北京工商大学经济学博士文库

假设检验问题中的频率与贝叶斯证据的和谐性

JIASHE JIANYAN WENTIZHONG DE PINLÜ YU BEIYESI ZHENGJU DE HEXIEXING

尹玉良◎著

知识产权出版社
全国百佳图书出版单位

北京工商大学经济学
博士文库

北京市哲学社会科学首都流通业研究基地资助出版（JD-2012-Y-05）

假设检验问题中的频率与贝叶斯证据的和谐性

尹玉良◎著

知识产权出版社
全国百佳图书出版单位

内容提要

本书研究假设检验问题中频率与贝叶斯（Bayes）证据的关系问题，分别研究单边、单点以及区间原假设的假设检验问题中两类证据的一致性或和谐性。对于单边假设检验问题，我们将频率与贝叶斯证据的和谐性推广到单边假设检验问题的广义p－值，研究广义p－值与原假设的后验概率的和谐性问题，得到了一类统计模型下两类证据的一致性；对于单点假设检验问题，针对传统贝叶斯检验方法下的贝叶斯证据，即贝叶斯因子或单点原假设成立的后验概率，会产生与经典频率证据即p－值不相和谐的所谓林德利（Lindley）悖论的现象，我们提出了一种新的贝叶斯检验方法，新的检验方法下的贝叶斯证据，具有检验和区间估计的对偶性，且从某种意义上解决了Lindley悖论；对于原假设为区间的假设检验问题，针对传统贝叶斯证据与经典频率证据之间所存在的不和谐现象，我们同样提出一种新的区间假设检验问题的贝叶斯方法，给出区间假设检验问题的新的贝叶斯证据。

责任编辑：兰涛

图书在版编目（CIP）数据

假设检验问题中的频率与贝叶斯证据的和谐性/尹玉良著.
—北京：知识产权出版社，2012.6
ISBN 978-7-5130-1149-5

Ⅰ.①假… Ⅱ.①尹… Ⅲ.①假设检验－研究
Ⅳ.①O212.1

中国版本图书馆CIP数据核字（2012）第038187号

假设检验问题中的频率与贝叶斯证据的和谐性

尹玉良　著

出版发行：知识产权出版社

社　　址：	北京市海淀区马甸南村1号	邮　　编：	100088
网　　址：	http://www.ipph.cn	邮　　箱：	bjb@cnipr.com
发行电话：	010-82000860转8101/8102	传　　真：	010-82000860转8240
责编电话：	010-82000860转8325	责编邮箱：	lantao@cnipr.com
印　　刷：	北京中献拓方科技发展有限公司	经　　销：	新华书店及相关销售网点
开　　本：	787mm×1092mm　1/16	印　　张：	9
版　　次：	2012年5月第1版	印　　次：	2012年5月第1次印刷
字　　数：	140千字	定　　价：	28.00元

ISBN 978-7-5130-1149-5/O·013(4027)

摘 要

本书研究假设检验问题中频率与贝叶斯(Bayes)证据的关系问题,分别研究单边、单点以及区间原假设的假设检验问题中两类证据的一致性或和谐性。

对于单边假设检验问题,我们将频率与贝叶斯证据的和谐性推广到单边假设检验问题的广义 p - 值,研究广义 p - 值与原假设的后验概率的和谐性问题,得到了一类统计模型下两类证据的一致性。而针对现有的单边假设检验问题中频率与贝叶斯证据的和谐性提法的局限性或片面性,我们给出一种更为合理的和谐性的定义。在新的和谐性定义下,我们首先得到了关于位置族参数的单边假设检验问题中两类证据的和谐性。其次,给出了方差未知时关于正态均值的单边假设检验问题的证据和谐性。最后,对于著名的贝伦斯 - 费希尔(Behrens - Fisher)问题,模拟研究表明,两类证据的和谐性结论仍然成立。

