第2卷 第2期 高能物理与核物理 Vol. 2, No. 2
1978 年 3 月 PHYSICA ENERGIAE FORTIS ET PHYSICA NUCLEARIS March 1978

参数估计的置信密度方法

李惕碚
(中国科学院高能物理研究所)

摘　　要

在分析参数估计的贝叶斯验后密度方法和非贝叶斯置信区间方法的基础上，由置信区间引出置信密度的概念和计算公式. 置信密度分布是非贝叶斯参数估计结果的完整描述，它有明确的概率意义. 以推断荷电粒子质量为例说明置信密度在高能物理实验数据处理中的应用.

一、引　　言

物理实验的重要任务是测定物理量的数值. 由于测量的偶然误差以及物理现象本身的随机性，实验观测值是随机变量，服从一定的概率分布，被测物理量是观测值概率分布中的参数. 实验只能取得观测值总体的一个样本. 由观测得到的样本推断分布参数的数值，必须利用数理统计学的参数估计方法.

有两种参数估计方法：贝叶斯方法和非贝叶斯方法. 如何评价这两种方法，长期以来就有争议，形成了统计学的两种观点、两个学派. 在实际应用时，这两种方法各有其优点，也各有其局限性.

如果被估计的参数本身也是随机变量，并且已知它的概率分布(验前分布)，用贝叶斯方法推断参数得到参数的验后分布. 验后分布是参数估计结果的一个完整的概率描述，所以贝叶斯方法又叫“概率法”. 观测值概率分布中的参数往往并非实验测定的最终目的，还需要利用参数估计结果作进一步的统计推断(例如误差传播问题)；利用参数的验后密度容易处理这一类问题，这是贝叶斯方法的一个重要优点.

被测物理量常常不是随机变量，而是一个未知的固定值；或者虽然是随机变量，但是不知道它的验前分布的准确形式：要在上述两种情况下应用贝叶斯方法，只有假定一个验前分布. 假定何种验前分布才合理？贝叶斯学派的统计学者们至今还说不清楚，因而验前分布的选取带有某种主观任意性. 对于数值固定的参数，其验后分布没有确切的概率意义.

目前通行的非贝叶斯方法叫做“统计法”或“置信区间法”，用参数的估计值和相应于某个选定置信水平的置信区间来表达参数估计结果. 置信水平和置信区间有确切的统计意义. 置信区间方法仅依据观测值进行推断，毋需假定任何验前的参数分布，这是它的

本文 1977 年 7 月 18 日收到.

主要优点. 但是估计值和置信区间并非被估参数的完整描述；对于很多重要的统计推断问题,置信区间方法难以应用,置信区间法无法严格处理误差传播问题.

在处理高能物理实验中或高能天文观测中的稀有事例这一类问题时，由于观测结果来之不易，要求数据处理方法尽可能严格，处理过程中尽可能不要损失有关的信息. 贝叶斯验后密度方法和非贝叶斯置信区间方法由于它们各自的缺点都不能完全满足上述要求. 因此，实验数学处理工作的实际需要要求有一种参数估计方法能够兼有两种方法的长处：既有非贝叶斯严格从观测数据出发作推断的优点，又能得出类似贝叶斯验后分布那样关于参数值的一个完全的概率描述.

本文首先简要地概述两种参数估计方法(更详细的介绍可以参阅文献[1—3])，在比较这两种方法和讨论两种估计结果的概率意义的基础上，由置信区间引出参数的置信密度的概念. 相应于一种求置信区间的一般方法，导出了一个求置信密度的公式. 类似于贝叶斯验后密度,置信密度是参数估计结果的一个完整的概率描述；同时，它又具有置信区间方法的优点,不需要假定任何验前分布，无论参数是随机变量还是固定值，都有明确的概率意义. 置信密度提供了应用贝叶斯方法时判断假设验前分布形式是否合理的一个标准. 最后,以测定荷电粒子质量的问题为例,说明如何应用置信密度方法实现对于参数推断问题的严格处理.

