots, \theta_k)$, 参数空间 Θ 是 k 维欧氏空间 $\mathbf{R}^k$ 中的一个含有内点的集合. 参数真值 θ_0 是 Θ 的一个内点. 假设该密度函数族满足定理 3.18 的四个条件.

考虑原假设 $H_0: \theta \in \Theta_0$ 对备择假设 $H_1: \theta \notin \Theta_0$, 其中 Θ_0 是 Θ 的一个子集, 且满足下述条件: 存在 r ($r < k$) 维欧氏空间 $\mathbf{R}^r$ 中的一个含有内点的集合 A , 以及定义在 A 上的 k 元三阶可导的函数 $g = (g_1, g_2, \dots, g_k)$, 使得 A 与 Θ_0 一一对应,

$$\Theta_0 = \{\theta : \theta = g(\varphi), \varphi \in A\}.$$

在参数真值 $\theta_0 \in \Theta_0$ 时, 假设 θ_0 对应于 A 中的点 $\varphi_0: \theta_0 = g(\varphi_0)$, φ_0 是 A 的内点.

在 $\theta = g(\varphi)$ 时, 记 $p(x; \theta) = p(x; g(\varphi))$ 为 $\tilde{p}(x; \varphi)$, 则函数族 $\{p(x; \theta) : \theta \in \Theta_0\}$ 可等价地写成 $\{\tilde{p}(x; \varphi) : \varphi \in A\}$ 的形式. 假设密度函数族 $\{\tilde{p}(x; \varphi) : \varphi \in A\}$ 满足定理 3.18 的四个条件.

在 $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ 时, 假设 $\boldsymbol{\theta}$ 的 MLE $\hat{\boldsymbol{\theta}}$ 是似然方程 $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0$ 的解, 其中

$$l(\boldsymbol{\theta}) = \ln \prod_{i=1}^n p(x_i; \boldsymbol{\theta}) = \sum_{i=1}^n \ln p(x_i; \boldsymbol{\theta}),$$

并且在 $n \rightarrow \infty$ 时, $\hat{\boldsymbol{\theta}}$ 依概率收敛于参数真值 $\boldsymbol{\theta}_0$. 在原假设为真, 即参数真值 $\boldsymbol{\theta}_0 \in$

$\boldsymbol{\Theta}_0$ 时, $\boldsymbol{\theta}$ 的 MLE $\hat{\boldsymbol{\theta}}_0 = \boldsymbol{g}(\hat{\boldsymbol{\varphi}}_0)$, 假设 $\hat{\boldsymbol{\varphi}}_0$ 是似然方程 $\frac{\partial \tilde{l}(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} = 0$ 的解, 其中

$$\tilde{l}(\boldsymbol{\varphi}) = \ln \prod_{i=1}^n \tilde{p}(x_i; \boldsymbol{\varphi}) = \sum_{i=1}^n \ln \tilde{p}(x_i; \boldsymbol{\varphi}),$$

并且在 $n \rightarrow \infty$ 时, $\hat{\boldsymbol{\varphi}}_0$ 依概率收敛于 $\boldsymbol{\varphi}_0$.

定义似然比统计量

$$\lambda(X) = \frac{\prod_{i=1}^n p(X_i; \hat{\boldsymbol{\theta}})}{\prod_{i=1}^n p(X_i; \hat{\boldsymbol{\theta}}_0)},$$

则在原假设 H_0 成立时, $2 \ln \lambda(X)$ 随 n 增大而依分布收敛于 $\chi^2(k-r)$.

证明: 下面的讨论都是在原假设 H_0 成立的条件下进行的, 并且沿袭定理 3.18 的符号.

令

$$\boldsymbol{V}(\boldsymbol{\theta}) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \frac{\partial \ln p(x_i; \boldsymbol{\theta})}{\partial \theta_1}, \dots, \sum_{i=1}^n \frac{\partial \ln p(x_i; \boldsymbol{\theta})}{\partial \theta_k} \right)^T,$$

则由多维中心极限定理知, 在 $n \rightarrow \infty$ 时,

$$\boldsymbol{V}(\boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \boldsymbol{I}(\boldsymbol{\theta}_0)), \quad (3.65)$$

从而由式 (3.61) 知, $\boldsymbol{I}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{V}(\boldsymbol{\theta}_0)$ 与 $\boldsymbol{D}(\boldsymbol{\theta}_0)$ 有相同的极限分布. 不仅如此, 由 MLE 渐近正态性定理的证明过程知, $\boldsymbol{I}^{-1}(\boldsymbol{\theta}_0) \boldsymbol{V}(\boldsymbol{\theta}_0)$ 与 $\boldsymbol{D}(\boldsymbol{\theta}_0)$ 仅相差一个依概率收敛于 0 的量.

因为密度函数族 $\{p(x; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ 满足定理 3.18 的四个条件, 则由式 (3.64) 知

$$2 \left(\sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}) - \sum_{i=1}^n \ln p(X_i; \boldsymbol{\theta}_0) \right) \text{ 与 } \boldsymbol{D}(\boldsymbol{\theta}_0)^T \boldsymbol{I}(\boldsymbol{\theta}_0) \boldsymbol{D}(\boldsymbol{\theta}_0),$$

从而

$$2 \left(\sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}) - \sum_{i=1}^n \ln p(X_i; \boldsymbol{\theta}_0) \right) \text{ 与 } \mathbf{V}^T(\boldsymbol{\theta}_0) \mathbf{I}^{-1}(\boldsymbol{\theta}_0) \mathbf{V}(\boldsymbol{\theta}_0)$$

仅相差一个依概率收敛于 0 的量, 它们有相同的极限分布.

同理, 因为密度函数族 $\{\tilde{p}(x; \boldsymbol{\varphi}) : \boldsymbol{\varphi} \in A\}$ 也满足定理 3.18 的四个条件, 所以

$$2 \left(\sum_{i=1}^n \ln \tilde{p}(X_i; \hat{\boldsymbol{\varphi}}) - \sum_{i=1}^n \ln \tilde{p}(X_i; \boldsymbol{\varphi}_0) \right) \text{ 与 } \mathbf{U}^T(\boldsymbol{\varphi}_0) \mathbf{J}^{-1}(\boldsymbol{\varphi}_0) \mathbf{U}(\boldsymbol{\varphi}_0)$$

仅相差一个依概率收敛于 0 的量, 它们有相同的极限分布, 其中

$$\mathbf{U}(\boldsymbol{\varphi}) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \frac{\partial \ln \tilde{p}(x_i; \boldsymbol{\varphi})}{\partial \varphi_1}, \dots, \sum_{i=1}^n \frac{\partial \ln \tilde{p}(x_i; \boldsymbol{\varphi})}{\partial \varphi_r} \right)^T,$$

$$\mathbf{J}(\boldsymbol{\varphi}) = (J_{ij}(\boldsymbol{\varphi}))_{r \times r},$$

其中

$$J_{ij}(\boldsymbol{\varphi}) = \int_x \frac{\partial \ln \tilde{p}(x; \boldsymbol{\varphi})}{\partial \varphi_i} \frac{\partial \ln \tilde{p}(x; \boldsymbol{\varphi})}{\partial \varphi_j} \tilde{p}(x; \boldsymbol{\varphi}) d\mu(x).$$

令 $\mathbf{B}(\boldsymbol{\varphi}) = (B_{ij}(\boldsymbol{\varphi}))_{r \times k}$, 其中 $B_{ij}(\boldsymbol{\varphi}) = \frac{\partial g_j(\boldsymbol{\varphi})}{\partial \varphi_i}$. 注意到 $\boldsymbol{\theta} = \mathbf{g}(\boldsymbol{\varphi})$, $\boldsymbol{\theta}_0 = \mathbf{g}(\boldsymbol{\varphi}_0)$. 从

而有

$$\begin{aligned} \mathbf{U}(\boldsymbol{\varphi}_0) &= \mathbf{B}(\boldsymbol{\varphi}_0) \mathbf{V}(\boldsymbol{\theta}_0), \\ \mathbf{J}(\boldsymbol{\varphi}_0) &= \mathbf{B}(\boldsymbol{\varphi}_0) \mathbf{I}(\boldsymbol{\theta}_0) \mathbf{B}^T(\boldsymbol{\varphi}_0), \end{aligned} \quad (3.66)$$

所以

$$2 \left(\sum_{i=1}^n \ln \tilde{p}(X_i; \hat{\boldsymbol{\varphi}}) - \sum_{i=1}^n \ln \tilde{p}(X_i; \boldsymbol{\varphi}_0) \right)$$

与

$$\mathbf{V}(\boldsymbol{\theta}_0)^T \mathbf{B}^T(\boldsymbol{\varphi}_0) \mathbf{J}^{-1}(\boldsymbol{\varphi}_0) \mathbf{B}(\boldsymbol{\varphi}_0) \mathbf{V}(\boldsymbol{\theta}_0)$$

仅相差一个依概率收敛于 0 的量, 它们有相同的极限分布.

注意到 $p(x; \boldsymbol{\theta}_0) = \tilde{p}(x; \boldsymbol{\varphi}_0)$, $p(x; \hat{\boldsymbol{\theta}}_0) = \tilde{p}(x; \hat{\boldsymbol{\varphi}}_0)$, 从而有

$$\begin{aligned}
2\ln \lambda(X) &= 2 \left(\sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}) - \sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}_0) \right) \\
&= 2 \left(\sum_{i=1}^n \ln p(X_i; \hat{\boldsymbol{\theta}}) - \sum_{i=1}^n \ln p(X_i; \boldsymbol{\theta}_0) \right) - \\
&\quad 2 \left(\sum_{i=1}^n \ln \tilde{p}(X_i; \hat{\boldsymbol{\varphi}}_0) - \sum_{i=1}^n \ln \tilde{p}(X_i; \boldsymbol{\varphi}_0) \right),
\end{aligned}$$

则 $2\ln \lambda(X)$ 与

$$\mathbf{V}(\boldsymbol{\theta}_0)^T [\mathbf{I}^{-1}(\boldsymbol{\theta}_0) - \mathbf{B}^T(\boldsymbol{\varphi}_0) \mathbf{J}^{-1}(\boldsymbol{\varphi}_0) \mathbf{B}(\boldsymbol{\varphi}_0)] \mathbf{V}(\boldsymbol{\theta}_0)$$

仅相差一个依概率收敛于 0 的量, 它们有相同的极限分布. 令 $\mathbf{W}(\boldsymbol{\theta}_0) = \mathbf{I}^{-\frac{1}{2}}(\boldsymbol{\theta}_0)$

$\mathbf{V}(\boldsymbol{\theta}_0)$, 则由式(3.65)知, $\mathbf{W}(\boldsymbol{\theta}_0) \xrightarrow{L} N(\mathbf{0}, \mathbf{I})$, 且 $2\ln \lambda(X)$ 与

$$\mathbf{W}^T(\boldsymbol{\theta}_0) (\mathbf{I} - \mathbf{G}(\boldsymbol{\theta}_0)) \mathbf{W}(\boldsymbol{\theta}_0)$$

有相同的极限分布, 其中

$$\mathbf{G}(\boldsymbol{\theta}_0) = \mathbf{I}^{\frac{1}{2}}(\boldsymbol{\theta}_0) \mathbf{B}^T(\boldsymbol{\varphi}_0) \mathbf{J}^{-1}(\boldsymbol{\varphi}_0) \cdot \mathbf{B}(\boldsymbol{\varphi}_0) \mathbf{I}^{\frac{1}{2}}(\boldsymbol{\theta}_0).$$

由式(3.66)知, $\mathbf{I} - \mathbf{G}(\boldsymbol{\theta}_0)$ 是幂等矩阵, 且 $\mathbf{I} - \mathbf{G}(\boldsymbol{\theta}_0)$ 的秩为

$$\begin{aligned}
k - \text{tr}(\mathbf{G}(\boldsymbol{\theta}_0)) &= k - \text{tr}(\mathbf{J}^{-1}(\boldsymbol{\varphi}_0) \mathbf{B}(\boldsymbol{\varphi}_0) \mathbf{I}(\boldsymbol{\theta}_0) \mathbf{B}^T(\boldsymbol{\varphi}_0)) \\
&= k - r,
\end{aligned}$$

所以 $2\ln \lambda(X)$ 的极限分布为 $\chi^2(k-r)$. 定理证毕.

§ 3.6.4 二维列联表的独立性检验

设有 X, Y 两个离散型随机变量, 分别取 r 个和 c 个值, 与此相联系的有一个二维列联表 $r \times c$. 作 n 次观测, 在 (i, j) 格的观测频数为 $n_{ij}, i=1, 2, \dots, r; j=1, 2, \dots, c$. 观测值落入 (i, j) 格的概率为 p_{ij} . 观测频数 $\{n_{ij}, i=1, 2, \dots, r; j=1, 2, \dots, c\}$ 服从多项分布, 其概率密度为

$$\frac{n!}{\prod_{i=1}^r \prod_{j=1}^c n_{ij}!} \prod_{i=1}^r \prod_{j=1}^c p_{ij}^{n_{ij}}.$$

由于 $\sum_{i=1}^r \sum_{j=1}^c p_{ij} = 1$, 所以参数空间

$$\Theta = \left\{ p_{ij}; i = 1, 2, \dots, r; j = 1, 2, \dots, c; p_{ij} \geq 0, \sum_{i=1}^r \sum_{j=1}^c p_{ij} = 1 \right\}$$

中独立参数只有 $rc-1$ 个. 故 Θ 是 $rc-1$ 维欧氏空间的一个含有内点的集合. p_{ij} 的 MLE 为 $\hat{p}_{ij} = n_{ij}/n$.

考虑独立性检验, 原假设 $H_0: p_{ij} = p_{i\cdot} \cdot p_{\cdot j}, i = 1, 2, \dots, r; j = 1, 2, \dots, c$, 其中 $p_{i\cdot}$ 和 $p_{\cdot j}$ 分别是 X 和 Y 的边缘分布. 原假设成立时, $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$, 所以其参数空间 Θ_0 与下述集合 A 一一对应:

$$A = \left\{ (p_{i\cdot}, p_{\cdot j}); i = 1, 2, \dots, r; j = 1, 2, \dots, c, p_{i\cdot} \geq 0, p_{\cdot j} \geq 0, \sum_{i=1}^r p_{i\cdot} = \sum_{j=1}^c p_{\cdot j} = 1 \right\}.$$

A 中独立参数只有 $r+c-2$ 个. 故 A 是 $r+c-2$ 维欧氏空间的一个含有内点的集合.

这时, p_{ij} 的 MLE 为 $\hat{p}_{i\cdot}, \hat{p}_{\cdot j} = \frac{n_{i\cdot}}{n}, \frac{n_{\cdot j}}{n}$, 其中 $n_{i\cdot} = \sum_{j=1}^c n_{ij}, n_{\cdot j} = \sum_{i=1}^r n_{ij}$.

该检验问题的似然比统计量为

$$\Lambda = \frac{\prod_{i=1}^r \prod_{j=1}^c \left(\frac{n_{ij}}{n} \right)^{n_{ij}}}{\prod_{i=1}^r \prod_{j=1}^c \left(\frac{n_{i\cdot}}{n} \cdot \frac{n_{\cdot j}}{n} \right)^{n_{ij}}},$$

由于 $(rc-1) - (r+c-2) = (r-1)(c-1)$, 故在原假设成立时, $2\ln \Lambda$ 的极限分布为 $\chi^2((r-1)(c-1))$. 在 $2\ln \Lambda \geq \chi^2_{1-\alpha}((r-1)(c-1))$ 时, 也就是在

$$2 \sum_{i=1}^r \sum_{j=1}^c n_{ij} \ln \frac{nn_{ij}}{n_{i\cdot} n_{\cdot j}} \geq \chi^2_{1-\alpha}((r-1)(c-1))$$

的时候, 拒绝原假设.

与例 3.16 相类似地, 在原假设 H_0 为真时, 可以推出 (见习题 3.19)

$$2\ln \Lambda = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n\hat{p}_{i\cdot}\hat{p}_{\cdot j})^2}{n\hat{p}_{i\cdot}\hat{p}_{\cdot j}} + \text{依概率收敛于 0 的量},$$

右边就是二维列联表的独立性检验问题的 χ^2 拟合优度检验统计量.

§ 3.6.5 三维列联表的条件独立性检验

首先看下面的例子. 表 3.2 是 1976—1977 年美国佛罗里达州的凶杀案件中, 326 个被告人的肤色和死刑判决情况的分类表^[6].

表 3.2 被告人的肤色和死刑判决情况的分类表

被告人	死刑		合计
	是	否	
白人	19	141	160
黑人	17	149	166
合计	36	290	326

二维列联表的独立性检验似然比统计量 $2\ln \Lambda = 0.2214$, 它比 $\chi^2(1)$ 分布的中位数 0.455 0 还要小, 不能拒绝被告人的肤色和死刑判决相互独立的原假设, 认为白人被判死刑的概率等于黑人被判死刑的概率. 事实上, 表 3.2 的数据显示, 白人被判死刑的比例较高, 等于 11.9%, 而黑人被判死刑的比例却较低, 等于 10.2%. 这显然与当时美国的现实是不相符的. 后来人们将被害人的肤色也考虑进去, 构造了三维列联表, 得到的结论就不一样了. 三维列联表如表 3.3 所示.

表 3.3 被告人与被害人的肤色以及死刑判决情况的分类表

被告人	被害人	死刑		死刑判决的比例
		是	否	
白人	白人	19	132	0.126
	黑人	0	9	0.000
黑人	白人	11	52	0.175
	黑人	6	97	0.058

如果被害人是白人, 那么白人被判死刑的比例与黑人被判死刑的比例分别为 12.6% 与 17.5%; 如果被害人是黑人, 那么白人被判死刑的比例与黑人被判死刑的比例分别为 0 与 5.8%. 这说明, 无论被害人是白人还是黑人, 黑人被判死刑的比例都是比较高的. 尤其是在被害人是黑人时, 白人没有被判死刑的. 这说明在美国有种族歧视, 死刑判决与肤色密切相关. 此例告诉我们, 在将三维列联表合并成二维列联表进行分析时, 由于删去了一个因子, 有可能产生问题. 人们有必要研究三维列联表, 或更高维列联表的统计分析.

设有 X, Y, Z 三个离散型随机变量, 与此相联系的有一个三维列联表 $r \times c \times t$. 做 n 次观测, 在 (i, j, k) 格的观测频数为 n_{ijk} , $i = 1, 2, \dots, r; j = 1, 2, \dots, c, k = 1, 2, \dots, t$. 观测值落入 (i, j, k) 格的概率为 p_{ijk} . $\{n_{ijk}, i = 1, 2, \dots, r; j = 1, 2, \dots, c, k = 1, 2, \dots, t\}$ 服从多项分布, 其概率密度为

$$\frac{n!}{\prod_{i=1}^r \prod_{j=1}^c \prod_{k=1}^t n_{ijk}!} \prod_{i=1}^r \prod_{j=1}^c \prod_{k=1}^t p_{ijk}^{n_{ijk}},$$

由于 $\sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^t p_{ijk} = 1$, 所以参数空间

$$\Theta = \left\{ p_{ijk} : i = 1, 2, \dots, r; j = 1, 2, \dots, c; k = 1, 2, \dots, t; p_{ijk} \geq 0, \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^t p_{ijk} = 1 \right\}$$

中独立参数只有 $rct-1$ 个. 故 Θ 是 $rct-1$ 维欧氏空间的一个含有内点的集合. p_{ijk} 的 MLE 为 $\hat{p}_{ijk} = n_{ijk}/n$.