对于单点假设检验问题,针对传统贝叶斯检验方法下的贝叶斯证据,即贝叶斯因子或单点原假设成立的后验概率,会产生与经典频率证据即p - 值不相和谐的所谓林德利(Lindley)悖论的现象,我们提出了一种新的贝叶斯检验方法。新的检验方法下的贝叶斯证据,具有检验和区间估计的对偶性,且从某种意义上解决了林德利悖论。对于许多经典的单点假设检验问题,如对位置参数、正态均值、二项分布成功概率、线性回归模型回归系数等的假设检验问题以及著名贝伦斯 - 费希尔问题,我们分别研究新的贝叶斯证据的合理性及其与相应频率证据的一致性或和谐性。

对于原假设为区间的假设检验问题,针对传统贝叶斯证据与经典频率证据之间所存在的不和谐现象,我们同样提出一种新的区间假设检验问题的贝叶斯方法,给出区间假设检验问题的新的贝叶斯证据。对于正态均值以及更为一般的位置参数的区间假设检验问题,我们得到无信息先验下新的贝叶斯证据与频率证据的一致性以及合理先验类下两类证据的和谐性。

Abstract

The relationship of the Bayesian and the frequentist evidence is considered in this book. We study the agreement or the reconcilability of these two kinds of evidence in testing a one-sided hypothesis, a point null hypothesis and an interval null hypothesis respectively.

For the problem of testing a one-sided hypothesis, we generalize the reconcilability of the Bayesian and the frequentist evidence to the situation of the generalized p-value and consider the reconcilability between the generalized p-value and the posterior probability that the null hypothesis is true. What's more, we propose a more reasonable definition of reconcilability of evidence for the problem of testing a one-sided hypothesis. Under this new definition, we can obtain the reconcilability of evidence in testing a location parameter of the location family. We also can obtain the reconcilability of evidence in testing a normal mean when the variance is unknown. Moreover, numerical results illustrate that evidence are reconcilable for the famous Behrens-Fisher problem.

For the problem of testing a point null hypothesis, Lindely's paradox shows that the conventional Bayesian evidence, the Bayes factor or the posterior probability of the null hypothesis being true, which is derived under the conventional Bayesian testing procedure, is usually irreconcilable with the frequentist p-value. We then propose a new testing procedure under the Bayesian framework. The new Bayesian evidence derived from this new testing procedure has the duality of testing and interval estimation and it resolves the Lindley's paradox in a sense. For a number of classical point null testing problems, such as the problem of testing a location parameter, a normal mean, the success probability of a binomial distribution and the coefficients in a normal regression model, and for the well known Behrens-Fisher problem, we study the rationality of the new Bayesian evidence and the agreement or the reconcilability of evidence respectively.

In testing an interval null hypothesis, the conventional Bayesian evidence is also usually irreconcilable with its frequentist counterpart, the p-value. We therefore provide a new Bayesian procedure and new Bayesian evidence for testing interval null hypotheses. For the problem of testing a normal mean, and more generally, for the problem of testing a location parameter in a location family, the agreement or the reconcilability of the new Bayesian evidence and the frequentist evidence are studied.

目　录

第一章　绪　论

本书研究假设检验问题中频率与贝叶斯证据的一致性或和谐性。这是一个涉及统计学中频率和贝叶斯两学派之争的研究领域，被公认为是统计学中的根本性的且带有哲学色彩的问题。本章首先介绍本书研究中所需要的一些预备知识，包括 p - 值检验、广义 p - 值以及贝叶斯检验等的相关背景。其次，简要介绍假设检验问题中频率与贝叶斯证据的一致性或和谐性的研究历史和现状。最后，概述本书的主要结果及整体结构。

§1.1　显著性检验与 p - 值

在经典统计推断中，一种行之有效的检验方法是显著性水平检验。它通过给出观测显著性水平即 p - 值来给出检验的证据度量。事实上，这一检验方法的使用可以追溯到 Pearson[1]，Gossett[2] 和 Fisher[3]。较之基于 Neyman - Pearson 基本引理的固定水平检验，显著性水平检验可为使用者提供更多的证据信息，而不是仅局限于一个 0.01 或者 0.05 的主观名义水平。而且，显著性水平检验的应用范围更加广泛。在离散分布族下，具有给定名义水平的非随机化的固定水平检验往往不可获得。因此，人们如果想要完成固定水平检验，只有借助于某些近似的手法。而显著性水平检验不会存在类似的问题。即使是在离散分布族下，通常也可以容易地获得精确的 p - 值。