我们采用符号 $p(x)$ 表示随机变量 x 的概率密度函数. 和普通的函数符号不同,括号前的 p 不限于某个特定的函数形式：随机变量 y 的概率密度函数用 $p(y)$ 表示；两个不同的随机变量 x 和 y 的概率密度函数 $p(x)$ 和 $p(y)$,函数形式不一定相同. 若随机变量 y 是随机变量 x 的单值函数

$$y = y(x), \tag{1.1}$$

则随机变量 x 落入区间 $[x, x + dx]$ 内的概率与随机变量 y 落入相应区间 $[y, y + dy]$ 的概率应该相等：

$$p(x)dx = p(y)dy. \tag{1.2}$$

由此得求随机变量函数分布的公式：

$$p(y) = p(x)\left|\frac{dx}{dy}\right|. \tag{1.3}$$

二、验后密度

本节介绍参数估计的贝叶斯方法.

若需测定的物理量是 θ, 实验观测值为 x. x 是一个随机变量，它的概率密度函数为 $p(x;\theta)$, θ 是观测值概率分布的参数. 如果参数 θ 本身也是一个随机变量,则 $p(x;\theta)$ 就是参数取某一特定值 θ 的条件下随机变量 x 的条件概率密度

$$p(x;\theta) \equiv p(x|\theta). \tag{2.1}$$

实验总共进行 N 次独立观测，得到 N 个具体的观测值 $x_1, x_2, \cdots, x_N$, 即得到随机变量 x 的一个容量为 N 的样本,在参数为特定值 θ 的条件下,出现一组观测结果

$$\mathbf{x} \equiv (x_1, x_2, \cdots, x_N) \tag{2.2}$$

的条件概率密度为

$$p(\mathbf{x}|\theta)=\prod_{i=1}^{N}p(x_i|\theta). \tag{2.3}$$

观测样本的密度函数（2.3）又叫**似然函数**，记作

$$L(\mathbf{x}|\theta)\equiv p(\mathbf{x}|\theta). \tag{2.4}$$

参数 θ 的密度函数 $p(\theta)$ 叫做参数的**验前密度**．验前密度表述实验观测前对于参数 θ 的知识．在取得一组观测结果 $\mathbf{x}=(x_1,\cdots,x_N)$ 的条件下，参数 θ 的条件概率密度 $p(\theta|\mathbf{x})$ 叫做 θ 的**验后密度**，利用关于条件概率的贝叶斯定理可以计算验后密度：

$$p(\theta|\mathbf{x})=\frac{p(\mathbf{x}|\theta)p(\theta)}{\int_{-\infty}^{\infty}p(\mathbf{x}|\theta)p(\theta)d\theta}=\frac{L(\mathbf{x}|\theta)p(\theta)}{\int_{-\infty}^{\infty}L(\mathbf{x}|\theta)p(\theta)d\theta}. \tag{2.5}$$

验后密度分布(2.5)是对于被估参数 θ 的一个完全的概率描述，它综和了实验观测前关于参数值的知识和观测结果中包含的信息，因而是实验观测后对于参数分布全部知识的概括．参数估计结果的完整报导是画出验后密度分布，如图 1．验后密度分布的概率意义是：任意区间 $[\theta_1,\theta_2]$ 上验后密度曲线下面的面积

$$\int_{\theta_1}^{\theta_2}p(\theta|\mathbf{x})d\theta=\xi, \tag{2.6}$$

就是在出现观测结果 $\mathbf{x}=(x_1,\cdots,x_N)$ 的条件下，参数值 θ 落入区间 $[\theta_1,\theta_2]$ 内的概率，即

$$\Pr(\theta_1\leqslant\theta\leqslant\theta_2|\mathbf{x})=\xi. \tag{2.7}$$

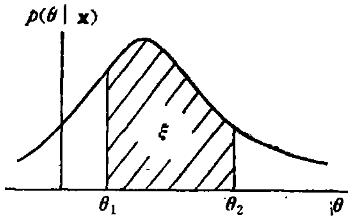

图 1　验后分布和它的概率意义

设想进行很多次实验，每次实验作 N 次观测，得到观测结果的一个序列

$$\mathbf{x}',\mathbf{x}'',\cdots.$$

若每次实验中参数 θ 是一个固定值，对于不同实验 θ 值是随机的，其概率密度为验前分布 $p(\theta)$．选出观测结果恰好是 $\mathbf{x}=(x_1,\cdots,x_N)$ 的那些结果，构成一个子序列．对于这个子序列，θ 的概率密度分布为验后分布 $p(\theta|\mathbf{x})$，θ 的值落入区间 $[\theta_1,\theta_2]$ 内的概率为 ξ．