这里有三个检验问题, 原假设分别是三个随机变量相互独立, 一个随机变量和另两个随机变量相互独立, 以及给定一个随机变量后另两个随机变量相互条件独立. 前两个问题类似于二维列联表的独立性检验问题, 故从略. 最后一个问题是有关条件独立性的检验, 下面来讨论这个检验问题. 这里仅考虑给定 Z 之后, X 和 Y 相互条件独立的检验问题.

若条件独立性成立, 则有

$$P(X = i, Y = j | Z = k) = P(X = i | Z = k)P(Y = j | Z = k),$$

从而有

$$p_{ijk} = \frac{p_{i \cdot k} p_{\cdot jk}}{p_{\cdot \cdot k}}, \quad (3.67)$$

其中 $p_{i \cdot k}$, $p_{\cdot jk}$ 和 $p_{\cdot \cdot k}$ 分别是 (X, Z) , (Y, Z) 和 Z 的边缘分布. 由此 (习题 3.21) 推得, 条件独立性成立时, p_{ijk} 的 MLE 为

$$\frac{\hat{p}_{i \cdot k} \hat{p}_{\cdot jk}}{\hat{p}_{\cdot \cdot k}} = \frac{n_{i \cdot k} n_{\cdot jk}}{nn_{\cdot \cdot k}}.$$

条件独立性的原假设成立时, 由式 (3.67) 知, 其参数空间 Θ_0 与下述集合 A 一一对应:

$$A = \left\{ (p_{i \cdot k}, p_{\cdot jk}, p_{\cdot \cdot k}) : i = 1, 2, \dots, r; j = 1, 2, \dots, c; k = 1, 2, \dots, t; \right.$$

$$\left. p_{i \cdot k} \geq 0, p_{\cdot jk} \geq 0, \sum_{i=1}^r p_{i \cdot k} = \sum_{j=1}^c p_{\cdot jk} = p_{\cdot \cdot k}, \sum_{k=1}^t p_{\cdot \cdot k} = 1 \right\},$$

A 中独立参数只有 $t(r+c-1)-1$ 个 (习题 3.21).

该检验问题的似然比统计量为

$$\Lambda = \frac{\prod_{i=1}^r \prod_{j=1}^c \prod_{k=1}^t \left(\frac{n_{ijk}}{n} \right)^{n_{ijk}}}{\prod_{i=1}^r \prod_{j=1}^c \prod_{k=1}^t \left(\frac{n_{i \cdot k} n_{\cdot jk}}{nn_{\cdot \cdot k}} \right)^{n_{ijk}}}.$$

由于 $(rct-1) - [t(r+c-1)-1] = t(r-1)(c-1)$, 故在原假设成立时, $2\ln A$ 的极限分布为 $\chi^2(t(r-1)(c-1))$. 在 $2\ln A \geq \chi^2_{1-\alpha}(t(r-1)(c-1))$ 时, 即在

$$2 \sum_{i=1}^r \sum_{j=1}^c \sum_{k=1}^t n_{ijk} \ln \frac{n_{i..} n_{.jk}}{n_{i.k} n_{.j.}} \geq \chi^2_{1-\alpha}(t(r-1)(c-1))$$

的时候, 拒绝原假设.

上述关于佛罗里达州的凶杀案件的例子中, 考虑给定被害人的肤色后, 被告人的肤色与死刑判决的条件独立性的检验问题(给定被告人的肤色后被害人的肤色与死刑判决的条件独立性的检验问题, 以及给定是否死刑判决后被告人的肤色与被害人的肤色的条件独立性的检验问题留作习题 3.22). 经计算, 似然比检验统计量 $2\ln A = 1.88$. 渐近 χ^2 分布的自由度为 2. 由于 $P\{\chi^2(2) \geq 1.88\} \approx 0.39$, 所以我们不拒绝条件独立性原假设. 认为给定被害人的肤色后, 被告人的肤色与死刑判决是条件独立的. 由此可以看到, 被害人的肤色与死刑判决有关, 被害人的肤色与被告人的肤色有关. 前者说明在美国有种族歧视.

与似然比检验法相类似的, 还有 Wald(沃尔德)检验法和 Rao 的 Score 检验法. 这两个检验方法的详细介绍见本章参考文献[3]和[4].

§ 3.7 U 统计量检验

§ 3.7.1 U 统计量

上述各种检验方法基本上适用于参数结构. 这些方法往往要求总体分布族的密度函数的数学形式已知, 且只含有限个未知参数. 但有的时候, 人们难以由经验或某种理论得到总体的参数结构, 而只能得到非参数结构. 因此有必要寻求非参数结构的检验方法. 首先看下面的例子.

例 3.17^[5] 令 X 表示某种电子元器件的寿命. 根据经验或某种理论, 常假设 X 服从指数分布, 或 Weibull 分布, 或对数正态分布等. 但有时很难作出这样的假设, 或人们不愿意有这样的假设, 而宁愿在一个较大的范围内考虑问题. 例如, 仅假设 X 为取非负值的连续型随机变量. 这样我们就涉及一个非参数结构.

如果对任意的 $s > 0, t > 0$, 都有 $P(X > s+t | X > t) < P(X > s)$, 则称该元件老化. 如果对任意的 $s > 0, t > 0$, 都有 $P(X > s+t | X > t) = P(X > s)$, 则称该元件无老化. 在可靠性统计中, 一个很重要的检验问题是:

原假设 H_0 : 元件无老化 对 备择假设 H_1 : 元件老化. (3.68)

若 X 取参数结构, 服从 Weibull 分布, 其分布函数为 $F(x) = 1 - \exp\left\{-\frac{x^\alpha}{\sigma}\right\}$, 则有

$$\begin{aligned} P(X > s) &= \exp\left\{-\frac{s^\alpha}{\sigma}\right\}, \\ P(X > s+t | X > t) &= \frac{P(X > s+t, X > t)}{P(X > t)} = \frac{P(X > s+t)}{P(X > t)} \\ &= \exp\left\{-\frac{(s+t)^\alpha - t^\alpha}{\sigma}\right\}. \end{aligned}$$

显然, 在 $\alpha=1$ 时, $s^\alpha = (s+t)^\alpha - t^\alpha$, 从而 $P(X > s+t | X > t) = P(X > s)$, 元件无老化; 而在 $\alpha > 1$ 时, $s^\alpha < (s+t)^\alpha - t^\alpha$, 从而 $P(X > s+t | X > t) < P(X > s)$, 元件老化. 故上述检验问题在 Weibull 分布场合就简化为关于形状参数 α 的检验问题, 原假设 $H_0: \alpha = 1$ 对备择假设 $H_1: \alpha > 1$. 受此启发, 我们在 X 取非参数结构时, 亦寻求一个参数, 使得检验问题化为关于这个参数的检验问题.

设 X 为取非负值的连续随机变量, 它的分布函数为 $F(x)$. 显然,

$$\begin{aligned} P(X > s+t | X > t) &= P(X > s) \Leftrightarrow \bar{F}(s+t) = \bar{F}(s)\bar{F}(t), \\ P(X > s+t | X > t) &< P(X > s) \Leftrightarrow \bar{F}(s+t) < \bar{F}(s)\bar{F}(t), \end{aligned}$$

其中 $\bar{F}(x) = 1 - F(x)$. 为此, 我们令

$$\omega = \int_0^\infty \int_0^\infty [\bar{F}(s)\bar{F}(t) - \bar{F}(s+t)] dF(s) dF(t),$$

则检验问题可简化为关于参数 ω 的检验问题

$$\text{原假设 } H_0: \omega = 0 \quad \text{对} \quad \text{备择假设 } H_1: \omega > 0. \quad (3.69)$$

由于

$$\begin{aligned} \int_0^\infty \int_0^\infty \bar{F}(s)\bar{F}(t) dF(s) dF(t) &= \int_0^1 \int_0^1 (1-u)(1-v) dudv = \frac{1}{4}, \\ \int_0^\infty \int_0^\infty \bar{F}(s+t) dF(s) dF(t) &= \int_0^\infty \int_0^\infty P(X_1 > x_2 + x_3) dF(x_2) dF(x_3) \\ &= P(X_1 > X_2 + X_3), \end{aligned}$$

所以

$$\omega = \frac{1}{4} - P(X_1 > X_2 + X_3),$$

其中 X_1, X_2, X_3 是该元器件的寿命试验的任意三次记录.

若令 $\theta = P(X_1 > X_2 + X_3)$, 则把检验问题 (3.69) 进一步简化为关于参数 θ 的检验问题

$$\text{原假设 } H_0: \theta = \frac{1}{4} \quad \text{对} \quad \text{备择假设 } H_1: \theta < \frac{1}{4}. \quad (3.70)$$

为了寻求检验统计量, 首先得讨论 θ 的估计问题. 这是非参数估计问题. 设 $X = (X_1, X_2, \dots, X_n)$ 是该元器件的寿命试验的 n 次记录, 即它是来自取非负值的连续随机变量的分布函数 $F(x)$ 的样本. 容易看出

$$\psi(X_1, X_2, X_3) = \begin{cases} 1, & X_1 > X_2 + X_3, \\ 0, & X_1 \leq X_2 + X_3 \end{cases}$$

是 θ 的一个无偏估计. 考虑到样本中各个分量的地位是相同的, 故定义对称函数

$$\phi(X_1, X_2, X_3) = \frac{\sum \psi(X_{i_1}, X_{i_2}, X_{i_3})}{3!},$$

它也是 θ 的一个无偏估计, 其中 (i_1, i_2, i_3) 是 $(1, 2, 3)$ 的一个置换, $\sum$ 对所有这种置换求和. $\psi(X_1, X_2, X_3)$ 和 $\phi(X_1, X_2, X_3)$ 都是 θ 的无偏估计, 但 $\phi(X_1, X_2, X_3)$ 的方差不可能比 $\psi(X_1, X_2, X_3)$ 的方差大 (习题 3.24). 进一步令

$$U(X_1, X_2, \dots, X_n) = \frac{\sum \phi(X_{i_1}, X_{i_2}, X_{i_3})}{\binom{n}{3}}, \quad (3.71)$$

其中 $1 \leq i_1 < i_2 < i_3 \leq n$, $\sum$ 对所有这种组合 (i_1, i_2, i_3) 求和. 则 $U(X_1, X_2, \dots, X_n)$ 仍是 θ 的一个无偏估计. 它的方差不可能比 $\phi(X_1, X_2, X_3)$ 的方差大 (习题 3.24). 由于在备择假设 H_1 为真时, θ 的值较小, 故很自然地, 我们在 $U(X_1, X_2, \dots, X_n) \leq c$ 时, 拒绝原假设, 认为元件老化. 如何确定临界值 c 的问题留待下面解决.

形如式 (3.71) 的统计量就是 §2.2.4 给出的 U 统计量.

§3.7.2 U 统计量的期望和方差

设有样本 $X = (X_1, X_2, \dots, X_n)$, 并设 ϕ 为 $X_1, X_2, \dots, X_m$ ($m \leq n$) 的对称函数, 令

$$U = U(X_1, X_2, \dots, X_n) = \binom{n}{m}^{-1} \sum \phi(X_{i_1}, X_{i_2}, \dots, X_{i_m}), \quad (3.72)$$

其中 $1 \leq i_1 < i_2 < \cdots < i_m \leq n$, Σ 对所有的组合 $(i_1, i_2, \cdots, i_m)$ 求和, 则 U (或 $U(X_1, X_2, \cdots, X_n)$) 是以 $\phi(X_1, X_2, \cdots, X_m)$ 为核的 U 统计量.

若令 $\theta = E(\phi(X_1, X_2, \cdots, X_m))$, 则 U 统计量的数学期望 $E(U) = \theta$. 为简化 U 统计量的方差的计算, 不妨假设 $\theta = 0$. 否则, 以 $\phi(X_1, X_2, \cdots, X_m) - \theta$ 为核构造 U 统计量. 下面, 我们来计算 U 的方差. 为此首先在 $X_1, X_2, \cdots, X_c (c \leq m)$ 给定的条件下, 计算 $\phi(X_1, X_2, \cdots, X_m)$ 的条件期望

$$\phi_c(x_1, x_2, \cdots, x_c) = E(\phi(X_1, X_2, \cdots, X_m) \mid X_1 = x_1, \cdots, X_c = x_c), \quad (3.73)$$

其中 $c = 1, 2, \cdots, m$. 则

$$\begin{aligned} E(\phi_c(X_1, X_2, \cdots, X_c)) &= E(E(\phi(X_1, X_2, \cdots, X_m) \mid X_1, X_2, \cdots, X_c)) \\ &= E(\phi(X_1, X_2, \cdots, X_m)) = 0. \end{aligned}$$

记

$$E(\phi_c(X_1, X_2, \cdots, X_c))^2 = \sigma_c^2, \quad c = 1, 2, \cdots, m, \quad (3.74)$$

若 $\phi(X_1, X_2, \cdots, X_m)$ 的方差有限, 则 σ_c^2 也有限 (习题 3.24). 由式 (3.72) 知, U 统计量的方差为

$$\text{Var}(U) = \binom{n}{m}^{-2} \sum E(\phi(X_{i_1}, X_{i_2}, \cdots, X_{i_m}) \phi(X_{j_1}, X_{j_2}, \cdots, X_{j_m})),$$

其中 $1 \leq i_1 < i_2 < \cdots < i_m \leq n, 1 \leq j_1 < j_2 < \cdots < j_m \leq n$, Σ 对所有的组合 $(i_1, i_2, \cdots, i_m)$ 和 $(j_1, j_2, \cdots, j_m)$ 求和. 显然, 在两集合 $\{i_1, i_2, \cdots, i_m\}$ 和 $\{j_1, j_2, \cdots, j_m\}$ 无公共元素时, 有

$$E(\phi(X_{i_1}, X_{i_2}, \cdots, X_{i_m}) \phi(X_{j_1}, X_{j_2}, \cdots, X_{j_m})) = 0,$$

所以

$$\text{Var}(U) = \binom{n}{m}^{-2} \sum_{c=1}^m M(c),$$

其中

$$M(c) = \sum E(\phi(X_{i_1}, X_{i_2}, \cdots, X_{i_m}) \phi(X_{j_1}, X_{j_2}, \cdots, X_{j_m})), \quad (3.75)$$

这里 $1 \leq i_1 < i_2 < \cdots < i_m \leq n, 1 \leq j_1 < j_2 < \cdots < j_m \leq n$, 且两集合 $\{i_1, i_2, \cdots, i_m\}$ 和 $\{j_1, j_2, \cdots, j_m\}$ 中恰有 $c (c = 1, 2, \cdots, m)$ 个元素相同, Σ 表示对所有这样的组合 $(i_1, i_2, \cdots, i_m)$ 和 $(j_1, j_2, \cdots, j_m)$ 求和.

由样本的独立同分布性知, 在两集合 $\{i_1, i_2, \cdots, i_m\}$ 和 $\{j_1, j_2, \cdots, j_m\}$ 中恰有 c

个元素相同时,有

$$\begin{aligned} & E(\phi(X_{i_1}, X_{i_2}, \dots, X_{i_m}) \phi(X_{j_1}, X_{j_2}, \dots, X_{j_m})) \\ &= E(\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c})), \end{aligned}$$

所以式(3.75)右边和式中的各个项都相等. 因此

$$M(c) = \binom{n}{m} \binom{m}{c} \binom{n-m}{m-c} E(\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c})),$$

在 $X_1 = x_1, \dots, X_c = x_c$ 给定的条件下, 计算 $\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c})$ 的条件期望. 由于在 $X_1 = x_1, \dots, X_c = x_c$ 给定时, $(X_1, X_2, \dots, X_m)$ 和 $(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c})$ 相互条件独立和同条件分布, 则有

$$\begin{aligned} & E(\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c}) \mid X_1 = x_1, \dots, X_c = x_c) \\ &= E(\phi(X_1, X_2, \dots, X_m) \mid X_1 = x_1, \dots, X_c = x_c) \cdot \\ & E(\phi(X_1, X_2, \dots, X_c, X_{m+1}, \dots, X_{2m-c}) \mid X_1 = x_1, \dots, X_c = x_c) \\ &= [\phi_c(x_1, x_2, \dots, x_c)]^2, \end{aligned}$$

所以

$$\begin{aligned} & E(\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c})) \\ &= E(E(\phi(X_1, X_2, \dots, X_m) \phi(X_1, \dots, X_c, X_{m+1}, \dots, X_{2m-c}) \mid X_1, \dots, X_c)) \\ &= E((\phi_c(X_1, X_2, \dots, X_c))^2) = \sigma_c^2, \end{aligned}$$

从而 $M(c) = \binom{n}{m} \binom{m}{c} \binom{n-m}{m-c} \sigma_c^2$. 故 U 统计量的方差为

$$\begin{aligned} \text{Var}(U) &= \binom{n}{m}^{-1} \sum_{c=1}^m \binom{m}{c} \binom{n-m}{m-c} \sigma_c^2 \\ &= \sum_{c=1}^m \binom{m}{c} \frac{m!}{(m-c)!} \frac{(n-m)(n-m-1)\cdots(n-2m+c+1)}{n(n-1)\cdots(n-m+1)} \sigma_c^2. \end{aligned}$$

由于 $(n-m)(n-m+1)\cdots(n-2m+c+1)$ 和 $n(n-1)\cdots(n-m+1)$ 分别是 n 的 $m-c$ 和 m 次多项式, 所以

$$\begin{aligned}
\text{Var}(U) &= \sum_{c=1}^m \binom{m}{c} \frac{m!}{(m-c)!} [n^{-c} + O(n^{-c-1})] \sigma_c^2 \\
&= m^2 [n^{-1} + O(n^{-2})] \sigma_1^2 + \sum_{c=2}^m \binom{m}{c} \frac{m!}{(m-c)!} [n^{-c} + O(n^{-c-1})] \sigma_c^2 \\
&= \frac{m^2}{n} \sigma_1^2 + O(n^{-2}). \tag{3.76}
\end{aligned}$$

在例 3.17 中, $m=3$. 当原假设 H_0 成立时, $\theta = \frac{1}{4}$, 则 $E(U) = \theta = \frac{1}{4}$. 由式 (3.73) 及样本的独立同分布性知

$$\begin{aligned}
\phi_1(x_1) &= E(\phi(X_1, X_2, X_3) \mid X_1 = x_1) \\
&= \frac{1}{3} P(x_1 > X_2 + X_3) + \frac{2}{3} P(X_2 > X_3 + x_1) \\
&= \frac{1}{3} \int_0^{x_1} F(x_1 - x_3) dF(x_3) + \frac{2}{3} \int_0^\infty \bar{F}(x_3 + x_1) dF(x_3),
\end{aligned}$$

当原假设成立时, 由于对任意的 $s>0, t>0$, 都有 $\bar{F}(s+t) = \bar{F}(s)\bar{F}(t)$, 则有

$$F(x_1 - x_3) = 1 - \frac{\bar{F}(x_1)}{\bar{F}(x_3)}, \quad \bar{F}(x_3 + x_1) = \bar{F}(x_3)\bar{F}(x_1),$$

从而

$$\begin{aligned}
\phi_1(x_1) &= \frac{1}{3} \int_0^{x_1} \left(1 - \frac{\bar{F}(x_1)}{\bar{F}(x_3)}\right) dF(x_3) + \frac{2}{3} \int_0^\infty \bar{F}(x_3) \bar{F}(x_1) dF(x_3) \\
&= \frac{1}{3} [F(x_1) + \bar{F}(x_1) \ln \bar{F}(x_1)] + \frac{1}{3} \bar{F}(x_1) \\
&= \frac{1}{3} [1 + \bar{F}(x_1) \ln \bar{F}(x_1)],
\end{aligned}$$

下面计算 $\phi_1(X_1)$ 的方差 σ_1^2 . 由于当原假设成立时,

$$\begin{aligned}
E(\phi_1(X_1)) &= \frac{1}{4}, \\
E(\phi_1(X_1))^2 &= \frac{1}{9} \int_0^\infty [1 + \bar{F}(x_1) \ln \bar{F}(x_1)]^2 dF(x_1)
\end{aligned}$$

$$= \frac{1}{9} \int_0^1 [1 + (1-u) \ln(1-u)]^2 du = \frac{31}{486},$$

所以

$$\sigma_1^2 = \frac{31}{486} - \left(\frac{1}{4}\right)^2 = \frac{5}{3888},$$

同理(习题 3.25), 在原假设成立时,

$$\phi_2(x_1, x_2) = \begin{cases} [F(x_1 - x_2) + \bar{F}(x_1 + x_2)]/3, & x_1 > x_2, \\ [F(x_2 - x_1) + \bar{F}(x_1 + x_2)]/3, & x_2 > x_1, \end{cases}$$

$$\sigma_2^2 = \frac{7}{1296};$$

$$\phi_3(x_1, x_2, x_3) = \phi(x_1, x_2, x_3),$$

$$\sigma_3^2 = \frac{1}{48}.$$

所以

$$\begin{aligned} \text{Var}(U) &= \frac{(n-3)(n-4)}{n(n-1)(n-2)} \frac{5}{432} + \frac{n-3}{n(n-1)(n-2)} \frac{7}{72} + \frac{1}{n(n-1)(n-2)} \frac{1}{8} \\ &= \frac{5}{432n} + O(n^{-2}). \end{aligned}$$

§ 3.7.3 U 统计量的渐近正态性

定理 3.20 对由式(3.72)定义的以 $\phi(X_1, X_2, \dots, X_m)$ 为核的 U 统计量, 在核函数 $\phi(X_1, X_2, \dots, X_m)$ 的数学期望为 θ 且方差有限时,

$$\sqrt{n}(U - \theta) \xrightarrow{L} N(0, m^2 \sigma_1^2), n \rightarrow \infty,$$

其中 $\sigma_1^2 > 0$, $\sigma_c^2 (c=1, 2, \dots, m)$ 由式(3.73)和式(3.74)给出.