对于单边假设检验问题

$$H_0:\theta \leqslant \theta_0 \quad v.s. \quad H_1:\theta > \theta_0, \tag{1.1}$$

如果检验变量 $T = T(X,\theta)$ 是随机增的，则我们可以给出 $H_0:\theta \leqslant \theta_0$ 的极端域

$$C_x = \{X:T(X,\theta) \geqslant T(x,\theta)\},$$

进而可以给出单边假设检验问题的 p - 值

$$p(x) = \sup_{\theta \leqslant \theta_0} P(X \in C_x \mid \theta).$$

而对于单点假设检验问题

$$H_0: \theta = \theta_0 \quad v.s. \quad H_1: \theta \neq \theta_0, \tag{1.2}$$

以及区间假设检验问题

$$H_0: |\theta - \theta_0| \leqslant a \quad v.s. \quad H_1: |\theta - \theta_0| > a, \tag{1.3}$$

如果存在一个检验统计量 $T(X)$，它的取值越大，表示越不支持原假设，那么，对于单点假设检验问题(1.2)，p - 值即观测显著性水平定义为

$$p(x) = P_{\theta=\theta_0}(T(X) \geqslant T(x)),$$

而对于区间假设检验问题(1.3)的 p - 值应定义为

$$p(x) = \sup_{|\theta-\theta_0| \leqslant a} P(T(X) \geqslant T(x)).$$

事实上，如果检验统计量 $T(X)$ 的分布不是关于 θ 对称的，那么关于单点假设检验问题的 p - 值的定义目前尚无定论。Gibbons 和 Pratt[4]以及 Cox 和 Hinkley[5]给出了问题(1.2)的一种 p - 值的定义。而 Weerahandi[6]给出了如下的较为一般的定义。

如果 T 是 θ 的一个检验统计量，C_t 是 T 的样本空间的一个子集，如果 C_t 满足，对于固定的 t，当 $\theta \geqslant \theta_0$ 时 $P(T \in C_t)$ 是 $\theta - \theta_0$ 的非降函数，而当 $\theta < \theta_0$ 时 $P(T \in C_t)$ 是 $\theta_0 - \theta$ 的非降函数，那么检验问题(1.3)的 p - 值定义为

$$p(x) = P(T \in C_{t_{obs}} \mid \theta = \theta_0).$$

显然，$p(x)$ 越小，则意味着数据中所提供的不支持原假设 H_0 的证据越强。

然而，在统计推断中还有一大类假设检验问题，由于讨厌参数的存在，经典 p - 值往往不易获得，甚至根本不存在。典型的例子如 Behrens - Fisher 问题。Weerahandi[7]使用广义 p - 值来对具有异方差的两个回归参数进行比较，并说明了它是一个无偏极端域的概率。受此启发，Tsui 和 Weerahandi[8]给出了广义 p - 值的定义及其构造方法。

定义 1.1　对于形如(1.1)的单边假设检验问题，设随机变量 X 具有概率分布 $P_\eta(\cdot)$，其中 $\eta = (\theta, \delta)$ 是参数空间 Ω 中的一个未知向量，$\theta = \theta(\eta)$ 是一个实值兴趣参数，δ 是讨厌参数，假设 Θ 是 θ 的参数空间，x 是 X 的观测值。那么，如果设 $T = T(X; x, \eta)$ 是 X，x 和 $\eta = (\theta, \delta)$ 的函数，则 T 被称为是广义检验变

量,如果它满足如下三个条件:

(a) $t = T(x;x,\eta)$ 不依赖于未知参数;

(b)对于具体的 θ, T 的概率分布与讨厌参数无关;

(c)对于固定的 x 和 δ, $P(T(X;x,\eta) \geqslant T(x;x,\eta) \mid \theta)$ 对于给定的 t 关于 θ 非降。

根据性质(a)-(c), $T(X;x,\eta)$ 的较大值 $T(x;x,\eta)$ 可以视为其分布在原假设 H_0 下的极端值,因此它们表明较强的不支持原假设的证据,于是广义 p-值可基于 $T(X;x,\eta)$ 定义如下:

$$\begin{aligned} p(x) &= \sup_{\theta \leqslant \theta_0} P(T(X;x,\eta) \geqslant T(x;x,\eta) \mid \theta) \\ &= P(T(X;x,\eta) \geqslant T(x;x,\eta) \mid \theta_0) \end{aligned} \tag{1.4}$$