如果参数并非随机变量，而是一个未知的固定值，或者虽然是随机变量，但不知道验前分布的准确形式，这两种情况下要应用贝叶斯方法推断参数，只有假定一个验前密度．经常采用下述**贝叶斯假设**[2]："如果缺乏参数值的验前知识，则假定它服从在全部可能取值区域上的均匀分布"．利用贝叶斯假设，对于观测结果 $\mathbf{x}$，参数 θ 的验后密度就是归一化后的似然函数：

$$p(\theta|\mathbf{x})=\frac{L(\mathbf{x}|\theta)}{\int L(\mathbf{x}|\theta)d\theta}. \tag{2.8}$$

在参数是固定值时，把它作为随机变量来处理是否合理？这种情况下，验后分布的概率意义是什么？采用均匀分布作为验前分布是否恰当？什么是验前密度假设合理与否的

标准？这些问题在数理统计学中一直争论不休，至今未能完全解决.

所以，在参数是随机变量而且其验前密度已知时，适宜于采用贝叶斯方法求出验后密度. 如果缺乏准确的验前知识，参数的验后密度没有准确的概率意义，并且在选取验前密度时存在着主观任意性.

三、估计值和置信区间

非贝叶斯估计方法采用观测值样本的某个函数作为参数的估计式. 对于一组实验观测值 $\mathbf{x} = (x_1, \cdots, x_N)$，估计式的值

$$\hat{\theta} = \hat{\theta}(x_1, \cdots, x_N) \tag{3.1}$$

就是参数 θ 的**估计值**. 估计值 $\hat{\theta}$ 是随机变量组 $\mathbf{x}$ 的函数，它也是一个随机变量. $\hat{\theta}$ 的概率密度函数 $p(\hat{\theta}; \theta)$ 可以用求随机变量函数分布的方法[1]由观测值的分布 $p(x; \theta)$ 导出.

通常选取使似然函数为最大的参数值作为参数的估计值，叫做参数的**最大似然估计值**[1,3]. 显然，解方程

$$\left.\frac{\partial \ln L(\mathbf{x}|\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}} = 0 \tag{3.2}$$

可求得 θ 的最大似然估计值 $\hat{\theta}$.

估计值 $\hat{\theta}$ 的精确程度（参数估计的误差）用选定置信水平 ξ 下的置信区间 $[\theta_1, \theta_2]$ 来表征. 这样来选择 θ_1 和 θ_2，使得随机区间 $[\theta_1, \theta_2]$ 内包含参数真值的概率为 ξ.

若参数真值 θ 是一个未知的固定值，实验观测取得样本 $\mathbf{x} = (x_1, \cdots, x_N)$，参数的估计值为 $\hat{\theta} = \hat{\theta}(x_1, \cdots, x_N)$，则对于选定的置信水平 ξ，求置信区间的一个办法是：解方程

$$\begin{aligned}\int_{-\infty}^{\hat{\theta}} p(\hat{\theta}; \theta_2)d\hat{\theta} &= 1 - \xi - \varepsilon \\ \int_{\hat{\theta}}^{\infty} p(\hat{\theta}; \theta_1)d\hat{\theta} &= \varepsilon\end{aligned} \tag{3.3}$$

（其中 ε 为任意满足 $0 \leqslant \varepsilon \leqslant 1 - \xi$ 的数），得 $\theta_1 = \theta_1(\mathbf{x})$，$\theta_2 = \theta_2(\mathbf{x})$. 可以证明[1,3]，随机区间 $[\theta_1(\mathbf{x}), \theta_2(\mathbf{x})]$ 满足置信区间的定义：

$$\Pr(\theta_1 \leqslant \theta \leqslant \theta_2|\theta) = \xi. \tag{3.4}$$

即在被估参数是任一固定值 θ 的条件下，随机区间 $[\theta_1(\mathbf{x}), \theta_2(\mathbf{x})]$ 内包含 θ 的概率为 ξ.