证明: 不妨假设 $\theta = 0$. 否则以 $\phi(X_1, X_2, \dots, X_m) - \theta$ 代替 $\phi(X_1, X_2, \dots, X_m)$.

显然, $\phi_1(X_1), \phi_2(X_2), \dots, \phi_n(X_n)$ 独立同分布, 其中函数 ϕ_1 由式(3.73)给

出. 令 $V = \frac{m}{n} \sum_{i=1}^n \phi_1(X_i)$, 则由中心极限定理知,

$$\sqrt{n}V \xrightarrow{L} N(0, m^2 \sigma_1^2), \quad n \rightarrow \infty.$$

因而,为证本定理,只需证明 $n \rightarrow \infty$ 时, $\sqrt{n}(U-V) \xrightarrow{P} 0$. 由于 $E(\sqrt{n}(U-V)) = 0$, 故只需证明 $n \rightarrow \infty$ 时, $\text{Var}(\sqrt{n}(U-V)) \rightarrow 0$. 由式(3.76)知,

$$\text{Var}(\sqrt{n}U) = n \left[\frac{m^2}{n} \sigma_1^2 + O(n^{-2}) \right] \rightarrow m^2 \sigma_1^2, \quad n \rightarrow \infty. \quad (3.77)$$

由于 $\phi_1(X_1)$ 的方差为 σ_1^2 , 所以

$$\text{Var}(\sqrt{n}V) = m^2 \sigma_1^2. \quad (3.78)$$

下面计算 $\text{Cov}(\sqrt{n}U, \sqrt{n}V)$.

$$\begin{aligned} \text{Cov}(\sqrt{n}U, \sqrt{n}V) &= E(U \cdot nV) = m \sum_{i=1}^n E(U \phi_1(X_i)) \\ &= mn E(U \phi_1(X_1)) \\ &= mn \binom{n}{m}^{-1} \sum E(\phi(X_{i_1}, X_{i_2}, \dots, X_{i_m}) \phi_1(X_1)), \end{aligned}$$

其中 $1 \leq i_1 < i_2 < \dots < i_m \leq n$, $\sum$ 表示对所有的组合 $(i_1, i_2, \dots, i_m)$ 求和. 若 $i_1 > 1$, 则 $E(\phi(X_{i_1}, X_{i_2}, \dots, X_{i_m}) \phi_1(X_1)) = 0$. 从而

$$\text{Cov}(\sqrt{n}U, \sqrt{n}V) = mn \binom{n}{m}^{-1} \sum E(\phi(X_1, X_{i_2}, \dots, X_{i_m}) \phi_1(X_1)),$$

其中 $1 < i_2 < \dots < i_m \leq n$, $\sum$ 表示对所有的组合 $(i_2, i_3, \dots, i_m)$ 求和. 因此

$$\text{Cov}(\sqrt{n}U, \sqrt{n}V) = mn \binom{n}{m}^{-1} \binom{n-1}{m-1} E(\phi(X_1, X_2, \dots, X_m) \phi_1(X_1)).$$

在 $X_1 = x_1$ 给定的条件下, 计算 $\phi(X_1, X_2, \dots, X_m) \phi_1(X_1)$ 的条件期望

$$E\{\phi(X_1, X_2, \dots, X_m) \phi_1(X_1) \mid X_1 = x_1\} = [\phi_1(X_1)]^2,$$

则

$$\text{Cov}(\sqrt{n}U, \sqrt{n}V) = mn \binom{n}{m}^{-1} \binom{n-1}{m-1} \sigma_1^2 = m^2 \sigma_1^2. \quad (3.79)$$

从而由式(3.77)、式(3.78)和式(3.79)知, 在 $n \rightarrow \infty$ 时,

$$\text{Var}(\sqrt{n}(U - V)) = \text{Var}(\sqrt{n}U) + \text{Var}(\sqrt{n}V) - 2\text{Cov}(\sqrt{n}U, \sqrt{n}V) \rightarrow 0,$$

定理证毕.

在例 3.17, $m=3, \sigma_1^2 = \frac{5}{3888}$, 在原假设 H_0 成立时,

$$\sqrt{n}\left(U - \frac{1}{4}\right) \xrightarrow{L} N\left(0, \frac{5}{432}\right),$$

故在大样本场合下检验问题 (3.68) 的拒绝域为

$$U \leq \frac{1}{4} + \sqrt{\frac{5}{432n}} U_\alpha,$$

其中 U_α 为标准正态分布 $N(0,1)$ 的 α 分位数.

例 3.18 (对称中心的检验) 设总体 X 的分布函数为 $F(x-\theta)$, 其中 $F(x)$ 是关于原点对称的连续型分布. 由此得到了一个非参数结构. 考虑检验问题

$$\text{原假设 } H_0: \theta = 0 \quad \text{对} \quad \text{备择假设 } H_1: \theta > 0, \quad (3.80)$$

即检验对称分布的对称中心在原点, 还是在原点的右边. 设 $X_1, X_2, \dots, X_n$ 是来自总体 X 的一个样本. 若 X 取参数结构, 例如服从正态分布 $N(\mu, \sigma^2)$, 那么上述检验问题就是关于均值 μ 的检验问题, 常用 t 检验法. 在 X 取非参数结构时, 注意到

$$\text{原假设 } H_0: \theta = 0 \text{ 成立} \Rightarrow P(X_1 + X_2 > 0) = 0.5;$$

$$\text{备择假设 } H_1: \theta > 0 \text{ 成立} \Rightarrow P(X_1 + X_2 > 0) > 0.5.$$

因而令 $\omega = P(X_1 + X_2 > 0)$. 那么, 上述检验问题可简化为关于参数 ω 的检验问题,

$$\text{原假设 } H_0: \omega = 0.5 \quad \text{对} \quad \text{备择假设 } H_1: \omega > 0.5. \quad (3.81)$$

显然,

$$\phi(X_1, X_2) = \begin{cases} 1, & X_1 + X_2 > 0, \\ 0, & X_1 + X_2 \leq 0 \end{cases}$$

是 ω 的一个无偏估计. 然后由对称函数 $\phi(X_1, X_2)$ 构造 U 统计量,

$$U = U(X_1, X_2, \dots, X_n) = \frac{\sum \phi(X_{i_1}, X_{i_2})}{\binom{n}{2}},$$

其中 $1 \leq i_1 < i_2 \leq n$, Σ 表示对所有的组合 (i_1, i_2) 求和. U 也是 ω 的一个无偏估计. 由于在备择假设成立时, ω 的值较大, 所以我们在 $U(X_1, X_2, \dots, X_n) \geq c$ 时, 拒绝原假设.

这里 $m=2$. 在原假设 H_0 为真时, $E(U(X_1, X_2, \dots, X_n)) = 0.5$, 并且不难推得 (习题 3.26),

$$\phi_1(x_1) = F(x_1), \quad \sigma_1^2 = \frac{1}{12},$$

$$\phi_2(x_1, x_2) = \phi(x_1, x_2), \quad \sigma_2^2 = \frac{1}{4},$$

所以

$$\text{Var}(U(X_1, X_2, \dots, X_n)) = \frac{2n-1}{6n(n-1)} = \frac{1}{3n} + O(n^{-2}).$$

在原假设成立时, $\sqrt{n}\left(U - \frac{1}{2}\right) \xrightarrow{L} N\left(0, \frac{1}{3}\right), n \rightarrow \infty$. 故在大样本场合下检验问题 (3.80) 的拒绝域为

$$U \geq \frac{1}{2} + \sqrt{\frac{1}{3n}} U_{1-\alpha}.$$

§ 3.7.4 两样本 U 统计量

定义 3.12 样本 $X_1, X_2, \dots, X_{n_1}$ 和 $Y_1, Y_2, \dots, Y_{n_2}$ 分别来自相互独立的总体 X 和 Y . 设在 $X_1, X_2, \dots, X_{m_1}$ 给定时, $\phi(X_1, X_2, \dots, X_{m_1}; Y_1, Y_2, \dots, Y_{m_2})$ 为 $Y_1, Y_2, \dots, Y_{m_2}$ 的对称函数, 而在 $Y_1, Y_2, \dots, Y_{m_2}$ 给定时, $\phi(X_1, X_2, \dots, X_{m_1}; Y_1, Y_2, \dots, Y_{m_2})$ 为 $X_1, X_2, \dots, X_{m_1}$ 的对称函数. 令

$$\begin{aligned} U &= U(X_1, X_2, \dots, X_{n_1}; Y_1, Y_2, \dots, Y_{n_2}) \\ &= \binom{n_1}{m_1}^{-1} \binom{n_2}{m_2}^{-1} \sum \phi(X_{i_1}, X_{i_2}, \dots, X_{i_{m_1}}; Y_{j_1}, Y_{j_2}, \dots, Y_{j_{m_2}}), \end{aligned} \quad (3.82)$$

其中 $1 \leq i_1 < \dots < i_{m_1} \leq n_1, 1 \leq j_1 < \dots < j_{m_2} \leq n_2$, Σ 表示对所有的组合 $(i_1, i_2, \dots, i_{m_1})$ 和 $(j_1, j_2, \dots, j_{m_2})$ 求和, 则称 U 是以 $\phi(X_1, X_2, \dots, X_{m_1}; Y_1, Y_2, \dots, Y_{m_2})$ 为核的两样本 U 统计量.

设 $E(\phi(X_1, X_2, \dots, X_{m_1}; Y_1, Y_2, \dots, Y_{m_2})) = \theta$, 则 U 统计量的数学期望

$$E(U) = \theta.$$

与单样本的情况相类似地,令

$$\begin{aligned} & \phi_{c_1 c_2}(x_1, x_2, \cdots, x_{c_1}; y_1, y_2, \cdots, y_{c_2}) \\ &= E(\phi(X_1, X_2, \cdots, X_{m_1}; Y_1, Y_2, \cdots, Y_{m_2}) | X_1 = x_1, \cdots, X_{c_1} = x_{c_1}; Y_1 = y_1, \cdots, Y_{c_2} = y_{c_2}), \\ & \sigma_{c_1 c_2}^2 = \text{Var}(\phi_{c_1 c_2}(X_1, X_2, \cdots, X_{c_1}; Y_1, Y_2, \cdots, Y_{c_2})), \end{aligned} \quad (3.83)$$

其中 $c_1 = 0, 1, \cdots, m_1; c_2 = 0, 1, \cdots, m_2$. 注意: 在 $c_1 = c_2 = 0$ 时,

$$\begin{aligned} \phi_{00}(x_1, x_2, \cdots, x_{c_1}; y_1, y_2, \cdots, y_{c_2}) &= E(\phi(X_1, X_2, \cdots, X_{m_1}; Y_1, Y_2, \cdots, Y_{m_2})) = \theta, \\ \sigma_{00} &= 0. \end{aligned}$$

若 $\phi(X_1, X_2, \cdots, X_{m_1}; Y_1, Y_2, \cdots, Y_{m_2})$ 的方差有限, 则 $\sigma_{c_1 c_2}^2$ 也有限. 与单样本的情况相类似地, 可以得到 U 统计量的方差, 它等于

$$\begin{aligned} \text{Var}(U) &= \binom{n_1}{m_1}^{-1} \binom{n_2}{m_2}^{-1} \sum_{c_1=0}^{m_1} \sum_{c_2=0}^{m_2} \binom{m_1}{c_1} \binom{n_1 - m_1}{m_1 - c_1} \binom{m_2}{c_2} \binom{n_2 - m_2}{m_2 - c_2} \sigma_{c_1 c_2}^2 \\ &= \frac{m_1^2}{n_1} \sigma_{10}^2 + \frac{m_2^2}{n_2} \sigma_{01}^2 + O(n_1^{-2} + n_2^{-2}). \end{aligned}$$

定理 3.21 对由式(3.82)定义的以 $\phi(X_1, X_2, \cdots, X_{m_1}; Y_1, Y_2, \cdots, Y_{m_2})$ 为核的两样本 U 统计量, 在核函数 $\phi(X_1, X_2, \cdots, X_{m_1}; Y_1, Y_2, \cdots, Y_{m_2})$ 的数学期望为 θ 且方差有限时,

$$\frac{U - \theta}{\sqrt{\frac{m_1^2 \sigma_{10}^2}{n_1} + \frac{m_2^2 \sigma_{01}^2}{n_2}}} \xrightarrow{L} N(0, 1), \quad n_1 \rightarrow \infty, \quad n_2 \rightarrow \infty,$$

其中 $\sigma_{10}^2 > 0, \sigma_{01}^2 > 0, \sigma_{c_1 c_2}^2 (c_1 = 0, 1, \cdots, m_1; c_2 = 0, 1, \cdots, m_2)$ 由式(3.83)给出.

证明与单样本的情况相类似^[5].

例 3.19 (位置参数的 Mann-Whitney (曼-惠特尼) 检验) 样本 $X_1, X_2, \cdots, X_{n_1}$ 和 $Y_1, Y_2, \cdots, Y_{n_2}$ 分别来自相互独立的连续随机变量总体 X 和 Y . 设 X 和 Y 的分布函数分别为 F 和 G . 考虑检验问题, 原假设 H_0 : 对任意的 x , 都有 $F(x) = G(x)$, 备择假设 H_1 : 对任意的 x , 都有 $F(x) > G(x)$.

在 X 和 Y 服从方差相等的正态分布时, 这就是比较两个正态分布的均值的

检验问题,常用 t 检验法.

在 $G(x) = F(x-a)$ 时,这就是关于位置参数 a 的检验问题,原假设即变为 $H_0: a=0$,而备择假设即变为 $H_1: a>0$. 这个检验问题常称为位置参数的检验问题,受人关注.

在原假设成立,即对任意的 x ,都有 $F(x) = G(x)$ 时,

$$P(X < Y) = \int_{-\infty}^{\infty} F(y) dG(y) = \int_0^1 u du = 0.5. \quad (3.84)$$

而在备择假设为真,即对任意的 x ,都有 $F(x) > G(x)$ 时,

$$P(X < Y) = \int_{-\infty}^{\infty} F(y) dG(y) > \int_{-\infty}^{\infty} G(y) dG(y) = \int_0^1 u du = 0.5. \quad (3.85)$$

在 $P(X < Y) > 0.5$ 时,这时我们称 X 随机地比 Y 小,记为“ $X \overset{r}{<} Y$ ”. 由式(3.84)和式(3.85)知,我们令 $\omega = P(X < Y)$. 那么上述检验问题可简化为关于参数 ω 的检验问题,原假设 $H_0: \omega = 0.5$ 对备择假设 $H_1: \omega > 0.5$. 显然,以

$$\phi(X_1, Y_1) = \begin{cases} 1, & X_1 < Y_1, \\ 0, & X_1 \geq Y_1 \end{cases}$$

为核构造的 U 统计量,

$$U = U(X_1, \dots, X_{n_1}; Y_1, \dots, Y_{n_2}) = \frac{\sum \phi(X_i; Y_j)}{n_1 n_2} \quad (3.86)$$

是 ω 的无偏估计,其中 $1 \leq i \leq n_1, 1 \leq j \leq n_2$, $\sum$ 表示对所有的 i 和 j 求和. 该 U 统计量是 Mann 和 Whitney 在 1947 年提出的,人们称它为 Mann-Whitney 统计量. 由于在备择假设为真时, ω 的值较大,所以我们在 $U \geq c$ 时,拒绝原假设,认为“对任意的 x ,都有 $F(x) > G(x)$ ”.

在原假设成立时, $\omega = 0.5$, 所以 $E(U) = 0.5$. 同时还有(习题 3.27)

$$\phi_{10}(x_1) = G(x_1) = F(x_1), \quad \phi_{01}(y_1) = \bar{F}(y_1), \quad \phi_{11}(x_1, y_1) = \phi(x_1, y_1);$$

$$\sigma_{10}^2 = \sigma_{01}^2 = \frac{1}{12}, \quad \sigma_{11}^2 = \frac{1}{4},$$

所以

$$\text{Var}(U) = \frac{n_1 + n_2 + 1}{12n_1 n_2}.$$

据定理 3.21 知,在 $n_1 \rightarrow \infty, n_2 \rightarrow \infty$ 时,

$$\frac{\sqrt{12n_1n_2}}{\sqrt{n_1+n_2}} \frac{U-0.5}{\sqrt{n_1+n_2}} \xrightarrow{L} N(0,1), \quad (3.87)$$

故在大样本场合下检验的拒绝域为

$$U \geq \frac{1}{2} + \sqrt{\frac{n_1+n_2}{12n_1n_2}} U_{1-\alpha},$$

这个检验称为 Mann-Whitney 检验,或位置参数的 Mann-Whitney 检验.

§ 3.8 秩 检 验

除了 U 统计量检验方法外,秩检验方法也是非常重要的非参数结构的检验方法.秩方法始于 1904 年 C. Spearman(斯皮尔曼)提出的秩相关方法,但是秩方法受到人们越来越大的重视是在 1945 年 F. Wilcoxon(威尔科克森)提出了秩和检验方法之后.秩和检验方法用于两个总体位置参数的比较问题.它在应用上有很大的意义,同时它还极大地推动了后来秩方法的发展.