$p(x)$ 越小,则表明不支持原假设 H_0 的证据越强。

广义 p-值概念的引入,很好地解决了因讨厌参数存在而经典 p-值难以获得这一问题。

Thursby[9]和 Griffiths 和 Judge[10]则将广义 p-值的概念应用到回归模型和混合模型的研究中。进一步,Hannig,Iyer 和 Patterson[11]和 Li,Xu 和 Li[12]从 Fiducial 推断的角度提供了一种构造广义 p-值的一般方法。

Fiducial 推断可在没有先验信息的情况下,给出参数的概率陈述即 Fiducial 分布,这一分布可将含于数据中的有关兴趣参数的信息包含其中。Fiducial 方法最初由 Fisher[13-14]提出,基于此方法,便可基于 Fiducial 分布来进行参数的统计推断。

Fisher 最早给出了一种 Fiducial 分布的定义,即在满足一定条件的单参数分布族 $\{F(x):\theta \in \Omega\}$ 中,称

$$f(\theta \mid x) = -\frac{\partial}{\partial \theta} F(x \mid \theta)$$

为参数 θ 的 Fiducial 密度。

Dawid 和 Stone[15]基于函数模型给出了如下的 Fiducial 分布的定义:

定义 1.2 假设数据 X,参数 θ 和误差 E 分别定义于空间 X, Θ 和 Ξ 上,若它们满足如下三个条件:

(i) X 可由 θ 和 E 唯一决定,其关系可表为 $X = \theta^{\circ} E$;

(ii) E 具有 Ξ 上的分布 P，且 P 与 θ 无关；

(iii) 只有 X 可被观测。

若对 θ 进行推断，则

(1) 称 $X = \theta^\circ E$ 为函数模型；

(2) 若对任意 $x \in \mathcal{X}, e \in \Xi$，方程 $x = \theta^\circ e$ 都有唯一解，设为 $\hat{\theta}_x(e)$，则称 $\Theta = \hat{\theta}_x(E)$，$E \sim P$ 为 Fiducial 模型，称 $\Theta = \hat{\theta}_x(E)$ 在 P 下的分布为 Fiducial 分布，称 $\hat{\eta}(\hat{\theta}_x(x))$ 为 $\eta = \eta(\theta)$ 的(边际) Fiducial 分布。

由定义 1.2 知，Dawid 和 Stone[15]通过结构方程的解来求得信仰分布，但这限于简单函数模型和正规函数模型。Xu 和 Li[16]将 Dawid 和 Stone[15]的工作进行推广，基于枢轴分布族、广义枢轴模型和 Fiducial 模型三个重要概念，利用拟合思想提出新的 Fiducial 分布的定义。

定义 1.3　设随机变量 $X \in \mathcal{X}$ 服从参数分布族 $\{P_\theta, \theta \in \Theta\}$，$\Theta$ 是欧氏空间的 Borel 集，若存在

(i) 定义在空间 Ξ 上的随机变量 E，E 具有已知分布 Q，Ξ 为 Borel 集；

(ii) 定义在 $\Theta \times \Xi$ 上的一个函数 $h(\theta,e)$，使得对所有的 $\theta \in \Theta$，当 $X \sim P_\theta$ 时，X 和 $h(\theta,E)$ 服从相同分布，