上述置信区间方法也适用于参数是随机变量的情况. 若参数 θ 是随机变量，概率密度函数为 $p(\theta)$，则利用全概率定理可得

$$\begin{aligned}\Pr(\theta_1 \leqslant \theta \leqslant \theta_2) &= \int \Pr(\theta_1 \leqslant \theta \leqslant \theta_2|\theta)p(\theta)d\theta \\ &= \xi \int p(\theta)d\theta = \xi.\end{aligned} \tag{3.5}$$

所以，对于随机的参数，非贝叶斯的置信区间 $[\theta_1, \theta_2]$ 和参数为固定值的情况有相同的概率意义. 置信区间 $[\theta_1, \theta_2]$ 与验前密度 $p(\theta)$ 无关，因此非贝叶斯的置信区间方法对参数的随机性质没有任何限制——这是置信区间方法的主要优点.

令 $\Delta\theta_1 = \hat{\theta} - \theta_1$，$\Delta\theta_2 = \theta_2 - \hat{\theta}$，则参数估计结果可以报导为

$$\theta = \hat{\theta}^{+\Delta\theta_2}_{-\Delta\theta_1} \quad (\text{置信水平 } \xi). \tag{3.6}$$

报导(3.6)不是对于被估参数的完整描述．它不像贝叶斯验后密度分布(图 1)那样表达出参数取各种不同值的可能性；除了一个估计值外，就只对于一个选定的置信水平ξ，给出一个置信区间．仅从一个估计值 $\hat{\theta}$，看不出参数取其它值的可能性；只用一个置信区间 $[\theta_1, \theta_2]$，也不能完全表示估计值的精确程度．由于上述弱点，参数 θ 的非贝叶斯估计结果(3.6)很难用来作进一步的统计推断．例如，实验数据处理中常见的误差传播问题：如果由观测值估计的参数是 α 和 β，而需要测定的物理量 θ 是参数 α 和 β 的函数

$$\theta = f(\alpha, \beta),$$

因此必须从对于参数 α 和 β 的估计结果出发，进一步推断 θ．通常的置信区间方法无法准确地处理这一类问题．

四、置 信 密 度

为了更完全地表达估计值的精确程度，可以设定大大小小各不相同的置信水平 ξ'、ξ''，…，求出相应的各个置信区间，用这一系列的置信区间报导对于参数 θ 的估计结果：

$$\theta = \hat{\theta}^{+\Delta\theta_2'}_{-\Delta\theta_1'} \quad (\text{置信水平 } \xi')$$

$$\theta = \hat{\theta}^{+\Delta\theta_2''}_{-\Delta\theta_1''} \quad (\text{置信水平 } \xi'')$$

$$\cdots\ \cdots \qquad \cdots$$

$$\cdots\ \cdots \qquad \cdots$$

这样作，可以部分弥补报导(3.6)的不完全．

被估参数的完整描述应该是类似于贝叶斯验后分布(图 1)那样一个关于参数值的概率分布．下面，我们引入置信分布函数作为对参数非贝叶斯估计结果的一个完整概括，然后使置信区间无穷小化导出置信密度分布．