秩和检验以及其他的秩方法都是建立在秩与秩统计量基础上的非参数方法.下面首先介绍有关秩与秩统计量的基本概念和基本结论.

§ 3.8.1 秩

设有独立同分布的样本 $x_1, x_2, \dots, x_n$. 不妨假设总体 X 是连续型随机变量,从而以概率 1 保证,样本 $x_1, x_2, \dots, x_n$ 互不相等.将样本由小到大排列成 $x_{(1)} < x_{(2)} < \dots < x_{(n)}$.若 $x_i = x_{(R_i)}$,则称 $x_i (i=1, 2, \dots, n)$ 在 $x_1, x_2, \dots, x_n$ 中的秩为 R_i ,简称 x_i 的秩为 $R_i, R_i=1, 2, \dots, n$. $\mathbf{R}=(R_1, R_2, \dots, R_n)$ 以及 $\mathbf{R}$ 任意的可测函数称为秩统计量.

$\mathbf{R}$ 服从离散型分布,它取 $n!$ 个值.由于样本 $x_1, x_2, \dots, x_n$ 独立同分布,所以 $\mathbf{R}$ 取任意一组值 $(r_1, r_2, \dots, r_n)$ 的概率都是 $1/n!$,其中 $(r_1, r_2, \dots, r_n)$ 是 $1, 2, \dots, n$ 的任意一个排列.这说明 $\mathbf{R}$ 服从均匀分布.由此可见,秩统计量的分布和总体服从什么样的分布是没有关系的.这就是秩方法之所以是非参数方法的缘由.

由于 $\mathbf{R}$ 服从均匀分布,所以单个的秩 $R_i (i=1, 2, \dots, n)$ 也服从均匀分布:

$$P(R_i = r) = \frac{1}{n}, \quad r = 1, 2, \dots, n,$$

从而我们有

定理 3.22 对任意的 $i=1, 2, \dots, n$, 都有

$$E(R_i) = \frac{n+1}{2}, \quad \text{Var}(R_i) = \frac{n^2-1}{12},$$

同单个的秩 R_i, R_i 和 $R_j (i \neq j)$ 的联合分布也是均匀分布:

$$P(R_i = r_1, R_j = r_2) = \frac{1}{n(n-1)}, \quad r_1 \neq r_2,$$

从而我们有

定理 3.23 对任意的 $1 \leq i < j \leq n$, 都有

$$\text{Cov}(R_i, R_j) = -\frac{n+1}{12}.$$

定理 3.22 和 3.23 的证明留作习题 3.29.

秩方法的基本思想是, 用 x_i 的秩 R_i 代替 x_i 作统计推断. R_i 可理解为 x_i 的得分. 一般地说, 我们可以定义一个计分函数 $a(r), r=1, 2, \dots, n$, 在 x_i 的秩为 R_i 时, 将 x_i 的得分定义为 $a(R_i)$. 和定理 3.22 和 3.23 相类似地, 我们有定理 3.24 和 3.25.

定理 3.24 对任意的 $i=1, 2, \dots, n$, 都有

$$E(a(R_i)) = \bar{a},$$

$$\text{Var}(a(R_i)) = \frac{1}{n} \sum_{i=1}^n (a(i) - \bar{a})^2,$$

其中 $\bar{a} = \frac{1}{n} \sum_{i=1}^n a(i)$.

定理 3.25 对任意的 $1 \leq i < j \leq n$, 都有

$$\text{Cov}(a(R_i), a(R_j)) = -\frac{1}{n(n-1)} \sum_{i=1}^n (a(i) - \bar{a})^2.$$

定理 3.24 和 3.25 的证明见习题 3.30.

秩方法的内容非常丰富. 本章就符号秩和检验与秩和检验这两个内容对秩方法作一简单介绍. 对秩方法, 乃至非参数方法有兴趣, 并且有志于研究这部分内容的人士要认真学习陈希孺院士等的专著《非参数统计》.

§ 3.8.2 符号秩和检验

在例 3.18 我们用 U 统计量的方法研究对称中心的检验问题. 这个问题也可以用秩方法来解决. 假设 $x_1, x_2, \dots, x_n$ 独立同分布, 其分布关于 θ 对称. 不妨假

设这个对称分布是连续分布. 由此知, 以概率 1 不仅能保证 $x_1, x_2, \dots, x_n$ 两两互不相等, 而且任何一个 $x_i (i = 1, 2, \dots, n)$ 都不等于 θ . 考虑对称中心 θ 的检验问题:

$$H_0: \theta = 0 \quad \text{对} \quad H_1: \theta < 0, \quad (3.88)$$

令

$$u_i = \begin{cases} 1, & x_i > 0, \\ 0, & \text{其他}, \end{cases} \quad i = 1, 2, \dots, n,$$

$$S^+ = \sum_{i=1}^n u_i = \# \{x_i > 0, i = 1, 2, \dots, n\}, \quad (3.89)$$

其中记号“#”表示计数. 显然, 我们可以在 S^+ 比较小的时候拒绝原假设, 认为 $\theta < 0$. 检验的 p 值为 $P(Z \leq S^+)$, 其中 $Z \sim b(n, 1/2)$, 以后, 我们将把诸如此类的 p 值计算公式简写为 $P(b(n, 1/2) \leq S^+)$. 这个检验方法称为符号检验. 符号检验仅仅使用样本数据是正数还是负数的信息, 而没有使用样本数据值的大小的信息. 它并不能有效地解决对称中心 θ 是否等于 0 的检验问题. 下面所述的符号秩和检验方法就是符号检验的改进.

符号秩和检验方法之所以是符号检验的改进, 就是因为它还用到了样本数据值的大小的信息. 我们首先将每一个样本都取绝对值 $|x_1|, |x_2|, \dots, |x_n|$. 由于总体服从连续分布, 我们同样能够以概率 1 保证 $|x_1|, |x_2|, \dots, |x_n|$ 两两也互不相等. 设 $|x_i| (i = 1, 2, \dots, n)$ 在 $|x_1|, |x_2|, \dots, |x_n|$ 中的秩为 R_i , $R_i = 1, 2, \dots, n$. 符号秩和检验的检验统计量为

$$W^+ = \sum_{i=1}^n u_i R_i. \quad (3.90)$$

显然, W^+ 不仅有样本数据是正数还是负数的信息, 而且还有样本数据值的大小的信息. 和符号检验相类似, 我们在 W^+ 比较小的时候拒绝原假设, 认为 $\theta < 0$. 为了求检验的 p 值和检验的临界值, 下面研究原假设 $H_0: \theta = 0$ 成立, 即总体的分布关于原点 0 对称时, 符号秩和检验统计量 W^+ 的分布的性质.

首先叙述关于原点 0 对称的连续分布的一个结论.

定理 3.26 假设连续随机变量 X 的分布关于原点 0 对称. 令

$$u = \begin{cases} 1, & X > 0, \\ 0, & \text{其他}, \end{cases}$$

则 $|X|$ 和 u 相互独立.

证明: u 的分布律为 $P(u=1)=P(u=0)=1/2$, 所以欲证明 $|X|$ 和 u 相互独立, 仅需证明下面两个等式对任意的 $a>0$ 都成立:

$$P(|X| \leq a, u=1) = P(|X| \leq a)P(u=1),$$

$$P(|X| \leq a, u=0) = P(|X| \leq a)P(u=0).$$

我们只证明前一个等式, 后一个等式可类似地证明. 由于

$$P(|X| \leq a, u=1) = P(|X| \leq a, X > 0) = P(0 < X \leq a),$$

$$P(|X| \leq a) = 2P(0 < X \leq a),$$

$$P(u=1) = \frac{1}{2},$$

故前一个等式得证. 定理 3.26 证毕.

由定理 3.26 知, 在样本 $x_1, x_2, \dots, x_n$ 独立同分布, 其分布是关于原点 0 对称的连续分布时, $|x_1|, u_1, |x_2|, u_2, \dots, |x_n|, u_n$ 相互独立. 我们知道 $u_1, u_2, \dots, u_n$ 之间相互独立, 而 $R_1, R_2, \dots, R_n$ 之间并不相互独立. 但是在总体 X 的分布是关于原点 0 对称的连续分布时, $u_1, u_2, \dots, u_n$ 和 $R_1, R_2, \dots, R_n$ 相互独立. 从而知, 不论 $R_1, R_2, \dots, R_n$ 是否给定, $u_1, u_2, \dots, u_n$ 总是独立同分布的, 而且 u_1 的分布总是 $P(u_1=0)=P(u_1=1)=0.5$.

W^+ 服从离散分布, 它可能取 $0, 1, 2, \dots, \frac{n(n+1)}{2}$ 各个值. 令 $t_{n,d}$ 表示从 $1, 2, \dots, n$ 这 n 个数中任取若干个数 (包括一个都不取), 其和恰为 d 的取法种数, 其中 $d=0, 1, 2, \dots, \frac{n(n+1)}{2}$. 例如, $t_{n,0}=t_{n,1}=t_{n,2}=1$; 因为 $3=1+2, 4=1+3$, 所以 $t_{n,3}=t_{n,4}=2$; 因为 $5=1+4=2+3$, 所以 $t_{n,5}=3$ 等. 则在总体 X 的分布是关于原点 0 对称的连续分布时, $R_1, R_2, \dots, R_n$ 给定后 $W^+=d$ 的条件概率:

$$P(W^+=d | R_1, R_2, \dots, R_n) = P\left(\sum_{i=1}^n u_i R_i = d \mid R_1, R_2, \dots, R_n\right) = \frac{t_{n,d}}{2^n},$$

$$d = 0, 1, 2, \dots, \frac{n(n+1)}{2}.$$

这个条件概率和 $R_1, R_2, \dots, R_n$ 没有关系. 这说明在 $R_1, R_2, \dots, R_n$ 给定后 $W^+=d$ 的条件概率就是 $W^+=d$ 的概率. 从而有如下定理.

定理 3.27 在总体的分布是关于原点 0 对称的连续分布时, W^+ 的概率分

布为

$$P(W^+ = d) = P\left(\sum_{i=1}^n u_i R_i = d\right) = \frac{t_{n,d}}{2^n}, \quad d = 0, 1, 2, \dots, \frac{n(n+1)}{2}. \quad (3.91)$$

检验问题(3.88)的 p 值为符号秩和检验统计量的值小于等于 W^+ 的概率. 由定理3.27知, p 值等于

$$\sum_{d \leq W^+} \frac{t_{n,d}}{2^n}.$$

在 n 不是很大时,可由符号秩和检验 W^+ 的累积概率表(参见文献[7]附录表4)查得 p 值.

如果从 $1, 2, \dots, n$ 这 n 个数中任取若干个数(包括一个都不取),其和恰为 d ,那么没有取到剩下来的数的和为 $n(n+1)/2-d$,所以从 $1, 2, \dots, n$ 这 n 个数中任取若干个数(包括一个都不取),其和恰为 d 的方法种数 $t_{n,d}$ 等于从 $1, 2, \dots, n$ 这 n 个数中任取若干个数(包括一个都不取),其和恰为 $n(n+1)/2-d$ 的方法种数 $t_{n, \frac{n(n+1)}{2}-d}$. 由此知,

$$P(W^+ = d) = P\left(W^+ = \frac{n(n+1)}{2} - d\right)$$

对所有的 $d=0, 1, 2, \dots$ 都成立. 这说明, W^+ 有如下定理.

定理 3.28 在总体的分布是关于原点0对称的连续分布时, W^+ 服从对称分布,对称中心为 $0, 1, 2, \dots, n(n+1)/2$ 的中点 $n(n+1)/4$.

如果所考虑的对称中心 θ 的检验问题为

$$H_0: \theta = 0 \quad \text{对} \quad H_1: \theta > 0, \quad (3.92)$$

则我们在 W^+ 比较大的时候拒绝原假设,认为 $\theta > 0$. 该检验问题的 p 值为符号秩和检验统计量的值大于等于 W^+ 的概率. 由定理3.27和3.28知, p 值等于

$$\sum_{d \geq W^+} \frac{t_{n,d}}{2^n} = \sum_{d \leq n(n+1)/2 - W^+} \frac{t_{n,d}}{2^n}.$$

在 n 不是很大时,由符号秩和检验 W^+ 的累积概率表可查得该检验问题的 p 值.

下面计算在总体的分布是关于原点0对称的连续分布时, W^+ 的期望和方差. 由定理3.26,以及定理3.22和3.23知,

$$E(W^+) = E\left(\sum_{i=1}^n u_i R_i\right) = nE(u_1 R_1)$$

$$= nE(u_1)E(R_1) = n \frac{1}{2} \frac{n+1}{2} = \frac{n(n+1)}{4}.$$

实际上,由定理 3.28 立即可以推得 W^+ 的期望 $E(W^+) = \frac{n(n+1)}{4}$,

$$\text{Var}(W^+) = n\text{Var}(u_1 R_1) + n(n-1)\text{Cov}(u_1 R_1, u_2 R_2).$$

由于

$$\begin{aligned}\text{Var}(u_1 R_1) &= E(u_1^2 R_1^2) - (E(u_1 R_1))^2 = E(u_1^2)E(R_1^2) - \left(\frac{n+1}{4}\right)^2 \\ &= \frac{1}{2} \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{16} = \frac{(n+1)(5n+1)}{48},\end{aligned}$$

$$\begin{aligned}\text{Cov}(u_1 R_1, u_2 R_2) &= E(u_1 u_2 R_1 R_2) - E(u_1 R_1)E(u_2 R_2) \\ &= E(u_1)E(u_2)E(R_1 R_2) - \left(\frac{n+1}{4}\right)^2 \\ &= \left(\frac{1}{2}\right)^2 \frac{\sum_{i \neq j} ij}{n(n-1)} - \frac{(n+1)^2}{16} \\ &= \frac{1}{4} \frac{(1 + \cdots + n)^2 - (1^2 + \cdots + n^2)}{n(n-1)} - \frac{(n+1)^2}{16} \\ &= -\frac{n+1}{48},\end{aligned}$$

所以

$$\begin{aligned}\text{Var}(W^+) &= n \frac{(n+1)(5n+1)}{48} + n(n-1) \frac{-n-1}{48} \\ &= \frac{n(n+1)(2n+1)}{24},\end{aligned}$$

从而 W^+ 有下述的性质.

定理 3.29 在总体的分布是关于原点 0 对称的连续分布时,

$$E(W^+) = \frac{n(n+1)}{4},$$

$$\text{Var}(W^+) = \frac{n(n+1)(2n+1)}{24}.$$

用下面的方法计算 W^+ 的期望和方差比较简单. 令 $S = \sum_{i=1}^n iu_i$. 同 W^+ 一样, S 也仅取 $0, 1, 2, \dots, n(n+1)/2$ 各个值, 并且

$$P(S = d) = P\left(\sum_{i=1}^n iu_i = d\right) = \frac{t_{n,d}}{2^n}, \quad d = 0, 1, 2, \dots, \frac{n(n+1)}{2}. \quad (3.93)$$

由式(3.91)和式(3.93)知, W^+ 和 S 有相同的概率分布律, 记为

$$W^+ \stackrel{d}{=} S, \quad (3.94)$$

符号“ $\stackrel{d}{=}$ ”表示等式两边的这两个随机变量同分布. 所以 W^+ 和 S 有相同的均值和方差.

W^+ 和随机变量 $u_1, u_2, \dots, u_n$ 和 $R_1, R_2, \dots, R_n$ 都有关系, 而 S 仅和随机变量 $u_1, u_2, \dots, u_n$ 有关系, 计算 S 的均值和方差比计算 W^+ 的容易得多. 由于 $E(u_i) = 1/2$, $\text{Var}(u_i) = 1/4$, $i = 1, 2, \dots, n$, 所以

$$E(S) = \frac{1}{2} \sum_{i=1}^n i = \frac{n(n+1)}{4},$$

$$\text{Var}(S) = \frac{1}{4} \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{24},$$

由此立即得到定理 3.29.

在 n 比较大时, 我们无法利用 W^+ 的累积概率表计算检验的 p 值. 这时, 需要应用 W^+ 的渐近分布. 下面我们将利用独立随机变量和的 Lyapunov (李雅普诺夫) 中心极限定理, 证明 S 有渐近正态性. 然后由式(3.94), 得到 W^+ 的渐近正态性.

独立随机变量和的 Lyapunov 中心极限定理如下.

Lyapunov 定理 设 $y_1, y_2, \dots$ 为独立随机变量序列, $E(y_i) = a_i$, $\text{Var}(y_i) = \sigma_i^2$, $i = 1, 2, \dots$. 令 $S_n = \sum_{i=1}^n y_i$, 则 $E(S_n) = \sum_{i=1}^n a_i$, $\text{Var}(S_n) = B_n^2$, 其中 $B_n^2 = \sum_{i=1}^n \sigma_i^2$.

将 S_n 标准化, 即令 $V_n = \frac{S_n - \sum_{i=1}^n a_i}{B_n}$, 则 $E(V_n) = 0$, $\text{Var}(V_n) = 1$. 如果当 $n \rightarrow \infty$ 时,

$$\frac{\sum_{i=1}^n E|y_i - a_i|^3}{B_n^{3/2}} \rightarrow 0,$$

则当 $n \rightarrow \infty$ 时, $V_n \xrightarrow{L} N(0, 1)$.

下面利用 Lyapunov 定理证明 S 的渐近正态性. 令 $y_i = iu_i, i = 1, 2, \dots$, 则 $y_1, y_2, \dots$ 为独立随机变量序列, $S = \sum_{i=1}^n y_i$. 由于 $E(y_i) = a_i, a_i = \frac{i}{2}, \text{Var}(y_i) = \sigma_i^2, \sigma_i^2 = \frac{i^2}{4}, i = 1, 2, \dots$, 所以 $E(S) = \frac{n(n+1)}{4}, \text{Var}(S) = B_n^2, B_n^2 = \frac{n(n+1)(2n+1)}{24}$. 由于

$$E \left| y_i - \frac{i}{2} \right|^3 = i^3 E \left| u_i - \frac{1}{2} \right|^3 = \frac{i^3}{8},$$

$$1^3 + 2^3 + \dots + n^3 = (1 + 2 + \dots + n)^2 = \frac{n^2(n+1)^2}{4},$$

所以当 $n \rightarrow \infty$ 时,

$$\frac{\sum_{i=1}^n E |y_i - a_i|^3}{B_n^{3/2}} = \frac{n^2(n+1)^2/32}{[n(n+1)(2n+1)/24]^{3/2}} \rightarrow 0,$$

则由独立随机变量和的 Lyapunov 中心极限定理知, S 有渐近正态性: 在样本容量 n 趋于无穷大时,

$$\frac{S - E(S)}{\sqrt{\text{Var}(S)}} = \frac{S - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \xrightarrow{L} N(0, 1),$$

从而由式(3.94)知, W^+ 也有渐近正态性.