则称

(1) $\mathcal{X}$ 上的分布族 $\{P_\theta, \theta \in \Theta\}$ 为枢轴分布族；

(2) 模型

$$X \overset{d}{=} h(\theta,E), \quad E \sim Q$$

为广义枢轴模型；

(3) E 称为枢轴随机元。

基于上述定义，Xu 和 Li[16]给出如下更为一般的 Fiducial 分布的定义。

定义 1.4　设 $X \sim P_\theta$，$\{P_\theta, \theta \in \Theta\}$ 为 $\mathcal{X}$ 上的枢轴分布族。存在广义枢轴模型

$$X \overset{d}{=} h(\theta,E), \quad E \sim Q,$$

令 $d(\cdot,\cdot)$ 表示空间 $\mathcal{X}$ 中的欧氏距离，对任意给定的观测值 $x \in \mathcal{X}$ 和 $e \in \Xi$，如果 θ 的函数 $d(x,h(\theta,e))$ 在 Θ 中具有唯一的最小值点，记为 $\hat{\theta}_x(e)$，即

$$\hat{\theta}_x(e) = \arg\min_{\theta \in \Theta} d(x,h(\theta,e)),$$

则称

$$\Theta = \hat{\theta}_x(E), \quad E \sim Q,$$

为参数 θ 的 Fiducial 模型,称在 $E \sim Q$ 下 $\hat{\theta}_x(E)$ 的分布为参数 θ 的 Fiducial 分布。

定义1.3和定义1.4中的枢轴元E是普通的随机变量,当x给定时,它的函数 $\hat{\theta}_x(E)$ 也是一个随机变量,所以 Fiducial 分布是 Kolmogorov 意义下的一个真正的概率测度,于是同样可以自然地定义兴趣参数 θ 的函数 $\eta(\theta)$ 的边际 Fiducial 分布。

由定义1.2, θ 的 Fiducial 分布应为

$$F_x(\theta) = P(\hat{\theta}_x(E) \leqslant \theta).$$

Li, Xu 和 Li[12]说明了如果定义1.2中的条件成立,且方程 $x = h(\theta, e)$ 对任给的 θ 和 x 在 Ξ 中有唯一解,则检验问题(1.1)的广义 p－值恰为其 Fiducial p－值, $p = F_x(\theta_0)$。

更多有关广义推断和 Fiducial 推断的文献见 Lindley[17], Fraser[19－20], Banard[21], Seidenfeld[22], Dawid 和 Wang[23], Weerahandi[24], Gamage 和 Weerahandi[25], Brown, Cai 和 DasGupta[26], Weerahandi[27]等。

§1.2 贝叶斯检验与贝叶斯证据

设 X 具有密度 $f(x \mid \theta)$,则对于假设检验问题

$$H_0: \theta \in \Theta_0 \quad v.s. \quad H_1: \theta \in \Theta_1, \tag{1.5}$$

贝叶斯学派通常用原假设成立的后验概率 $P(H_0 \mid x)$ 或者加权似然比即贝叶斯因子

$$B(x) = \frac{\int_{H_0} f(x \mid \theta)\pi_0(\theta)\,\mathrm{d}\theta}{\int_{H_1} f(x \mid \theta)\pi_1(\theta)\,\mathrm{d}\theta}$$

来作为原假设成立与否的证据度量,其中, $\pi_0(\theta)$ 和 $\pi_1(\theta)$ 分别为兴趣参数在 H_0 和 H_1 下的先验密度。

对于单点假设检验问题 (1.2),为了给出原假设成立的后验概率,我们必须为原假设 $H_0:\theta = \theta_0$ 赋予一个非零的先验概率 π_0。若参数在 H_1 上的先验密度为 $\pi(\theta)$,那么可得检验问题(1.2)的贝叶斯因子以及原假设成立的后验概率分别为

$$B(x) = \frac{f(x \mid \theta_0)}{m_\pi(x)}$$

和

$$P(H_0 \mid x) = \left[1 + \frac{1-\pi_0}{\pi_0}\frac{m_\pi(x)}{f(x \mid \theta_0)}\right]^{-1} = \left[1 + \frac{1-\pi_0}{\pi_0}\frac{1}{B(x)}\right]^{-1},$$

其中

$$m_\pi(x) = \int f(x \mid \theta)\pi(\theta)\,\mathrm{d}\theta.$$

而对于区间假设检验问题(1.3),若设参数先验密度为 $\pi(\theta)$,并设原假设 H_0 和备择假设 H_1 的先验概率分别为 c 和 $1-c$,于是 $\pi(\theta)$ 可记为

$$\pi(\theta) \propto c\pi_0(\theta)I\{\mid \theta - \theta_0 \mid \leqslant a\} + (1-c)\pi_1(\theta)I\{\mid \theta - \theta_0 \mid > a\},$$