若参数的估计值为 $\hat{\theta}$，选定置信水平 ξ，用区间估计方法可以得到置信区间 $[\theta_1, \theta_2]$．固定置信区间的下确界 $\theta_1 = -\infty$，则区间 $(-\infty, \theta]$ 的置信水平 ξ_θ 可以视为区间上确界 θ 的函数

$$\xi_\theta = \Xi(\theta). \tag{4.1}$$

由置信水平的定义，ξ_θ 是随机区间 $(-\infty, \theta]$ 内包含参数真值的概率，显然 $\Xi(\theta)$ 是单调非减函数，并且

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

$\Xi(\theta)$ 至少是右连续的，因为对于任意 $\theta' > \theta$，$\Xi(\theta') - \Xi(\theta)$ 是半开区间 $(\theta, \theta']$ 的置信水平，即随机区间 $(\theta, \theta']$ 内包含参数真值 θ_0 的概率

$$\Xi(\theta') - \Xi(\theta) = \Pr(\theta < \theta_0 \leqslant \theta'),$$

则

$$\lim_{\theta' \to \theta+0} [\Xi(\theta') - \Xi(\theta)] = \lim_{\theta' \to \theta+0} \Xi(\theta') - \Xi(\theta) = \lim_{\theta' \to \theta+0} \Pr(\theta < \theta_0 \leqslant \theta') = 0,$$

所以

$$\lim_{\theta'\to\theta+0} \Xi(\theta') = \Xi(\theta).$$

函数 $\Xi(\theta)$ 满足分布函数的充要条件，我们把 $\Xi(\theta)$ 叫做**置信分布函数**，记作

$$\mathrm{P}_c(\theta|\hat{\theta}) \equiv \Xi(\theta). \tag{4.2}$$

如果置信分布函数是连续型分布函数，存在一个非负函数 $p_c(\theta|\hat{\theta})$，对于任意实数 θ，有

$$\mathrm{P}_c(\theta|\hat{\theta}) = \int_{-\infty}^{\theta} p_c(\theta|\hat{\theta})d\theta, \tag{4.3}$$

则把函数

$$p_c(\theta|\hat{\theta}) = \frac{\partial \mathrm{P}_c(\theta|\hat{\theta})}{\partial\theta} = \frac{\partial\Xi(\theta)}{\partial\theta} \tag{4.4}$$

叫做参数 θ 的**置信密度***.

显然，任意区间 $[\theta_1, \theta_2]$ 上置信密度曲线(图 2)下的面积 ξ 就是置信区间 $[\theta_1, \theta_2]$ 的置信水平

$$\xi = \int_{\theta_1}^{\theta_2} p_c(\theta|\hat{\theta})d\theta. \tag{4.5}$$

任意 θ 值附近一个无穷小区间 $[\theta, \theta + d\theta]$ 的置信水平 $d\xi$ 为

$$d\xi = p_c(\theta|\hat{\theta})d\theta. \tag{4.6}$$

置信分布函数 $\mathrm{P}_c(\theta|\hat{\theta})$ 或置信密度 $p_c(\theta|\hat{\theta})$ 是非贝叶斯参数估计结果的一个完整概括. $\mathrm{P}_c(\theta|\hat{\theta})$ 和 $p_c(\theta|\hat{\theta})$ 不仅依赖于估计值，而且依赖于选取何种估计式以及如何由观测样本和估计式求置信区间，即依赖于参数估计的方法和结果. 记号 $\mathrm{P}_c(\theta|\hat{\theta})$ 和 $p_c(\theta|\hat{\theta})$ 中的条件 $\hat{\theta}$ 不仅表示估计值，而且应该理解为代表整个参数估计方法.

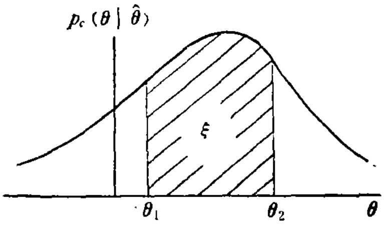

图 2 置信密度曲线

如果置信区间是用解方程(3.3)的方法求得的，估计式的分布对于任何 $\hat{\theta} > -\infty$ 满足

$$\lim_{\theta\to-\infty} \int_{\hat{\theta}}^{\infty} p(\hat{\theta}; \theta)d\hat{\theta} = 0,$$