定理 3.30 如果总体的分布是关于原点 0 对称的连续分布, 则在样本容量 n 趋于无穷大时, W^+ 有渐近正态性:

$$\frac{W^+ - E(W^+)}{\sqrt{\text{Var}(W^+)}} = \frac{W^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \xrightarrow{L} N(0, 1).$$

由定理 3.30, 在大样本场合, 检验问题(3.88)的 p 值为

$$\begin{aligned} P\left(N\left(\frac{n(n+1)}{4}, \frac{n(n+1)(2n+1)}{24}\right) \leq W^+\right) \\ = \Phi\left(\frac{W^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}\right), \end{aligned}$$

而检验问题(3.92)的 p 值为

$$P\left(N\left(\frac{n(n+1)}{4}, \frac{n(n+1)(2n+1)}{24}\right) \geq W^+\right)$$

$$= \Phi\left(\frac{n(n+1)/4 - W^+}{\sqrt{n(n+1)(2n+1)/24}}\right).$$

一般地说,我们可以定义一个计分函数 $a(r)$, $r=1, 2, \dots, n$, 在 x_i 的秩为 R_i 时, 将 x_i 的得分定义为 $a(R_i)$. 从而将符号秩和检验统计量 $W^+ = \sum_{i=1}^n u_i R_i$ 推广为 $\sum_{i=1}^n u_i a(R_i)$. 为了将这个推广后的符号秩和检验统计量用于检验问题 (3.88) 和 (3.90), 计分函数 $a(r)$ 必须非增, 或非减. 计分函数 $a(r)$ 除了取为 $a(r)=r$ 得到符号秩和检验统计量外, 还有以下一些特例:

(1) 取 $a(r)=1, r=1, 2, \dots, n$, 得符号检验 $S^+ = \sum_{i=1}^m u_i = \#\{x_i > 0, i=1, 2, \dots, n\}$;

(2) 假设 $\xi_1, \xi_2, \dots, \xi_n$ 为来自标准正态分布 $N(0, 1)$ 的容量为 n 的样本, 将 $|\xi_1|, |\xi_2|, \dots, |\xi_n|$ 由小到大排列成 $|\xi|_{(1)} < |\xi|_{(2)} < \dots < |\xi|_{(n)}$. 取 $a(r) = E(|\xi|_{(r)})$, $r=1, 2, \dots, n$. 这个计分函数称为正态期望计分.

(3) 记标准正态分布 $N(0, 1)$ 的分布函数为 $\Phi(x)$, 其反函数记为 $\Phi^{-1}(x)$. 取 $a(r) = \Phi^{-1}\{1/2 + r/[2(n+1)]\}$, $r=1, 2, \dots, n$. 这个计分函数称为正态分位计分.

在总体的分布是关于原点 0 对称的连续分布时, 和式 (3.94) 相类似地, 可以证明: $\sum_{i=1}^n u_i a(R_i) \stackrel{d}{=} \sum_{i=1}^n u_i a(i)$. 据此, 可以求得 $\sum_{i=1}^n u_i a(R_i)$ 的期望和方差, 以及它们有渐近正态性的一个充分条件.

§ 3.8.3 位置参数的秩和检验

在例 3.19, 我们用 U 统计量的方法解位置参数的检验问题. 这个 U 统计量检验方法称为 Mann-Whitney 检验. 位置参数的检验问题也可以用秩方法来解决. 这个秩方法称为 Wilcoxon 秩和检验.

样本 $x_1, x_2, \dots, x_m$ 和 $y_1, y_2, \dots, y_n$ 分别来自相互独立的连续随机变量总体 X 和 Y , 分别记 X 和 Y 的分布函数为 $F(x)$ 和 $G(y)$. 假设 $G(y) = F(y-a)$, 其中 $-\infty < a < \infty$. 在 $a > 0$ 时, 由于对任意的 c 都有 $F(c) = G(c+a)$, 所以将 X 的分布函数 $F(x)$ 向右平行移动 a 个单位即得到 Y 的分布函数 $G(y)$. 类似地, 在 $a < 0$ 时, 将 X 的分布函数 $F(x)$ 向左平行移动 a 个单位即得到 Y 的分布函数 $G(y)$. a 称为位置参数.

由于 X 的分布函数为 $F(x)$, 所以 $X+a$ 的分布函数为 $F(x-a)$. 这说明 $X+a$ 和 Y 有相同的分布函数, $X+a \stackrel{d}{=} Y$. 因而在 $a > 0$ 时, 如果将合样本

$(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n)$ 由小到大排列, $y_1, y_2, \dots, y_n$ 倾向于排在右边. 若记 $y_j (j=1, 2, \dots, n)$ 在合样本中的秩为 $R_j (R_j=1, 2, \dots, N, N=m+n)$, 则在 $a>0$ 时, Y 样本 $y_1, y_2, \dots, y_n$ 的秩和

$$W_y = \sum_{j=1}^n R_j \quad (3.95)$$

倾向于取比较大的值. 考虑位置参数 a 的检验问题:

$$H_0: a = 0 \quad \text{对} \quad H_1: a > 0. \quad (3.96)$$

除了在例 3.19, 用 Mann-Whitney U 统计量方法外, 还可以用秩方法解决这个检验问题. 秩方法取 Y 样本的秩和 W_y 作为检验统计量, 在 W_y 比较大的时候拒绝原假设, 认为 $a>0$. 我们也可以取 X 样本的秩和 W_x 作为检验统计量, 在 W_x 比较小的时候拒绝原假设, 认为 $a>0$. 由于 $W_x + W_y = 1 + 2 + \dots + N = N(N+1)/2$, 所以这两个检验方法是等价的. 考虑位置参数 a 的另一个检验问题:

$$H_0: a = 0 \quad \text{对} \quad H_1: a < 0. \quad (3.97)$$

同样取 Y 样本的秩和 W_y 作为检验统计量, 在 W_y 比较小的时候拒绝原假设, 认为 $a<0$. 这个检验方法就是著名的 Wilcoxon 秩和检验.

下面研究原假设 $H_0: a=0$ 成立, $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, 检验统计量 W_y 的分布的性质.

W_y 服从离散分布, 它依次取 $n(n+1)/2, n(n+1)/2+1, \dots, n(n+1)/2+mn$, 令 $t_{m,n}(d)$ 表示从 $1, 2, \dots, N=m+n$ 这 N 个数中任取 n 个数, 其和恰为 d 的方法种数. 例如, $t_{m,n}(n(n+1)/2) = t_{m,n}(n(n+1)/2+1) = 1, t_{m,n}(n(n+1)/2+2) = 2$, 等等.

定理 3.31 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, W_y 的概率分布为

$$P(W_y = d) = \frac{t_{m,n}(d)}{\binom{N}{n}}, \quad d = \frac{n(n+1)}{2}, \dots, \frac{n(n+1)}{2} + mn.$$

由定理 3.31, 检验问题(3.96)和(3.97)的 p 值分别为

$$\sum_{d \geq W^+} \frac{t_{m,n}(d)}{\binom{N}{n}} \quad \text{和} \quad \sum_{d \leq W^+} \frac{t_{m,n}(d)}{\binom{N}{n}}.$$

假设从 $1, 2, \dots, N$ 这 N 个数中取出的 n 个数为 $a_1, a_2, \dots, a_n$, 其和 $\sum_{i=1}^n a_i = d$. 若

令 $b_i = (N+1) - a_i, i = 1, 2, \dots, n$. 显然 $1 \leq b_i \leq N, i = 1, 2, \dots, n$, 并且它们的和为 $\sum_{i=1}^n b_i = n(N+1) - d$. 这说明从 $1, 2, \dots, N$ 这 N 个数中任取 n 个数, 其和恰为 d 与其和恰为 $n(N+1) - d$ 的方法种数相等; $t_{m,n}(d) = t_{m,n}(n(N+1) - d)$. 故由定理 3.31 推得定理 3.32.

定理 3.32 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, W_y 服从对称分布, 对称中心为 $n(n+1)/2, n(n+1)/2+1, \dots, n(n+1)/2+mn$ 的中点 $n(N+1)/2$.

由定理 3.31 和 3.32, 人们构造了 Wilcoxon 秩和检验统计量分位点表 (见孙山泽著《非参数统计讲义》(附表 4)). 查表可以得到检验的 p 值和临界值.

一般地说, 我们可以定义一个计分函数 $a(r), r = 1, 2, \dots, N$, 在 y_i 的秩为 R_i 时, 将 y_i 的得分定义为 $a(R_i)$. 从而将秩和检验统计量 $W_y = \sum_{i=1}^n R_i$ 推广为 $\sum_{i=1}^n a(R_i)$. 为了将这个推广后的秩和检验统计量用于检验问题 (3.96) 和 (3.97), 计分函数 $a(r)$ 必须非增或非减. 计分函数 $a(r)$ 除了取为 $a(r) = r$ 得到 Wilcoxon 秩和检验统计量外, 还有以下一些特例:

(1) Fisher-Yates (费希尔-耶茨) 检验: 假设 $\xi_1, \xi_2, \dots, \xi_N$ 为来自标准正态分布 $N(0, 1)$ 的容量为 n 的样本, 将 $\xi_1, \xi_2, \dots, \xi_N$ 由小到大排列成 $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(N)}$, 取 $a(r) = E(\xi_{(r)})$. 这个计分函数是 Fisher 和 Yates 于 1938 年提出的.

(2) Van der Waerden (范德瓦尔登) 检验: 记标准正态分布 $N(0, 1)$ 的分布函数为 $\Phi(x)$, 其反函数记为 $\Phi^{-1}(x)$. 取 $a(r) = \Phi^{-1}(r/(N+1)), r = 1, 2, \dots, N$. 这个计分函数是 Van der Waerden 于 1952 年提出的.

下面计算在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, $\sum_{i=1}^n a(R_i)$ 的期望和方差. 由定理 3.24 和 3.25 知

$$E\left(\sum_{i=1}^n a(R_i)\right) = nE(a(R_1)) = n\bar{a}, \quad \bar{a} = \frac{1}{N} \sum_{r=1}^N a(r), \quad (3.98)$$

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n a(R_i)\right) &= n\text{Var}(a(R_1)) + n(n-1)\text{Cov}(a(R_1), a(R_2)) \\ &= n \frac{\sum_{r=1}^N (a(r) - \bar{a})^2}{N} - n(n-1) \frac{\sum_{r=1}^N (a(r) - \bar{a})^2}{N(N-1)} \\ &= \frac{mn}{N(N-1)} \sum_{r=1}^N (a(r) - \bar{a})^2. \end{aligned} \quad (3.99)$$

在 $a(r)=r$ 时, 因为 $\bar{a}=(N+1)/2$, $\sum_{r=1}^N (a(r)-\bar{a})^2 = N(N^2-1)/12$, 故由式 (3.98) 和式 (3.99) 得到秩和检验统计量 W_y 的期望和方差. 从而有下述的性质:

定理 3.33 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, 秩和检验统计量 W_y 的期望和方差分别为

$$E(W_y) = n(N+1)/2,$$

$$\text{Var}(W_y) = mn(N+1)/12.$$

在例 3.19 我们用 U 统计量的方法解位置参数的检验问题. 这个 U 统计量为

$$U(x_1, x_2, \dots, x_n) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \phi(x_i, y_j),$$

$$\phi(x_i, y_j) = \begin{cases} 1, & x_i < y_j, \\ 0, & \text{其他.} \end{cases}$$

可以证明

$$\sum_{i=1}^m \sum_{j=1}^n \phi(x_i, y_j) = W_y - \frac{n(n+1)}{2}.$$

这说明例 3.19 的 U 统计量与秩和检验统计量 W_y (见式 (3.95)) 有着某种关系. 利用这个关系我们可以在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, 由 U 统计量的期望和方差 (见例 3.19) 计算 W_y 的期望和方差. 由于例 3.19 的 U 统计量具有渐近正态性, 所以秩和检验统计量 W_y 也具有渐近正态性. 这就是如下的定理.

定理 3.34 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, 在样本容量 m 和 n 都趋于无穷大时, W_y 有渐近正态性:

$$\frac{W_y - E(W_y)}{\sqrt{\text{Var}(W_y)}} = \frac{W_y - n(N+1)/2}{\sqrt{nm(N+1)/12}} \xrightarrow{L} N(0, 1).$$

由定理 3.34, 在大样本场合, 检验问题 (3.96) 的 p 值为

$$\begin{aligned} & P(N(n(N+1)/2, mn(N+1)/12) \geq W_y) \\ &= \Phi\left(\frac{n(N+1)/2 - W_y}{\sqrt{mn(N+1)/12}}\right). \end{aligned}$$

而检验问题 (3.97) 的 p 值为

$$P(N(n(N+1)/2, mn(N+1)/12) \leq W_y) \\ = \Phi\left(\frac{W_y - n(N+1)/2}{\sqrt{mn(N+1)/12}}\right).$$

计分函数 $a(r)$ 满足的条件和所考虑的检验问题的类型有关. 在上面, 我们考虑位置参数的检验问题, 计分函数 $a(r)$ 必须非增或非减, 而如果考虑尺度参数的检验问题, 计分函数 $a(r)$ 就要求满足另外的条件了. 作为位置参数检验问题的补充, 下面简单讨论尺度参数的检验问题.

§ 3.8.4 尺度参数的秩检验

设样本 $x_1, x_2, \dots, x_m$ 和 $y_1, y_2, \dots, y_n$ 分别来自相互独立的连续随机变量总体 X 和 Y . X 和 Y 的取值范围相同, 都是以原点 0 为对称中心的一个区间, 分别记 X 和 Y 的分布函数为 $F(x)$ 和 $G(y)$. 假设 $G(y) = F(y/\sigma)$, 其中 $\sigma > 0$. 不难证明: $Y \stackrel{d}{=} \sigma X$, σ 称为尺度参数. 在 $\sigma > 1$ 时, 由于对任意的 $c > 0$ 都有 $\sigma c > c$, 而对任意的 $c < 0$ 都有 $\sigma c < c$, 所以若将合样本 $(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n)$ 由小到大排列, $y_1, y_2, \dots, y_n$ 倾向于排在左右两边. 若记 $y_j (j=1, 2, \dots, n)$ 在合样本中的秩为 R_j ($R_j=1, 2, \dots, N, N=m+n$), 则在 $\sigma > 1$ 时, R_j 的值倾向于比较大或比较小. 如果考虑尺度参数 σ 的检验问题:

$$H_0: \sigma = 1 \quad \text{对} \quad H_1: \sigma > 1,$$

并取 $\sum_{i=1}^n a(R_i)$ 为检验统计量, 则计分函数 $a(r)$ 必须单峰或单谷. 随着 r 从 1, 2 增加到 N , $a(r)$ 先上升然后下降, 或者先下降然后上升. 尺度参数检验问题的计分函数 $a(r)$ 的选取, 通常有以下 5 种方法:

(1) Mood(穆德)检验: $a(r) = (r - (N+1)/2)^2$, $r=1, 2, \dots, N$.

(2) Ansari-Bradley(安塞利-布拉德利)检验:

$$N=2k \text{ 为偶数时, } a(r) = \begin{cases} r, & r=1, 2, \dots, k, \\ N-r+1, & r=k+1, k+2, \dots, N; \end{cases}$$

$$N=2k+1 \text{ 为奇数时, } a(r) = \begin{cases} r, & r=1, 2, \dots, k, k+1, \\ N-r+1, & r=k+2, k+3, \dots, N. \end{cases}$$

(3) Copan(科潘)检验: 假设 $\xi_1, \xi_2, \dots, \xi_N$ 为来自标准正态分布 $N(0, 1)$ 的容量为 n 的样本, 将 $\xi_1, \xi_2, \dots, \xi_N$ 由小到大排列成 $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(N)}$, 取 $a(r) = E(\xi_{(r)}^2)$, $r=1, 2, \dots, N$.

(4) Klotz(克洛茨)检验: 记标准正态分布 $N(0, 1)$ 的分布函数为 $\Phi(x)$, 其

反函数记为 $\Phi^{-1}(x)$, 取 $a(r) = [\Phi^{-1}(r/(N+1))]^2, r=1, 2, \dots, N$.

(5) Siegel-Turkey (西格尔-图基) 检验: 取 $a(1) = N, a(N) = N-1, a(N-1) = N-2, a(2) = N-3, a(3) = N-4, a(N-2) = N-5, a(N-3) = N-6, a(4) = N-7, a(5) = N-8, \dots$.

根据式(3.98)和式(3.99), 可以在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, 计算上述这些检验统计量的期望和方差. 这里特别要指出的是, 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同为连续分布时, Siegel-Turkey 检验统计量和 Wilcoxon 秩和检验统计量有相同的分布, 它们有相同的期望和方差, 以及相同的渐近正态性. 在应用 Siegel-Turkey 检验统计量解尺度参数检验问题时, 查 Wilcoxon 秩和检验统计量分位点表, 可以得到 Siegel-Turkey 检验的 p 值和临界值.

§ 3.8.5 线性秩统计量

最后讨论位置参数和尺度参数检验问题中所使用过的检验统计量的渐近正态性问题.

假设 $x_1, x_2, \dots, x_N$ 同为连续分布, 记 $x_i (i=1, 2, \dots, N)$ 在 $x_1, x_2, \dots, x_N$ 中的秩为 $R_i, R_i=1, 2, \dots, N$, 则

$$L = \sum_{i=1}^N c(i) a(R_i)$$

称为线性秩统计量, 其中 $c(i)$ 是定义在 $\{1, 2, \dots, N\}$ 上的某个特定的函数, 又称为回归系数, $a(r), r=1, 2, \dots, N$ 是计分函数. 上述有关位置参数和尺度参数的各个检验统计量都属于线性秩统计量. 譬如, 将两样本 $x_1, x_2, \dots, x_m$ 和 $y_1, y_2, \dots, y_n$ 合在一起, 并将 $y_1, y_2, \dots, y_n$ 视为 $x_{m+1}, x_{m+2}, \dots, x_N$. 如果在 $i=1, 2, \dots, m$ 时取回归系数 $c(i)=0$, 而在 $i=m+1, m+2, \dots, N$ 时, 取回归系数 $c(i)=1$, 并取计分函数

$a(r)=r, r=1, 2, \dots, N$, 则有 $L = \sum_{i=1}^N c(i) a(R_i) = \sum_{j=1}^n R_j$. 这就是式(3.95)的 Wilcoxon 秩和检验统计量 W_y .

下面计算线性秩统计量的期望和方差. 记

$$\bar{c} = \frac{1}{N} \sum_{i=1}^N c(i), \quad \bar{a} = \frac{1}{N} \sum_{r=1}^N a(r),$$

则由定理 3.24 和 3.25 知

$$E(L) = E\left(\sum_{i=1}^N c(i) a(R_i)\right) = \bar{a} \sum_{i=1}^N c(i) = N \bar{a} \bar{c},$$

$$\begin{aligned}\text{Var}(L) &= \text{Var} \left(\sum_{i=1}^N c(i) a(R_i) \right) \\ &= \frac{\sum_{r=1}^N (a(r) - \bar{a})^2}{N} \sum_{i=1}^N c^2(i) - \frac{\sum_{r=1}^N (a(r) - \bar{a})^2}{N(N-1)} \sum_{i \neq j} c(i) c(j),\end{aligned}$$

由于

$$\sum_{i \neq j} c(i) c(j) = (N\bar{c})^2 - \sum_{i=1}^N c^2(i),$$

所以

$$\text{Var}(L) = \frac{1}{N-1} \sum_{r=1}^N (a(r) - \bar{a})^2 \sum_{i=1}^N (c(i) - \bar{c})^2.$$

线性秩统计量的渐近正态性经过了很多统计学家的共同努力,最终得到了著名的 Hajek (哈耶克) 定理. 下面简要介绍线性秩统计量的渐近正态性.