其中, $\pi_0(\theta)$ 和 $\pi_1(\theta)$ 是参数分别在 H_0 和 H_1 下的先验分布, I 表示示性函数。

于是,检验的贝叶斯因子以及原假设 H_0 成立的后验概率分别为

$$B(x) = \frac{\int_{\mid \theta-\theta_0 \mid \leqslant a} f(x \mid \theta)\pi_0(\theta)\,\mathrm{d}\theta}{\int_{\mid \theta-\theta_0 \mid > a} f(x \mid \theta)\pi_1(\theta)\,\mathrm{d}\theta}$$

和

$$P(H_0 \mid x) = \left[1 + \frac{1-c}{c}\frac{1}{B(x)}\right]^{-1}.$$

Dempster[28 -29]提出了直接运用原假设与备择假设的似然比来进行单点假设检验的贝叶斯方法。Aitkin[30]对此方法进行推广并给出了另外一种关于单点原假设的评估方法。针对 Aitkin[30]所给检验方法的局限性,Aitkin, Boys 和 Chadwick[31]提出了一种通过后验似然比进行的贝叶斯单点假设检验

的方法。

Basu[32]所给出的基于 HPD 集的后验概率的单点原假设的检验方法,可以避免给单点赋予非零概率质量这一颇有争议的做法。这一检验方法对于解决 DasGupta,Ghosh 和 Zen[33]中的许多有关推断、回归等的问题是很有帮助的。

Pereira 和 Stern[34]则给出一种完全贝叶斯显著性检验(FBST)的方法。FBST 同样遵循 Basu[35],Birnbaum[36],Finetti[37 - 39],Good[40],Kempthorne[41 -42],Royall[43]等将似然作为统计信息的载体的原则。另一方面,FBST 还与 Gaskins[44]及 Kokott[45]的"Onus Proband"判决原则相一致。Madruga,Esteves 和 Wechsler[46]又说明了 FBST 与 Rubin[47]所提出的贝叶斯决策理论的一致性。Madruga,Pereira 和 Stern[48]则研究了单点原假设 FBST 方法,并说明了这一方法在参数空间中仿射变换下的不变性。

另外,作为一种处理讨厌参数的一般方法,Meng[49]将 Rubin[50]的工作加以推广,提出了后验预测 p - 值的概念。

定义 1.5　对于一个原假设 $H_0:\theta \in \Theta_0$,设 $T(X)$ 为一个检验统计量,x^{rep} 表示 x 的一个复制,则检验的后验预测 p - 值定义为

$$p_B(x) = P(T(x^{rep}) \geqslant T(x) \mid x,H_0),$$

式中是在 x^{rep} 的原假设 H_0 的条件下的后验预测分布即

$$f(x^{rep} \mid x,H_0) = \int_{\Theta_0} f(x^{rep} \mid \theta)\Pi_0(\mathrm{d}\theta \mid x)$$

下求概率,其中 $\Pi_0(\theta \mid x)$ 是 θ 在 H_0 下的后验分布。

Meng[49]将后验预测 p - 值的概念应用到关于正态均值以及多重归因推断等典型的单点检验的例子中,并研究了它的频率性质。后验预测 p - 值可以视为经典 p - 值在参数的后验分布下的平均,因此它是一种基于 p - 值检验的准贝叶斯检验方法。相关文献还有 Rubin[50],Meng 和 Rubin[51]以及 Horra 和 Rodriguez - Bernal[52]等。

§1.3　参数消去方法

在频率框架下,我们通过构建广义 p - 值来解决讨厌参数的问题。而有时

为了获得贝叶斯证据,我们也经常受到讨厌参数的困扰,这就使得我们不得不考虑统计推断中的一个至关重要的问题,即参数消去问题。

假设 X 是一个可观测的随机向量,其分布函数和密度函数分别为 $P_\eta(\cdot)$ 和 $f(x \mid \eta)$,其中 $\eta = (\theta,\delta)$ 是参数空间 Ω 中的一个未知参数向量,$\theta = \theta(\eta)$ 为一实值兴趣参数,δ 为讨厌参数。于是 η 的似然函数为 $L(\theta,\delta) = f(x \mid \theta,\delta)$。