则由 (3.3)，对于区间 $(-\infty, \theta]$ 的置信水平 ξ_θ，有

$$\int_{-\infty}^{\hat{\theta}} p(\hat{\theta}; \theta)d\hat{\theta} = 1 - \xi_\theta,$$

由 (4.1)、(4.2) 和 (4.4)，相应于求置信区间的这种方法，置信分布函数为

$$\mathrm{P}_c(\theta|\hat{\theta}) = 1 - \int_{-\infty}^{\hat{\theta}} p(\hat{\theta}; \theta)d\hat{\theta}, \tag{4.7}$$

置信密度为

$$p_c(\theta|\hat{\theta}) = \frac{\partial \mathrm{P}_c(\theta|\hat{\theta})}{\partial\theta} = -\int_{-\infty}^{\hat{\theta}} \frac{\partial p(\hat{\theta}; \theta)}{\partial\theta} d\hat{\theta}. \tag{4.8}$$

* 对于离散型的置信分布函数，参数只能取有限个或可数个数值 $\theta_i (i = 1, 2, \cdots)$，可定义置信概率函数

$$p_c(\theta_i|\hat{\theta}) = \Xi(\theta_i) - \Xi(\theta_{i-1}). \tag{4.4'}$$

本文仅讨论连续型的情况.

置信密度分布有明确的概率意义：估计值 $\hat{\theta}$ 是随机变量，因而对于不同的观测结果，置信密度曲线的形状也各不相同，可以把置信密度曲线视为“随机曲线”．选定置信水平 ξ 或 $d\xi$，由各条曲线得到满足（4.5）或（4.6）的很多区间 $[\theta_1, \theta_2]$ 或 $[\theta, \theta+d\theta]$，其中平均有份额为 ξ 或 $d\xi$ 的区间包含参数真值（为了保持置信水平 $d\xi$ 不变，无穷小区间的“长度” $d\theta$ 是随机的，反比于置信密度 $p_c(\theta|\hat{\theta})$）．置信密度的这种频率解释可以通过大量实验观测进行验证．

贝叶斯验后密度曲线是一条固定的曲线，图 1 所示贝叶斯置信区间 $[\theta_1, \theta_2]$ 是一个固定区间．验后分布 $p(\theta|\mathbf{x})$ 是对于以出现观测结果 $\mathbf{x}$ 为标准挑选出来的子序列，随机变量 θ 的概率密度分布．测得一组特定观测值 $\mathbf{x}$ 的条件下，参数值落入固定区间 $[\theta_1, \theta_2]$ 内的概率

$$\Pr(\theta_1 \leqslant \theta \leqslant \theta_2|\mathbf{x}) = \xi.$$

对于连续随机变量 x，不可能严格地挑选出观测结果恰好为 $\mathbf{x}=(x_1, \cdots, x_N)$ 的子序列[1]．因此，验后密度的概率解释中应该去掉特定观测结果的限制：将 $[\theta_1, \theta_2]$ 视为随机区间，它的范围依赖于观测结果——对于不同观测结果 $\mathbf{x}$，有不同的验后密度分布 $p(\theta|\mathbf{x})$，因而验后密度曲线也是“随机曲线”，满足条件（2.6）的区间 $[\theta_1, \theta_2]$ 范围也不同．对于所有可能的观测结果，随机区间 $[\theta_1, \theta_2]$ 包含参数值 θ 的概率也等于 ξ，因为由全概率定理有：

$$\Pr(\theta_1 \leqslant \theta \leqslant \theta_2) = \sum_{\mathbf{x}} \Pr(\theta_1 \leqslant \theta \leqslant \theta_2|\mathbf{x})\Pr(\mathbf{x}) = \xi \sum_{\mathbf{x}} \Pr(\mathbf{x}) = \xi. \tag{4.9}$$

式中 $\sum_{\mathbf{x}}$ 表示对所有不同的观测结果求和．

所以，只有在统计的意义上，将贝叶斯验后密度曲线视为“随机曲线”，将贝叶斯置信区间视为随机区间，贝叶斯估计结果才有确切的可以用实验予以验证的概率解释．

置信密度是非贝叶斯参数估计的一个完整报导．非贝叶斯的置信密度 $p_c(\theta|\hat{\theta})$ 具有与贝叶斯的验后密度 $p(\theta|\mathbf{x})$ 类似的概率意义，但是对于参数 θ 的随机性质没有任何限制：同参数的验前分布无关；适用于随机的参数，也适用于参数是未知的固定值．