条件 N 若

$$\frac{\max_{1 \leq i \leq N} (c(i) - \bar{c})^2}{N \sum_{i=1}^N (c(i) - \bar{c})^2} \rightarrow 0, \quad N \rightarrow \infty,$$

则称回归系数 $\{c(i)\}$ 满足条件 N. 条件 N 是 Noether 于 1949 年提出的. 同样, 若

$$\frac{\max_{1 \leq i \leq N} (a(i) - \bar{a})^2}{N \sum_{i=1}^N (a(i) - \bar{a})^2} \rightarrow 0, \quad N \rightarrow \infty,$$

则称计分函数 $\{a(i)\}$ 满足条件 N.

条件 M 令

$$d_{ij} = |(c(i) - \bar{c})(a(j) - \bar{a})|, \quad i, j = 1, 2, \dots, N,$$

$$d^2 = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N d_{ij}^2 = N \sum_{i=1}^N (c^2(i) - \bar{c})^2 \sum_{i=1}^N (a^2(i) - \bar{a})^2.$$

如果对任意的 $\varepsilon > 0$, 都有

$$\frac{\sum |d_{ij}| \geq \varepsilon d^2}{d^2} \rightarrow 0, \quad N \rightarrow \infty,$$

则称回归系数和计分函数 $\{c(i), a(i)\}$ 满足条件 M. 条件 M 是 Motoo (木村) 于

1955 年提出的.

可以证明: 若回归系数和计分函数 $\{c(i), a(i)\}$ 满足条件 M, 则它们都满足条件 N. 这个结论的逆不真.

Hajek 定理 假设 $x_1, x_2, \dots, x_N$ 同为连续分布. 记 $x_i (i=1, 2, \dots, N)$ 在 $x_1, x_2, \dots, x_N$ 中的秩为 R_i . 假设线性秩统计量 $L = \sum_{i=1}^N c(i) a(R_i)$ 的回归系数 $\{c(i)\}$ 和计分函数 $\{a(i)\}$ 都满足条件 N, 则在样本容量 $N \rightarrow \infty$ 时,

$$\frac{L - E(L)}{\sqrt{\text{Var}(L)}} = \frac{L - N\bar{a}\bar{c}}{\sqrt{\sum_{i=1}^N (c(i) - \bar{c})^2 \sum_{i=1}^N (a(i) - \bar{a})^2 / (N-1)}} \xrightarrow{L} N(0, 1)$$

的充分条件为, 回归系数和计分函数 $\{c(i), a(i)\}$ 满足条件 M.

这个定理是 Hajek 于 1966 年证明的. 条件 M 的验证比较复杂, 为此人们引入平方可积计分函数. 如果计分函数 $a(i)$ 具有下述的形式:

$$a(i) = b(N) \varphi\left(\frac{i}{N+1}\right), \quad i = 1, 2, \dots, N,$$

其中 $b(N)$ 仅与 N 有关, 函数 $\varphi(t)$ 是 $(0, 1)$ 上两个非降平方可积函数之差, 且 $\varphi(t)$ 不等于某一常数, 则称该计分函数 $a(i)$ 为平方可积计分函数.

可以证明, 如果回归系数 $\{c(i)\}$ 满足条件 N, 计分函数 $a(i)$ 为平方可积计分函数, 则回归系数和计分函数 $\{c(i), a(i)\}$ 满足条件 M. 由此可见, 平方可积计分函数必满足条件 N.

Hajek 定理的推广 假设 $x_1, x_2, \dots, x_N$ 同为连续分布. 记 $x_i (i=1, 2, \dots, N)$ 在 $x_1, x_2, \dots, x_N$ 中的秩为 R_i . 如果线性秩统计量 $L = \sum_{i=1}^N c(i) a(R_i)$ 的回归系数 $c(i)$ 满足条件 N, 且计分函数 $a(i)$ 为平方可积计分函数, 则

$$\frac{L - E(L)}{\sqrt{\text{Var}(L)}} = \frac{L - N\bar{a}\bar{c}}{\sqrt{\sum_{i=1}^N (c(i) - \bar{c})^2 \sum_{i=1}^N (a(i) - \bar{a})^2 / (N-1)}} \xrightarrow{L} N(0, 1).$$

通常使用的计分函数往往都是平方可积计分函数. 例如, Wilcoxon 秩和检验统计量 W_y 的计分函数 $a(i) = i, i=1, 2, \dots, N$, 它具有形式 $a(i) = (N+1) \varphi(i/(N+1))$, 其中 $\varphi(t) = t, 0 < t < 1$. $\varphi(t)$ 是平方可积计分函数. 由 Hajek 定理的推广, 不难验证 Wilcoxon 秩和检验统计量 W_y 的渐近正态性.

§ 3.9 置换检验

除了 U 统计量检验与秩检验方法,置换检验是又一个非参数结构的检验方法. 置换检验的详细介绍见本章参考文献[5]与[8].

§ 3.9.1 置换检验的引进

早期人们使用的是机械式打字机. 20 世纪 70 年代后期出现的电脑打字机,很快就取代了机械式打字机,80 年代电脑打字机风靡世界. 没多久,随着个人计算机(PC)技术的发展,越来越多的人用个人计算机打字. 用电脑打字机打字和用个人计算机打字,哪一个的打字速度快? 现在几乎没有人用电脑打字机打字,大家都说用个人计算机打字快. 但在 30 多年前个人计算机刚普及时,这的确是个问题. 表 3.4 是三位秘书分别用电脑打字机和个人计算机的打字速度(字数/min). 这批数据能否说明用个人计算机的打字速度快?

表 3.4 有重复观察值的 2 个处理 3 个区组的区组设计 单位:字数/min

打字机类别(处理)	秘书(区组)			合计
	1	2	3	
1. 电脑打字机	72,58	59,52	52,51,59	403
2. 个人计算机	76,66	57,60	65,57,64	445

不同的秘书的打字速度可能有所不同,有的甚至有很大的差别. 例如,由表 3.4 可以看到,第 1 位秘书不论用电脑打字机,还是用个人计算机,打字的速度都比第 2、3 位秘书快一些. 表 3.4 是一个有重复观察值的 $k=2$ 个处理和 $b=3$ 个区组的区组设计. 若有正态分布假定,则可用方差分析的 F 检验去求解. 若难以对分布形式作出假定,则可考虑使用非参数的置换检验方法.

表 3.4 的假设检验问题的原假设为“电脑打字机和个人计算机打字一样快”,备择假设为“个人计算机打字快”. 本章 § 3.5,求解多参数指数型分布族的假设检验问题使用的是条件检验方法. 置换检验其实也是条件检验. 我们是在区组 1 的 4 个观察值给定为 72,58,76 与 66,区组 2 的 4 个观察值给定为 59,52,57 与 60,区组 3 的 6 个观察值给定为 52,51,59,65,57 与 64 的条件下,求解表 3.4 的假设检验问题. 为简化起见,特记上述条件为“条件 H”. 由于条件 H 仅给定区组 1 的 4 个观察值为 72,58,76 与 66,并没有说在处理 1 与 2 里分别是哪两个观察值. 若原假设成立,电脑打字机和个人计算机打字一样快,则区组 1 的处理 1

与处理2里的观察值可能是观察到的“72,58”与“76,66”,也有可能是“72,76”与“58,66”等. 相互置换,共有 $C_4^2=6$ 种不同的置换方法. 类似地,区组2的4个观察值也有6种不同的置换方法. 区组3的6个观察值一分为二,在处理1与2里分别有3个,共有 $C_6^3=20$ 种不同的置换方法. 总而言之,表3.4有 $6 \times 6 \times 20 = 720$ 种不同的置换方法. 在原假设成立,电脑打字机和个人计算机打字一样快时,每一个置换出现的可能性都等于 $1/720$,服从均匀分布. 不论原有的是什么样的分布,置换总是服从均匀分布,这就是置换检验之所以是非参数方法的缘由.

这个区组设计的检验问题只有两个处理,并且同一个区组的不同处理的重复观察次数相等,因而可用处理1的7个观察值的和 x_{1+} 作为检验统计量:

$$x_{1+} = x_{111} + x_{112} + x_{121} + x_{122} + x_{131} + x_{132} + x_{133} = 403,$$

其中 x_{ijk} 表示第 i 个处理第 j 个区组的第 k 个观察值,在 x_{1+} 比较小的时候拒绝原假设,认为处理1(即用电脑打字机)打字慢,处理2(即用个人计算机)打字快. 根据现有观察值算得 $x_{1+}=403$,它是不是一个比较小的数,小到可以认为用个人计算机打字快. 这就需要计算 $x_{1+}=403$ 的 p 值,也就是计算概率 $P(x_{1+} \leq 403)$. 置换检验的 p 值,其实是条件 p 值,是在条件 H 下,原假设成立时,计算 x_{1+} 小于等于 403 的概率.

在条件 H 下,表3.4有720种不同的置换方法. 每一个置换都有 x_{1+} 的一个计算结果,共有720个计算结果. 原假设成立时, x_{1+} 的这720个计算结果服从均匀分布,每一个结果出现的概率都是 $\frac{1}{720}$. 经逐一置换搜索排序, x_{1+} 小于等于 403 的有35个. 所以 $x_{1+}=403$ 的 p 值为

$$P(x_{1+} \leq 403) = \frac{35}{720} \approx 0.0486.$$

若取显著性水平 $\alpha=0.05$,则我们认为用电脑打字机打字慢,而用个人计算机打字快.

显然,这种类型的置换检验的 p 值与观察到的数值有关,因而无法构造检验的临界值表. 下面介绍的区组设计的 Friedman(弗里德曼)检验也是置换检验,但它基于秩方法,其 p 值仅与处理数 k 和区组数 b 有关,因而可以构造它的检验的临界值表.

§ 3.9.2 Friedman 检验

为比较 A, B, C, D 和 E 5 种药物注射后产生的皮肤疱疹的大小,选取 6 只家兔并给每只家兔先后分别按随机排列的次序注射这 5 种药物. 试验结果(疱疹面积单位: mm^2) 见表 3.5. 这批数据能否说明这 5 种药物注射后产生的皮肤疱

疹大小有差异?

表 3.5 家兔注射药物后产生的皮肤疱疹的面积 单位: mm²

药物类别(处理)	家兔(区组)					
	1	2	3	4	5	6
1	73	73	67	61	69	79
2	83	81	99	82	82	87
3	74	60	76	77	68	76
4	58	64	64	71	77	74
5	77	75	73	59	85	82

表 3.5 是 $k=5$ 个处理和 $b=6$ 个区组的区组设计. Friedman 检验将同一个区组内的 $k=5$ 个样本观察值由小到大排列, 然后求出观察值在区组内的秩. 把表 3.5 中的每一个数都用它在区组内的秩代替, 从而得到表 3.6.

表 3.6 家兔注射药物后产生的皮肤疱疹面积观察值的秩 单位: mm²

药物类别(处理)	家兔(区组)						和	均值
	1	2	3	4	5	6		
1	2	3	2	2	2	3	14	$\bar{R}_1 \approx 2.33$
2	5	5	5	5	4	5	29	$\bar{R}_2 \approx 4.83$
3	3	1	4	4	1	2	15	$\bar{R}_3 = 2.50$
4	1	2	1	3	3	1	11	$\bar{R}_4 \approx 1.83$
5	4	4	3	1	5	4	21	$\bar{R}_5 = 3.50$

对于有 3 个或更多个处理的区组设计, 我们用“处理间平方和”作为检验统计量. 表 3.6 的处理间平方和

$$\begin{aligned}
 SS &= b \sum_{i=1}^k (\bar{R}_i - \bar{R})^2 \\
 &= 6 (2.33-3)^2 + 6 (4.83-3)^2 + 6 (2.5-3)^2 + 6 (1.83-3)^2 + 6 (3.5-3)^2 \\
 &= 34,
 \end{aligned}$$

其中 $\bar{R}_i$ 是第 i ($i=1, 2, \dots, 5$) 个处理的秩的平均, $\bar{R} = (k+1)/2 = 3$ 是秩的总平均. 显然, 在处理间平方和 SS 的值比较大的时候, 我们认为这 5 种药物注射后产生的皮肤疱疹大小有差异. 现在的问题是, $SS = 34$ 是不是比较大. 这就需要计算

SS=34 的 p 值,也就是在 5 种药物注射后产生的皮肤疱疹大小无差异时,计算概率 $P(SS \geq 34)$. Friedman 检验的 p 值,也是条件 p 值,是在表 3.6 各个区组的秩都给定为 1,2,3,4 与 5 的条件下,计算 SS=34 的 p 值.由此可见,Friedman 检验的 p 值仅与处理数 k 和区组数 b 有关,因而可以构造它的检验临界值表,见参考文献[7]与[9].

由于仅给定区组的 5 个秩为 1,2,3,4 与 5,并没有说它是如何排列的.例如区组 1 这 5 个秩原有的排列(从处理 1 依次到处理 5)为 2,5,3,1 与 4,但也有可能置换成另外一种排列.区组 1 的置换有 $5! = 120$ 个.6 个区组总的来说有 $(5!)^6 = 2.99 \times 10^{12}$ 个置换.每一个置换都有 SS 的一个计算结果.对如此多的置换,逐一置换搜索排序极其困难,但也不是不可能的. Friedman 检验临界值表列有 5 个处理 6 个区组的 $\alpha = 0.05$ 与 $\alpha = 0.01$ 的临界值.若处理过多,或区组过多,在检验临界值表中不能查到临界值,这时就需要使用渐近分布近似地计算 p 值.下面讨论 Friedman 检验统计量的渐近分布问题.

Friedman 检验的一般提法如下:设有一个 k 个处理 b 个区组的区组设计.假设第 i 个处理第 j 个区组的观察值 x_{ij} 的分布函数为 $F_j(x - \theta_i)$, $i = 1, 2, \dots, k, j = 1, 2, \dots, b$. 这也就是说,在同一个区组里,不同处理的观察值的分布函数的形状都相同,仅有可能的是位置参数不同.而不同区组内的观察值的分布函数的形状有可能不同. Friedman 检验用于这 k 个处理的效应有没有差异,也就是位置参数 $\theta_1, \theta_2, \dots, \theta_k$ 是否全都相等的检验问题.它的原假设和备择假设分别为

$$H_0: \theta_1 = \theta_2 = \dots = \theta_k, \quad H_1: \theta_1, \theta_2, \dots, \theta_k \text{ 不全相等.}$$

假设对每一个 $i = 1, 2, \dots, k$ 与每一个 $j = 1, 2, \dots, b$, 第 i 个处理第 j 个区组都只有一个观察值 x_{ij} , 故共有 $N = bk$ 个观察值. 假设所有的 x_{ij} 互相独立, 且同一个区组的 k 个观察值互不相等. 记 x_{ij} 在第 j 个区组的 k 个观察值 $x_{1j}, \dots, x_{kj}$ 中的秩为 R_{ij} , 则有 $R_{ij} = 1, 2, \dots, k$. Friedman 检验是在各个区组的秩都给定为 1, 2, $\dots, k$ 时的置换检验. 这 k 个处理之间的平方和为

$$SS = b \sum_{i=1}^k (\bar{R}_i - \bar{R})^2,$$

其中

$$\begin{aligned} \bar{R}_i &= \sum_{j=1}^b R_{ij}/b \text{ 是第 } i \text{ 个处理的 } b \text{ 个观察值的秩的平均, } i = 1, 2, \dots, k, \\ \bar{R} &= \sum_{i=1}^k \sum_{j=1}^b R_{ij}/N = b(1 + 2 + \dots + k)/N = (k + 1)/2 \text{ 是总的平均, 它是个常数.} \end{aligned}$$

将 SS 变换为

$$Q = \frac{12}{k(k+1)} SS = \frac{12b}{k(k+1)} \sum_{i=1}^k (\bar{R}_i - (k+1)/2)^2,$$

这个 Q 可以作为检验统计量. 我们在 Q 的值比较大的时候拒绝原假设 H_0 , 认为这 k 个位置参数 $\theta_1, \theta_2, \dots, \theta_k$ 不全都相等, 处理有差异. Friedman 检验临界值表就是以 Q 作为检验统计量而制作的.

接下来证明, 在原假设 H_0 为真时, 这 k 个位置参数全都相等, $\theta_1 = \theta_2 = \dots = \theta_k$ 时, 若 k 固定, 而 $b \rightarrow \infty$, 则 Friedman 置换检验的检验统计量 Q 渐近服从 $\chi^2(k-1)$ 分布, 即

$$Q \xrightarrow{L} \chi^2(k-1).$$

处理效应无差异, $\theta_1 = \theta_2 = \dots = \theta_k$ 时, 对每一个 $j = 1, 2, \dots, b$, $(R_{1j}, R_{2j}, \dots, R_{kj})$ 服从均匀分布

$$P(R_{1j} = a_1, \dots, R_{kj} = a_k) = 1/k!,$$

$(a_1, a_2, \dots, a_k)$ 是 $(1, 2, \dots, k)$ 的任意一个排列. 则由定理 3.22 与 3.23 知,

$$E(R_{ij}) = (k+1)/2,$$

$$\text{Cov}(R_{ij}, R_{sj}) = \begin{cases} (k^2 - 1)/12, & i = s, \\ -(k+1)/12, & i \neq s. \end{cases}$$

记第 j 个区组前 $k-1$ 个秩构成的向量为 $\mathbf{Y}_j$,

$$\mathbf{Y}_j = \begin{pmatrix} R_{1j} \\ \vdots \\ R_{(k-1)j} \end{pmatrix},$$

则 $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_b$ 独立同分布, 且

$$E(\mathbf{Y}_1) = \begin{pmatrix} (k+1)/2 \\ \vdots \\ (k+1)/2 \end{pmatrix},$$

$$\text{Cov}(\mathbf{Y}_1) = \boldsymbol{\Sigma} = (\sigma_{ij}), \text{ 其中 } \sigma_{ij} = \begin{cases} (k^2 - 1)/12, & i = j, \\ -(k+1)/12, & i \neq j. \end{cases}$$

故由中心极限定理知

$$\mathbf{Z} = \sqrt{b} \begin{pmatrix} \bar{R}_1 - (k+1)/2 \\ \vdots \\ \bar{R}_{k-1} - (k+1)/2 \end{pmatrix} \xrightarrow{L} N_{k-1}(\mathbf{0}, \boldsymbol{\Sigma}), b \rightarrow \infty \text{ 时},$$

从而有

$$\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z} \xrightarrow{L} \chi^2(k-1),$$

由于

$$\Sigma = \left(-\frac{k+1}{12} \right) \mathbf{a}\mathbf{a}^T + \frac{k(k+1)}{12} \mathbf{I},$$

其中 $\mathbf{a}$ 是元素全为数字 1 的 $k-1$ 维向量, $\mathbf{I}$ 是单位矩阵. 由此不难算得

$$\Sigma^{-1} = \frac{12}{k(k+1)} (\mathbf{a}\mathbf{a}^T + \mathbf{I}) = (\rho_{ij}),$$

其中 $\rho_{ij} = \begin{cases} 24/k(k+1), & i=j, \\ 12/k(k+1), & i \neq j. \end{cases}$ 因而

$$\mathbf{Z}^T \Sigma^{-1} \mathbf{Z} = \frac{12}{k(k+1)} \mathbf{Z}^T (\mathbf{a}\mathbf{a}^T + \mathbf{I}) \mathbf{Z} = \frac{12}{k(k+1)} ((\mathbf{Z}^T \mathbf{a})^2 + \mathbf{Z}^T \mathbf{Z}),$$

由于 $\bar{R}_1 + \cdots + \bar{R}_{k-1} + \bar{R}_k = k(k+1)/2$, 所以

$$\begin{aligned} \mathbf{Z}^T \mathbf{a} &= \sqrt{b} (\bar{R}_1 + \cdots + \bar{R}_{k-1} - (k-1)(k+1)/2) = \sqrt{b} ((k+1)/2 - \bar{R}_k), \\ (\mathbf{Z}^T \mathbf{a})^2 &= b (\bar{R}_k - (k+1)/2)^2, \end{aligned}$$

显然 $\mathbf{Z}^T \mathbf{Z} = b \sum_{i=1}^{k-1} (\bar{R}_i - \bar{R})^2$, 所以 $(\mathbf{Z}^T \mathbf{a})^2 + \mathbf{Z}^T \mathbf{Z}$ 等于处理间平方和 SS,

$$\mathbf{Z}^T \Sigma^{-1} \mathbf{Z} = \frac{12}{k(k+1)} ((\mathbf{Z}^T \mathbf{a})^2 + \mathbf{Z}^T \mathbf{Z}) = \frac{12}{k(k+1)} \text{SS} = Q.$$

这就证明了所要的结果: $Q \xrightarrow{L} \chi^2(k-1)$. 注: 为什么用第 j 个区组的前 $k-1$ 个秩, 而不用全部 k 个秩构成向量 $\mathbf{Y}_j$?