Fisher 曾基于构造分布不依赖于讨厌参数的枢轴量来解决讨厌参数存在的问题。然而,处理讨厌参数的较为主流和自然的方法是基于似然的方法。通过关于 Lebesgue 测度进行积分消去讨厌参数 δ 的方法曾被十九世纪主张"逆概率"法的贝叶斯 – Laplace 学派当做是一种明显且合理的方法,这就是所谓的 uniform – integrated 似然

$$L^U(\theta) = \int L(\theta,\delta)\,\mathrm{d}\delta .$$

除了简单积分法,人们也一直在寻求其它更为有效的基于似然的消除讨厌参数的方法。这其中最简单的便是将讨厌参数替换为它们的极大似然估计,从而得到下面的剖面似然

$$\hat{L}(\theta) = \sup_{\delta} L(\theta,\delta) .$$

如果参数的充分统计量的一个或多个分量存在只依赖于兴趣参数 θ 的边际分布或条件分布,此边际分布或条件分布成为 θ 的边际似然或条件似然。这是另外一种似然方法,它实际上是 Cox 所提出的更为一般的偏似然的一种特殊情形。

另外,还有一种通过积分得到的似然

$$L^B(\theta) = \int L(\theta,\delta)\pi(\delta \mid \theta)\,\mathrm{d}\delta ,$$

其中,$\pi(\delta \mid \theta)$ 为给定 θ 下的 δ 的条件先验密度。

显然,上述方法各有利弊,也都存在丢失信息的风险。而我们更加倾向于使用积分似然

$$L^B(\theta) = \int L(\theta,\delta)\pi(\delta \mid \theta)\,\mathrm{d}\delta$$

来处理讨厌参数。这主要基于如下两点。其一,这种积分似然可被视为是一种纯粹的贝叶斯方法;其二,基于上述积分似然得到的后验分布与 θ 的边际后验分布相同。

事实上,基于积分似然得到的 θ 的后验分布可简单表示为

$$\pi(\theta \mid x) = \frac{\pi(\theta)\int \pi(\delta \mid \theta) f(x \mid \theta,\delta)\,\mathrm{d}\delta}{\int \pi(\theta)\int \pi(\delta \mid \theta) f(x \mid \theta,\delta)\,\mathrm{d}\delta \mathrm{d}\theta},$$

另一方面, θ 和 δ 的边际后验分布为

$$\pi(\theta,\delta \mid x) = \frac{\pi(\theta,\delta) f(x \mid \theta,\delta)}{\iint \pi(\theta,\delta) f(x \mid \theta,\delta)\,\mathrm{d}\delta \mathrm{d}\theta}$$

于是 θ 的边际后验分布为

$$\pi^m(\theta \mid x) = \frac{\int \pi(\theta,\delta) f(x \mid \theta,\delta)\,\mathrm{d}\delta}{\iint \pi(\theta,\delta) f(x \mid \theta,\delta)\,\mathrm{d}\delta \mathrm{d}\theta}$$

$\pi(\theta \mid x)$ 和 $\pi^m(\theta \mid x)$ 的等价性是显然的。

上述结论说明,基于 $\pi(\theta \mid x)$ 和 $\pi^m(\theta \mid x)$ 可以给出相同的贝叶斯后验概率。也即虽然出发点不同,我们却得到了相同的结果。为简单起见,在寻求原假设的后验概率时,本书均采用基于 θ 的边际后验密度 $\pi^m(\theta \mid x)$。

§1.4 先验预测 p - 值与后验预测 p - 值

本节介绍两个重要概念,先验预测 p - 值和后验预测 p - 值。它们既可以作为特殊的贝叶斯证据,又可以作为消除讨厌参数的一般方法。

§1.4.1 先验预测 p - 值

给定参数 θ 的先验分布 $\pi(\theta)$,便可给出如下的贝叶斯证据即贝叶斯先验预测 p - 值

$$pp(x) = P^{m(x)}(X \geqslant x)$$

其中

$$m(x) = \int_{\Theta} f(x \mid \theta)\pi(\theta)\,\mathrm{d}\theta$$

通过交换积分顺序可得