如果对参数 θ 用贝叶斯方法作估计，并且取验前分布为

$$p(\theta) = \frac{p_c(\theta|\hat{\theta})}{L(\mathbf{x}|\theta)} \Big/ \int \frac{p_c(\theta|\hat{\theta})}{L(\mathbf{x}|\theta)} d\theta, \tag{4.10}$$

则贝叶斯的验后密度与非贝叶斯的置信密度一致：因为，将（4.10）代入（2.5），有

$$p(\theta|\mathbf{x}) = \frac{L(\mathbf{x}|\theta)p(\theta)}{\int L(\mathbf{x}|\theta)p(\theta)d\theta} = p_c(\theta|\hat{\theta}). \tag{4.11}$$

所以，任何一个非贝叶斯推断都对应于采用了某种验前分布的一个贝叶斯推断．

在缺乏参数验前知识的情况下应用贝叶斯方法，必须假定一个验前分布．由假设验前分布得出的验后密度与某个非贝叶斯估计的置信密度一致，可以作为假设验前分布合理性的一个标准．

若观测值 x 服从正态分布，方差 σ^2 已知，被测物理量 θ 是观测值分布的期待值，即

$$p(x;\theta)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2\right]\equiv n(x;\theta,\sigma^2). \tag{4.12}$$

实验观测得样本 $\mathbf{x}=(x_1,\cdots,x_N)$. θ 的最大似然估计值 $\hat{\theta}$ 是样本平均值

$$\hat{\theta}=\bar{x}=\frac{\sum_{i=1}^{N}x_i}{N}, \tag{4.13}$$

正态样本平均值 $\bar{x}$ 服从期待值 θ 方差 $\frac{\sigma^2}{N}$ 的正态分布,所以

$$p(\hat{\theta};\theta)=n\left(\hat{\theta};\theta,\frac{\sigma^2}{N}\right)=\frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{N}{2}\left(\frac{\hat{\theta}-\theta}{\sigma}\right)^2\right].$$

用公式 (4.8) 求 θ 的置信密度:

$$\begin{aligned}p_c(\theta|\hat{\theta})&=-\int_{-\infty}^{\hat{\theta}}\frac{\partial}{\partial\theta}p(\hat{\theta};\theta)d\hat{\theta}=\frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{N}{2}\left(\frac{\theta-\hat{\theta}}{\sigma}\right)^2\right]\\&=n\left(\theta;\hat{\theta},\frac{\sigma^2}{N}\right)=n\left(\theta;\bar{x},\frac{\sigma^2}{N}\right).\end{aligned} \tag{4.14}$$

如果用贝叶斯方法估计 θ, 并且采用贝叶斯假设,令验前分布 $p(\theta)$ 为 $(-\infty,\infty)$ 区间上的均匀分布,则由 (2.8), θ 的验后密度为:

$$\begin{aligned}p(\theta|\mathbf{x})&=\frac{L(\mathbf{x}|\theta)}{\int L(\mathbf{x}|\theta)d\theta}=\frac{\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma}\right)^2\right]}{\int_{-\infty}^{\infty}\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma}\right)^2\right]d\theta}\\&=\frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{N}{2}\left(\frac{\theta-\bar{x}}{\sigma}\right)^2\right]=n\left(\theta;\bar{x},\frac{\sigma^2}{N}\right).\end{aligned} \tag{4.15}$$

采用均匀验前密度,正态分布期待值的验后密度 (4.15) 与非贝叶斯的置信密度 (4.14) 一致. 因此,在估计正态分布的期待值时,采用贝叶斯假设是合适的.

五、荷电粒子质量测定

本节以从动量和游离观测推断荷电粒子质量的问题为例,说明置信密度方法在高能物理实验数据处理中的应用.