上述那个药物注射后产生的皮肤疱疹的大小有没有差异的区组设计, 算得的处理间平方和 SS 为 34. 由此算得 $Q = 13.6$. 查 Friedman 检验临界值表, 得到 $k=5$ 个处理, $b=6$ 个区组时的显著性水平 $\alpha=0.01$ 的临界值为 11.867, 所以根据已有数据我们认为这 5 种药物注射后产生的皮肤疱疹大小有差异. 我们也可以使用 Q 渐近服从 $\chi^2(k-1)$ 分布的性质计算 p 值. 由于 $k=5$, 从而算得检验的 p 值为 $P(\chi^2(4) \geq 13.6) = 0.009$.

这个区组设计的处理效应有没有差异的问题还可以用 Hodges-Lehmann (霍奇斯-莱曼) 检验方法去求解. 它也是基于秩方法的置换检验, 其详细介绍见参考文献 [7] 或 [9].

与 Friedman 检验相类似地, § 3.8.3 的位置参数的秩和检验与 § 3.8.4 的尺度参数的秩检验都可视为置换检验.

§ 3.9.3 C 检验

区组设计的处理, 除了有没有差异的检验问题, 还有趋势性的检验问题. 看下面的例子.

某药物有镇痛效果. 一般来说, 镇痛药物的剂量越大, 镇痛时间就越长. 为检

验该药物有没有这样的效果,取它的 6 个剂量:1 mg,2 mg,3 mg,4 mg,5 mg 和 6 mg 做试验.有 9 位患者参加试验.对每一位患者分别先后按随机排列的次序使用该药物的这 6 个剂量镇痛,镇痛时间(单位:min)见表 3.7.这是一个 $k=6$ 个处理, $b=9$ 个区组的区组设计,检验处理的效应有没有上升的趋势.这个问题可以用 Page(佩奇)检验方法去求解.它也是基于秩方法的置换检验,其详细介绍见参考文献[7]与[9].下面介绍如何用 C 检验的方法去求解这个问题.有关 C 检验的详细介绍可参阅王静龙和李贵军于 2008 年发表在《应用概率统计》上的文章:《完全区组设计下基于 Aligned 秩的有方向检验问题》.C 检验的名称取自于陈希孺院士的姓的首字母.

表 3.7 某药物的剂量和镇痛时间

剂量/mg	患者的镇痛时间/min								
	1	2	3	4	5	6	7	8	9
1	36	62	53	105	36	118	42	51	114
2	51	91	81	63	46	65	108	63	51
3	71	40	67	49	62	126	123	55	30
4	63	51	75	65	63	96	32	86	109
5	82	33	116	107	42	122	69	41	97
6	128	81	38	33	104	112	102	121	86
均值 $\bar{x}_j$	71.8	59.7	71.7	70.3	58.8	106.5	79.3	69.5	81.2

各个区组的观察值的均值 $\bar{x}_j$ 见表 3.7 的最后一行.由此可以看到,不同患者的镇痛时间有长有短,这里面有区组效应.同 Hodges-Lehmann 检验,C 检验按下面的方法消除区组效应.将观察值 x_{ij} 调整为 $x'_{ij} = x_{ij} - \bar{x}_j$.计算 x'_{ij} 在所有的 x'_{ij} 中的秩 R_{ij} ,见表 3.8. Hodges 和 Lehmann 称调整后的观察值的秩为 Aligned 秩.表 3.8 的最右边一列是第 i 个处理里 9 个秩的和 R_{i+} ;而最下边一行是第 j 个区组里 6 个秩的和 R_{+j} .表 3.8 的最右边一列, $R_{1+} < R_{2+} < \cdots < R_{6+}$,随着剂量的增加,患者的镇痛时间的秩和在增加.由此能否下结论,该镇痛药物的剂量越大,镇痛时间就越长?这样下的结论是有风险的.基于和或平均数的相关系数有缺陷,它有可能夸大了相关程度.有关这个问题的详细说明见参考文献[9]的附录 15.

同 Page 检验,C 检验的检验统计量取为

$$C = \sum_{i=1}^6 iR_{i+} = 5\,560,$$

$R_{i+} = \sum_{j=1}^9 R_{ij}$ 是第 i 个处理的 9 个秩的和. 在 C 的值比较大的时候, 我们认为该镇痛药物的剂量越大, 镇痛时间就越长. 现在的问题是, $C=5\ 560$ 是不是比较大. 这就需要计算它的 p 值, 也就是在镇痛时间与药物剂量无关时, 计算概率 $P(C \geq 5\ 560)$. C 检验也是置换检验, 是在各个区组的秩都如表 3.8 那样给定的条件下的置换检验. 例如表 3.8 中第 1 个区组的秩依次为 6, 13, 28, 22, 36, 54. C 检验仅给定第 1 个区组的秩为这 6 个, 但它可任意置换排列. 区组 1 有 $6! = 720$ 个置换. 表 3.8 有 9 个区组, 共有 $720^9 = 5.199\ 87 \times 10^{25}$ 个置换. 所以 C 检验的 p 值, 是条件 p 值. 我们是在各个区组的秩都如表 3.8 那样给定的条件下, 计算 $C=5\ 560$ 的 p 值. 由此可见, C 检验的 p 值不仅与处理数 k 和区组数 b 有关, 还与区组内的秩有关, 因而难以构造它的检验临界值表. 在处理数 k 与区组数 b 都不大时, 可逐一置换搜索排序计算 p 值. 而在处理数 k 或区组数 b 比较大时, 通常使用渐近分布近似地计算 p 值. 下面讨论 C 检验统计量的渐近分布问题.

表 3.8 观察值经过调整后的秩

剂量/mg	患者的镇痛时间/min									R_{i+}
	1	2	3	4	5	6	7	8	9	
1	6	29	15	48	11	37	4	16	47	213
2	13	46	35	24	19	3	45	25	8	218
3	28	14	27	12	30	41	50	18	1	221
4	22	23	31	26	32	20	2	40	44	240
5	36	10	51	49	17	38	21	9	39	270
6	54	42	7	5	52	34	43	53	33	323
R_{+j}	159	164	166	164	161	173	165	161	172	

C 检验的一般提法同 Friedman 检验. 设有一个 k 个处理 b 个区组的区组设计. 假设第 i 个处理在第 j 个区组的观察值 x_{ij} 的分布函数为 $F_j(x - \theta_i)$, $i = 1, 2, \dots, k, j = 1, 2, \dots, b$. 与 Friedman 检验不同的是, C 检验用于区组设计的趋势性检验问题:

单调上升趋势检验问题 $H_0: \theta_1 = \theta_2 = \dots = \theta_k, H_1: \theta_1 \leq \theta_2 \leq \dots \leq \theta_k$, 且 $\theta_1 < \theta_k$;

单调下降趋势检验问题 $H_0: \theta_1 = \theta_2 = \dots = \theta_k, H_1: \theta_1 \geq \theta_2 \geq \dots \geq \theta_k$, 且 $\theta_1 > \theta_k$.

假设第 i 个处理第 j 个区组都只有一个观察值 x_{ij} . 消除区组效应, 将观察值 x_{ij} 调整为 $x'_{ij} = x_{ij} - \bar{x}_j$, $\bar{x}_j$ 是第 j 个区组的观察值的均值. 假设所有调整后的观察值 x'_{ij} 都

互不相等. 记 x'_{ij} 在所有的 x'_{ij} 中的秩为 R_{ij} ($R_{ij} = 1, 2, \dots, bk$). 所有的 R_{ij} 也都互不相等. C 检验是置换检验, 是第 j ($j = 1, 2, \dots, b$) 个区组的 k 个秩给定为 $R_{1j}, \dots, R_{kj}$ 的条件下的置换检验. 取 C 检验的检验统计量为

$$C = \sum_{i=1}^k iR_{i+},$$

$R_{i+} = \sum_{j=1}^b R_{ij}$ 是第 i 个处理的 b 个秩的和. 在 C 比较大时认为有单调上升趋势, 而在 C 比较小时认为有单调下降趋势. 接下来证明, 在原假设 H_0 为真时, 这 k 个位置参数全都相等, $\theta_1 = \theta_2 = \dots = \theta_k$ 时, 若 k 固定, 而 $b \rightarrow \infty$, 检验统计量 C 有渐近正态性:

$$\frac{C - \mu}{\sigma} \xrightarrow{L} N(0, 1),$$

其中 $\mu = E(C)$, $\sigma = \sqrt{D(C)}$.

将 $C = \sum_{i=1}^k iR_{i+}$ 改写为 $C = \sum_{j=1}^b Y_j$, 其中 $Y_j = \sum_{i=1}^k iR_{ij}$, $j = 1, 2, \dots, b$. 由于 C 检验是每一个区组的 k 个秩都各自给定的条件下的置换检验, 所以 $Y_1, Y_2, \dots, Y_b$ 相互独立. 下面利用独立随机变量和的 Lyapunov 定理证明 C 的渐近正态性. 当原假设为真, 处理无差异时, 在第 j 个区组的 k 个秩给定为 $R_{1j}, R_{2j}, \dots, R_{kj}$ 的条件下, 则由定理 3.24 与 3.25 知,

$$E(R_{ij}) = R_{.j},$$

$$\text{Cov}(R_{ij}, R_{sj}) = \begin{cases} \sum_{i=1}^k (R_{ij} - R_{.j})^2 / k, & i = s, \\ - \sum_{i=1}^k (R_{ij} - R_{.j})^2 / (k(k-1)), & i \neq s, \end{cases}$$

其中 $R_{.j} = R_{+j}/k$ 与 $R_{+j} = \sum_{i=1}^k R_{ij}$ 分别是第 j 个区组的 k 个秩的平均与和. 由此知

$$\mu_j = E(Y_j) = \frac{k(k+1)}{2} R_{.j}, \sigma_j^2 = D(Y_j) = \frac{k(k+1)}{12} \sum_{i=1}^k (R_{ij} - R_{.j})^2,$$

因而

$$\begin{aligned} \mu &= E(C) = \sum_{j=1}^b \mu_j = \frac{k(k+1)}{2} \sum_{j=1}^b R_{.j} = \frac{(k+1)bk(bk+1)}{4}, \\ \sigma^2 &= D(C) = \sum_{j=1}^b \sigma_j^2 = \frac{k(k+1)}{12} \sum_{j=1}^b \sum_{i=1}^k (R_{ij} - R_{.j})^2 \\ &= \frac{k(k+1)}{12} \sum_{j=1}^b \left(\sum_{i=1}^k R_{ij}^2 - kR_{.j}^2 \right) \end{aligned}$$

$$= \frac{bk^2(k+1)(bk+1)(2bk+1)}{72} - \frac{k+1}{12} \sum_{j=1}^b R_{+j}^2.$$

为了用独立随机变量和的 Lyapunov 定理证明检验统计量 C 的渐近正态性, 首先给出两个结论.

由于

$$\begin{aligned} |Y_j - \mu_j| &= \left| \sum_{i=1}^k iR_{ij} - \frac{k(k+1)}{2} R_{\cdot j} \right| \leq \sum_{i=1}^k iR_{ij} + \frac{k(k+1)}{2} R_{\cdot j} \\ &\leq bk \left(\sum_{i=1}^k i + \frac{k(k+1)}{2} \right) = bk^2(k+1), \end{aligned}$$

所以 $E(|Y_j - \mu_j|^3) \leq k^6(k+1)^3 b^3$, 从而得到下面的结论.

结论 1 存在一个仅与 k 有关的常数 d_1 , 使得 $\sum_{j=1}^b E(|Y_j - \mu_j|^3) \leq d_1 b^4$.

假设所有调整后的观察值 x'_{ij} 中, 与原点 0 最接近的那个调整后的观察值为 m . 又假设 m 在所有的 x'_{ij} 中的秩为 R . 由于是将观察值 x_{ij} 调整为 $x'_{ij} = x_{ij} - \bar{x}_{\cdot j}$, $\bar{x}_{\cdot j}$ 是第 j 个区组的观察值的均值, 并假设所有调整后的观察值 x'_{ij} 都互不相等, 所以在每一个区组都至少有一个调整后的观察值大于等于 m 的绝对值 $|m|$, 且都至少有一个调整后的观察值小于等于 $-|m|$. 由此可见,

$$\begin{aligned} \sigma_j^2 &= \frac{k(k+1)}{12} \sum_{i=1}^k (R_{ij} - R_{\cdot j})^2 \geq \frac{k(k+1)}{12} \cdot \min_{1 \leq i \leq k} (R_{ij} - R)^2, \\ \sigma^2 &= \sum_{j=1}^b \sigma_j^2 \geq \frac{k(k+1)}{12} \sum_{j=1}^b \min_{1 \leq i \leq k} (R_{ij} - R)^2, \end{aligned}$$

由于所有的 R_{ij} 都互不相等, 所以

在 b 为奇数时, $\sigma^2 \geq \frac{k(k+1)}{12} \left(1^2 + 2^2 + \cdots + \left(\frac{b-1}{2} \right)^2 \right) \times 2 = \frac{k(k+1)}{144} (b^3 - b)$;

在 b 为偶数时,

$\sigma^2 \geq \frac{k(k+1)}{12} \left(\left(\frac{b}{2} \right)^2 + \left(1^2 + 2^2 + \cdots + \left(\frac{b-2}{2} \right)^2 \right) \times 2 \right) = \frac{k(k+1)}{72} (b^3 - b^2 + b)$.

由于是在 k 固定, $b \rightarrow \infty$ 时, 证明检验统计量 C 的渐近正态性, 故不妨假设 b 充分大, 从而得下面的结论.

结论 2 存在一个仅与 k 有关的常数 d_2 , 使得 $\sigma^2 \geq d_2 b^3$.

根据结论 1 与 2, 则在 $b \rightarrow \infty$ 时

$$\frac{E(|Y_j - \mu_j|^3)}{\sigma^3} \leq \frac{d_1 b^4}{(\sqrt{d_2} b^3)^3} = \frac{d_1}{(\sqrt{d_2})^3} \frac{1}{\sqrt{b}} \rightarrow 0.$$

根据独立随机变量和的 Lyapunov 定理, 检验统计量 C 有渐近正态性.

下面用检验统计量 C 的渐近正态性, 求解表 3.7 与表 3.8 的 $k=6$ 个处理, $b=9$ 个区组的区组设计的单调上升趋势的检验问题. 经计算

$$\text{检验统计量 } C = \sum_{i=1}^6 iR_{i+} = 5\,560,$$

$$\mu = E(C) = \frac{(k+1)bk(bk+1)}{4} = 5\,197.5,$$

$$\sigma^2 = D(C) = \frac{bk^2(k+1)(bk+1)(2bk+1)}{72} - \frac{k+1}{12} \sum_{j=1}^b R_{+j}^2 = 45\,803.92,$$

$$\frac{C - E(C)}{\sqrt{D(C)}} = 1.694,$$

C 检验的 p 值为 $P(N(0,1) \geq 1.694) = 0.045$.

若取显著性水平 $\alpha=0.05$, 则拒绝原假设, 认为镇痛药物的剂量越大, 镇痛时间就越长.

参考文献



习题三

3.1 电话交换台单位时间内接到的呼唤次数服从 Poisson 分布 $P(\lambda)$, $\lambda > 0$. λ 为单位时间内接到的平均呼唤次数. 设 $x = (x_1, x_2, \dots, x_{10})$ 是该电话交换台的 10 次记录. 考虑假设检验问题: $H_0: \lambda \geq 1$ 对 $H_1: \lambda < 1$. 取显著性水平为 $\alpha = 0.05$.

3.2 设 $X = (X_1, X_2, \dots, X_n)$ 是来自正态分布族 $\{N(0, \sigma^2) : 0 < \sigma^2 < \infty\}$ 的样本, 考虑 $H_0: \sigma^2 = 1$ 对 $H_1: \sigma^2 = \sigma_1^2 (\sigma_1^2 > 1)$ 的检验问题. 取显著性水平为 $\alpha (0 < \alpha < 1)$. 试求其 MPT.

3.3 设 $X = (X_1, X_2, \dots, X_n)$ 是来自均匀分布族 $\{R(0, \theta) : \theta > 0\}$ 的样本, 考虑 $H_0: \theta = 1$ 对 $H_1: \theta = \theta_1 (\theta_1 < 1)$ 的检验问题. 取显著性水平为 $\alpha (0 < \alpha < 1)$. 试求其 MPT.

3.4 (N-P 基本引理的补遗) 设 P_{θ_0} 和 P_{θ_1} 是可测空间 $(\mathcal{X}, \mathcal{B})$ 上的两个不同的概率测度, 关于某个 σ 有限的测度 μ , 有

$$p(x; \theta_0) = \frac{dP_{\theta_0}}{d\mu}, \quad p(x; \theta_1) = \frac{dP_{\theta_1}}{d\mu},$$

则在检验问题 (3.9) 中, 如果 $\phi(x)$ 是显著性水平为 α 的 MPT, 则它不一定满足式 (3.10), 但在 $E_{\theta_1}(\phi(X)) < 1$ 的时候, 它必满足式 (3.10).

3.5 设样本 $X = (X_1, X_2, \dots, X_{100})$ 来自二点分布族 $\{b(1, p) : 0 \leq p \leq 1\}$. 试求检验问题: $H_0: p \leq 0.01$ 对 $H_1: p > 0.01$ 的显著性水平 $\alpha = 0.05$ 的 UMPT.

3.6 (定理 3.7 的推广) 设概率密度族 $\{p(x; \theta) : \theta \in \Theta \subseteq \mathbf{R}\}$ 关于 X 具有非降 MLR. 试证明:

(1) 若 $X_1, X_2, \dots, X_n$ 是来自该 MLR 分布族的一个样本, 并且 $\psi(x_1, x_2, \dots, x_n)$ 关于每一个 x_i 都是非降的, 则 $E_{\theta}(\psi(X_1, X_2, \dots, X_n))$ 是 θ 的一个非降函数;

(2) 若函数 $\psi(x)$ 具有性质: 存在点 x_0 使得在 $x < x_0$ 时, $\psi(x) \leq 0$, 而在 $x > x_0$ 时, $\psi(x) \geq 0$, 则 $E_{\theta}(\psi(X))$ 或者总是正的, 或者总是负的, 或者存在点 θ_0 使得在 $\theta < \theta_0$ 时, $E_{\theta}(\psi(X)) \leq 0$, 而在 $\theta > \theta_0$ 时, $E_{\theta}(\psi(X)) \geq 0$.

3.7 设 X 和 Y 的密度函数分别为 $f(x)$ 和 $g(y)$. 试证明: 若 $f(x)/g(x)$ 是 x 的非降函数, 则 $E(X) \geq E(Y)$.

3.8 设 $X = (X_1, X_2, \dots, X_n)$ 是来自带有位置参数的指数分布总体的样本. 总体的密度函数为

$$p(x; \theta) = \begin{cases} \exp\{-(x - \theta)\}, & x \geq \theta, \\ 0, & \text{其他,} \end{cases}$$

考虑如下的检验问题: $H_0: \theta = 0$ 对 $H_1: \theta > 0$. 试构造显著性水平为 $\alpha (0 < \alpha < 1)$ 的 UMPT.

3.9 设 $X = (X_1, X_2, \dots, X_{10})$ 是来自 Pareto 分布总体的样本. Pareto 分布的密度函数为

$$p(x; \theta) = \begin{cases} \frac{\beta}{\theta} \left(\frac{\theta}{x}\right)^{\beta+1}, & x \geq \theta, \\ 0, & \text{其他,} \end{cases}$$

其中 $\beta = 2$ 已知. 考虑如下的检验问题: $H_0: \theta = 1$ 对 $H_1: \theta > 1$. 试构造显著性水平 $\alpha = 0.1$ 的 UMPT.

3.10 设 $T(x)$ 的密度函数如式 (3.26) 所示, 即

$$p(t; \theta) = c(\theta) \exp \{ \theta t \} h(t),$$

试证明: $E_{\theta}(T) = -\frac{c'(\theta)}{c(\theta)}$.

3.11 设 $T(x)$ 的密度函数如式(3.26)所示, 其中 $c(\theta) > 0$. 单边假设检验问题: $H_0: \theta \leq \theta_0$ 对 $H_1: \theta > \theta_0$ 的显著性水平为 $\alpha (0 < \alpha < 1)$ 的 UMPT $\phi(T)$ 如式(3.19)所示, 即

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

其中常数 $r (0 \leq r \leq 1)$ 和 c 由式(3.20), 即由 $E_{\theta_0}(\phi(T)) = \alpha$ 确定. 我们知道这个检验的势函数 $g(\theta) = E_{\theta}(\phi(T(X)))$ 是非降的, 且在集合 $\{\theta: 0 < g(\theta) < 1\}$ 上是严增的. 试证明: $g'(\theta_0) > 0$. (提示: 为了证明 $g'(\theta_0) > 0$, 考虑如下的问题: 在满足条件 $E_{\theta_0}(\phi(T)) = \alpha$ 的检验函数 $\phi(T)$ 中寻找一个检验, 使得其势函数在点 θ_0 处的导数 $g'(\theta_0)$ 达到最大, 或等价地使得 $E_{\theta_0}(\phi(T)T)$ 最大.)

3.12 设样本 X 服从多参数指数型分布(3.47). 试证明: $H_0: \theta = \theta_0$ 对 $H_1: \theta \neq \theta_0$ 的双边假设检验问题(III)的显著性水平为 α 的 UMPUT 由式(3.53)给出, 其中 $r_i(t) (0 \leq r_i(t) \leq 1)$ 和 $c_i(t)$ 由式(3.50)和式(3.53)确定.

3.13 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自正态总体 $N(\mu, \sigma^2)$, 其中 $-\infty < \mu < \infty$, $\sigma^2 > 0$. 在 $\mu = 0$ 时, 试证明 w 的分布与参数 σ^2 无关, 且 w 的分布关于原点对称, 其中

$$w = \frac{\bar{X}}{\sqrt{\sum_{i=1}^n X_i^2}}.$$

3.14 设样本 $X = (X_1, X_2, \dots, X_m)$ 和样本 $Y = (Y_1, Y_2, \dots, Y_n)$ 互相独立. 它们分别来自正态总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$, 其中: $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$, $\sigma^2 > 0$. 在 $\mu_1 = \mu_2 = \mu$ 时, 试证明: V 的分布与参数 σ^2 和 μ 无关, V 的分布关于原点对称, 其中

$$V = \frac{U}{\sqrt{T_2 - \frac{T_1^2}{m+n}}}, \quad U = \bar{Y} - \bar{X},$$

$$T_1 = m\bar{X} + n\bar{Y}, \quad T_2 = \sum_{i=1}^m X_i^2 + \sum_{i=1}^n Y_i^2.$$

3.15 设样本 $X = (X_1, X_2, \dots, X_m)$ 和样本 $Y = (Y_1, Y_2, \dots, Y_n)$ 互相独立. 它们分别来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$, 其中 $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$, $\sigma_1^2 > 0, \sigma_2^2 > 0$.

(1) 试证明: 在 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 时,

$$V = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2} \sim Be\left(\frac{n-1}{2}, \frac{m-1}{2}\right).$$

(2) 试证明: 检验问题 $H_0: \sigma_1^2 = \sigma_2^2$ 对 $H_1: \sigma_1^2 \neq \sigma_2^2$ 的 UMPUT 的拒绝域为 $\{x, y: v \leq c_1 \text{ 或 } v \geq c_2\}$, 其中 c_1 和 c_2 由以下两式确定:

$$\int_{c_1}^{c_2} Be\left(x \mid \frac{n-1}{2}, \frac{m-1}{2}\right) dx = 1 - \alpha,$$

$$\int_{c_1}^{c_2} \left(x - \frac{n-1}{m+n-2}\right) Be\left(x \mid \frac{n-1}{2}, \frac{m-1}{2}\right) dx = 0.$$

(3) 试证明: 该 UMPUT 的拒绝域可等价地写为 $\{x, y: F \leq c_1 \text{ 或 } F \geq c_2\}$, 其中 c_1 和 c_2 由以下两式确定:

$$\int_{c_1}^{c_2} F(x \mid n-1, m-1) dx = 1 - \alpha,$$

$$\int_{c_1}^{c_2} \frac{x-1}{(n-1)x + (m-1)} F(x \mid n-1, m-1) dx = 0.$$

3.16 设样本 $X = (X_1, X_2, \dots, X_n)$ 来自双参数指数分布总体, 其密度函数为

$$p(x; \mu, \sigma) = \begin{cases} \frac{1}{\sigma} \exp\left\{-\frac{x-\mu}{\sigma}\right\}, & x \geq \mu, \\ 0, & x < \mu, \end{cases}$$

(1) 试求检验问题 $H_0: \sigma = 1$ 对 $H_1: \sigma \neq 1$ 的 UMPUT;

(2) 试求检验问题 $H_0: \mu = 0$ 对 $H_1: \mu \neq 0$ 的 UMPUT.

3.17 (定理 3.18 的补遗) 试证明:

$$I_{ij}(\theta) = E_{\theta} \left(\frac{\partial \ln f(X; \theta)}{\partial \theta_i} \frac{\partial \ln f(X; \theta)}{\partial \theta_j} \right) = -E_{\theta} \left(\frac{\partial^2 \ln f(X; \theta)}{\partial \theta_i \partial \theta_j} \right).$$

3.18 设 $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 为来自二元正态分布 $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ 的样本, 其中 ρ 是相关系数. 试求检验问题 $H_0: \rho=0$ 对 $H_1: \rho \neq 0$ 的似然比检验.

3.19 试证明下述结论:

(1) 接例 3.16,

$$\begin{aligned} 2\ln \Lambda &= 2 \sum_{i=1}^r n_i \ln \frac{n_i}{np_{0i}} \\ &= \sum_{i=1}^r \frac{(n_i - np_{0i})^2}{np_{0i}} + \text{依概率收敛于 0 的量}; \end{aligned}$$

(2) 接 § 3.6.4,

$$\begin{aligned} 2\ln \Lambda &= 2 \sum_{i=1}^r \sum_{j=1}^c n_{ij} \ln \frac{nn_{ij}}{n_{i\cdot} n_{\cdot j}} \\ &= \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n \hat{p}_{i\cdot} \hat{p}_{\cdot j})^2}{n \hat{p}_{i\cdot} \hat{p}_{\cdot j}} + \text{依概率收敛于 0 的量}. \end{aligned}$$

3.20 设有 X, Y 两个离散型随机变量. X 表示人的性别, “ $X=1$ ”表示男性, “ $X=2$ ”表示女性; Y 表示是否色盲, “ $Y=1$ ”表示正常, “ $Y=2$ ”表示色盲. 令 $p_{ij} = P(X=i, Y=j), i, j=1, 2$. 这时 $\theta = (p_{11}, p_{12}, p_{21})$, 其中 $0 \leq p_{11}, p_{12}, p_{21} \leq 1$, 且 $p_{11} + p_{12} + p_{21} \leq 1$. 所以参数空间 Θ 为 $\mathbf{R}^3$ 中的一个单纯形.

有 1 000 人按性别与是否色盲分类如下

性别	正常	色盲
男	442	38
女	514	6

按遗传学模型, 数据应有下列相对应的概率

性别	正常	色盲
男	$\frac{p}{2}$	$\frac{1-p}{2}$
女	$\frac{p^2}{2} + p(1-p)$	$\frac{(1-p)^2}{2}$

其中 $0 \leq p \leq 1$. 问数据与模型是否相符?

3.21 设有 X, Y, Z 三个离散型随机变量, 与此相联系的有一个三维列联表 $r \times c \times t$. 作 n 次观测, 在 (i, j, k) 格的观测频数为 $n_{ijk}, i=1, 2, \dots, r; j=1, 2, \dots, c; k=1, 2, \dots, t$. 观测值落入 (i, j, k) 格的概率为 p_{ijk} .

(1) 试证明:在给定 Z 之后, X 和 Y 条件相互独立的时候, p_{ijk} 的 MLE 为

$$\hat{p}_{ijk} = \frac{n_{i \cdot k} n_{\cdot jk}}{nn_{\cdot \cdot k}}.$$

(2) $A = \{ (p_{i \cdot k}, p_{\cdot jk}, p_{\cdot \cdot k}) : i = 1, 2, \dots, r; j = 1, 2, \dots, c; k = 1, 2, \dots, t; p_{i \cdot k} \geq 0, p_{\cdot jk} \geq 0, \sum_{i=1}^r p_{i \cdot k} = \sum_{j=1}^c p_{\cdot jk} = p_{\cdot \cdot k}, \sum_{k=1}^t p_{\cdot \cdot k} = 1 \}$, A 中独立参数为什么有 $t(r+c-1)-1$ 个?

3.22 接 § 3.6.5, 试解下列检验问题:

(1) 原假设 H_0 : 给定被告的肤色后被害人的肤色与死刑判决相互条件独立;

(2) 原假设 H_0 : 给定是否死刑判决后被告的肤色与被害人的肤色相互条件独立.

3.23 下表是某大学秋季招生情况的数据. 令 A = 是否录取, B = 系别, C = 性别.

系别(B)	性别(C)	是否录取(A)	
		是	否
B_1	男	353	207
	女	17	8
B_2	男	120	205
	女	202	391
B_3	男	138	279
	女	131	244
B_4	男	53	138
	女	94	299
B_5	男	22	351
	女	24	317

(1) 检验: A 和 C 的相互独立性. 该大学秋季招生有没有性别歧视? 如果有, 哪种性别的录取率高?

(2) 检验: 给定 B 之后, A 和 C 的条件独立性. 该大学秋季招生有没有性别歧视?

3.24 试证明:

(1) 以 $\phi(X_1, X_2, \dots, X_m)$ 为核的 U 统计量的方差不可能比核 $\phi(X_1, X_2, \dots,$

X_m) 的方差大;

(2) 若 U 统计量的核 $\phi(X_1, X_2, \dots, X_m)$ 的方差有限, 则 $\phi_c(X_1, X_2, \dots, X_c)$ 的方差 σ_c^2 也有限, 其中 $\phi_c(X_1, X_2, \dots, X_c)$ 如式 (3.73) 所示.

3.25 接例 3.17, 在原假设成立时, 试证明:

$$\phi_2(x_1, x_2) = \begin{cases} [F(x_1 - x_2) + \bar{F}(x_1 + x_2)]/3, & x_1 > x_2, \\ [F(x_2 - x_1) + \bar{F}(x_1 + x_2)]/3, & x_2 > x_1, \end{cases}$$

$$\sigma_2^2 = \frac{7}{1296},$$

$$\phi_3(x_1, x_2, x_3) = \phi(x_1, x_2, x_3),$$

$$\sigma_3^2 = \frac{1}{48}.$$

3.26 接例 3.18, 在原假设成立时, 试证明:

$$\phi_1(x_1) = F(x_1), \quad \sigma_1^2 = \frac{1}{12},$$

$$\phi_2(x_1, x_2) = \phi(x_1, x_2), \quad \sigma_2^2 = \frac{1}{4},$$

$$\text{Var}(U(X_1, X_2, \dots, X_n)) = \frac{2n-1}{6n(n-1)}.$$

3.27 接例 3.19, 在原假设成立时, 试证明:

$$\phi_{10}(x_1) = G(x_1), \quad \phi_{01}(y_1) = \bar{F}(y_1), \quad \phi_{11}(x_1, y_1) = \phi(x_1, y_1),$$

$$\sigma_{10}^2 = \sigma_{01}^2 = \frac{1}{12}, \quad \sigma_{11}^2 = \frac{1}{4},$$

$$\text{Var}(U(X_1, X_2, \dots, X_{n_1}; Y_1, Y_2, \dots, Y_{n_2})) = \frac{n_1 + n_2 + 1}{12n_1n_2}.$$

3.28 (Kendall τ 检验) 设 $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 是来自连续总体 (X, Y) 的样本. 欲检验假设: 原假设 $H_0: X$ 和 Y 相互独立对备择假设 $H_1: X$ 和 Y 正相关. 1938 年 Kendall 构造了以 $\phi((X_1, Y_1), (X_2, Y_2))$ 为核的 U 统计量, 其中

$$\phi((X_1, Y_1), (X_2, Y_2)) = \begin{cases} 1, & (X_1 - X_2)(Y_1 - Y_2) > 0, \\ 0, & \text{其他.} \end{cases}$$

(1) 在原假设成立时, 试证明: $E(U) = 0.5$;

(2) 这里将正相关理解为

$$P(Y_1 - Y_2 > 0 \mid X_1 - X_2 > 0) > 0.5,$$

$$P(Y_1 - Y_2 < 0 \mid X_1 - X_2 < 0) > 0.5,$$

在原假设不成立时, 试证明: $E(U) > 0.5$;

(3) 试在大样本场合下, 给出检验的拒绝域;

(4) 上述检验问题可认为是“有方向的”检验问题, 而下述检验问题, 原假设 $H_0: X$ 和 Y 相互独立, 对备择假设 $H_1: X$ 和 Y 不相互独立, 可认为是“无方向的”检验问题. 令

$$\theta = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(x, y) - F_x(x)F_y(y)]^2 dF(x, y),$$

基于样本构造参数 θ 的 U 统计量(假定样本容量足够大), 并基于该 U 统计量给出这个“无方向的”检验问题的解, 其中 $F(x, y)$ 是 (X, Y) 的联合分布函数, $F_x(x)$ 和 $F_y(y)$ 分别是 X 和 Y 的边缘分布函数.

3.29 设 $x_1, x_2, \dots, x_n$ 独立同连续型分布. 记 $x_i (i=1, 2, \dots, n)$ 在 $x_1, x_2, \dots, x_n$ 中的秩为 $R_i, R_i=1, 2, \dots, n$. 试证明:

$$(1) E(R_i) = (n+1)/2, i=1, 2, \dots, n;$$

$$(2) \text{Var}(R_i) = (n^2-1)/12, i=1, 2, \dots, n;$$

$$(3) \text{Cov}(R_i, R_j) = -(n+1)/12, 1 \leq i < j \leq n.$$

3.30 设 $x_1, x_2, \dots, x_n$ 独立同连续型分布. 记 $x_i (i=1, 2, \dots, n)$ 在 $x_1, x_2, \dots, x_n$ 中的秩为 $R_i, R_i=1, 2, \dots, n$. 设有计分函数 $a(r), r=1, 2, \dots, n$. 令

$$\bar{a} = \frac{1}{n} \sum_{i=1}^n a(i). \text{ 试证明:}$$

$$(1) E(a(R_i)) = \bar{a}, i=1, 2, \dots, n;$$

$$(2) \text{Var}(a(R_i)) = \frac{1}{n} \sum_{i=1}^n (a(i) - \bar{a})^2, i=1, 2, \dots, n;$$

$$(3) \text{Cov}(a(R_i), a(R_j)) = -\frac{1}{n(n-1)} \sum_{i=1}^n (a(i) - \bar{a}), 1 \leq i < j \leq n.$$

3.31 关于连续型对称分布的对称中心的检验问题, 可以用 U 统计量的方法(见 § 3.7.3), 也可以用符号检验和符号秩和检验方法(见 § 3.8.2). 试证明, 它们之间有下列两个关系:

(1) 试证明: $\binom{n}{2} U = W^+ - S^+$, 其中 U 见例 3.18, W^+ 见式 (3.90), S^+ 见式 (3.89);

(2) 试证明: $\binom{n}{2} U \stackrel{d}{=} \sum_{i=1}^n (i-1)u_i$, 并根据这个结果求 U 的期望和方差,

验证 U 的渐近正态性;

(3) 试证明: $\sqrt{3n} \frac{W^+ - n(n+1)/4}{\binom{n}{2}}$ 和 $\sqrt{3n}(U-1/2)$ 仅相差一个依概率收敛于 0 的量;

(4) 试由 U 的渐近正态性(见例 3.18), 证明 W^+ 的渐近正态性.

3.32 关于位置参数的检验问题, 可以用 U 统计量的方法(见 § 3.7.4), 也可以用秩和检验方法(见 § 3.8.3). 试证明它们之间有下述的关系:

$$mnU = W_y - \frac{n(n+1)}{2}.$$

3.33 在 $x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_n$ 同连续型分布时, 试求下列尺度参数检验问题的检验统计量的期望和方差:

(1) Ansari-Bradley 检验统计量;

(2) Mood 检验统计量.

3.34 试根据 Hajek 定理, 验证下列检验统计量的渐近正态性:

(1) Wilcoxon 秩和检验统计量 W_y ;

(2) Ansari-Bradley 检验统计量;

(3) Mood 检验统计量.

3.35 如 § 3.9.1 小节, 与 Friedman 检验相类似地, § 3.8.3 小节的位置参数的秩和检验可视为置换检验. 这个检验置换的是数据的秩. 其实, 数据本身也可以置换. 例如, 某种羊毛在进行某种工艺处理之前与处理之后, 各随机抽取一个样本, 测得其含脂率如下:

处理前: 0.20, 0.24, 0.66, 0.42, 0.12;

处理后: 0.13, 0.07, 0.21, 0.08, 0.19.

问该处理能否降低含脂率? 倘若处理前后含脂率无差异, 则处理前样本的含脂率可能是观察到的 (0.20, 0.24, 0.66, 0.42, 0.12), 也可能是 (0.13, 0.24, 0.66, 0.42, 0.12) 等, 共有 $C_{10}^5 = 252$ 个不同的置换. 记处理前后样本均值分别为 $\bar{X}$ 与 $\bar{Y}$. 由样本观察值算得的 $\bar{X} - \bar{Y} = 32.8 - 13.6 = 19.2$. 显然, 在 $\bar{X} - \bar{Y} = 19.2$ 比较大的时候拒绝原假设. 认为该处理能降低含脂率.

试求该置换检验的解(取显著性水平 $\alpha = 0.05$).