在观测基本粒子的实验中,常利用公式

$$\mathrm{M}=\frac{\mathrm{p}}{\beta\gamma} \tag{5.1}$$

测定荷电粒子质量 M. p 是粒子动量, $\beta=\frac{v}{c}$, v 是粒子速度, c 是光速, $\gamma=\frac{1}{\sqrt{1-\beta^2}}$. 例如,测定形成磁云室照片上某根径迹的粒子质量的办法是: 测量径迹曲率以导出粒子动量估计值 $\hat{\mathrm{p}}$, 测量径迹上的水珠数以导出粒子速度的估计值 $\widehat{\beta\gamma}$ (本文中把 $\beta\gamma$ 简称为速

度），然后计算粒子质量的估计值 $\hat{M}=\dfrac{\hat{p}}{\widehat{\beta\gamma}}$. 质量估计值的标准误差 $\sigma(\hat{M})$ 通常利用所谓误差传播公式由动量估计值的标准误差 $\sigma(\hat{p})$ 和速度估计值的标准误差 $\sigma(\widehat{\beta\gamma})$ 导出：

$$\frac{\sigma(\hat{M})}{M}\simeq\sqrt{\left(\frac{\sigma(\hat{p})}{p}\right)^2+\left(\frac{\sigma(\widehat{\beta\gamma})}{\beta\gamma}\right)^2}. \tag{5.2}$$

使用误差传播公式的条件是

$$\frac{\sigma(\hat{p})}{p}\ll 1,\quad \frac{\sigma(\widehat{\beta\gamma})}{\beta\gamma}\ll 1.$$

在高能物理实验中，上述条件经常不能满足. 例如云南高山站的大型磁云室拍得的一张宇宙线高能作用事例[4]，其中 c 粒子的动量估计值 $\hat{p}=110\,\mathrm{GeV}/c$，而云室最大可测动量 $p_{max}=48\,\mathrm{GeV}/c$. $\dfrac{\sigma(\hat{p})}{p}\simeq\dfrac{\hat{p}}{p_{max}}>1$,不满足使用误差传播公式的条件. 这种情况下要推断粒子质量,必须应用数理统计学的参数估计方法[5].

待测定的某个荷电粒子质量M是一个未知的固定值，应该用非贝叶斯估计方法. 但是,要从对 p 和 $\beta\gamma$ 这两个量的估计结果利用函数关系(5.1)进一步推断质量 M，置信区间方法无能为力. 应用置信密度方法可以对粒子质量作出准确和完整的概率推断.

1. 由动量和速度的置信密度求质量的置信密度

如果粒子速度 $\beta\gamma$ 已知,则由 (5.1),动量 p 和质量M是一一对应的函数

$$p=M\cdot\beta\gamma.$$

在 $\beta\gamma$ 已知的条件下，对于动量估计值 $\hat{p}$，质量M的置信密度 $p_c(M|\hat{p},\beta\gamma)$ 可以由动量 p 的置信密度 $p_c(p|\hat{p})$ 利用求函数分布的公式 (1.3) 导出：

$$p_c(M|\hat{p},\beta\gamma)=p_c(p|\hat{p})\left|\frac{dp}{dM}\right|=p_c(p|\hat{p})\beta\gamma. \tag{5.3}$$

用求全概率的办法,可得 $\beta\gamma$ 未知的情况下,对于动量和速度估计值 $\hat{p}$ 和 $\widehat{\beta\gamma}$，质量 M 的置信密度

$$\begin{aligned}p_c(M|\hat{p},\widehat{\beta\gamma})&=\int_{-\infty}^{\infty}p_c(M|\hat{p},\beta\gamma)p_c(\beta\gamma|\widehat{\beta\gamma})d(\beta\gamma)\\&=\int_{-\infty}^{\infty}p_c(\beta\gamma|\widehat{\beta\gamma})p_c(p=M\beta\gamma|\hat{p})\beta\gamma d(\beta\gamma).\end{aligned} \tag{5.4}$$

由(5.4)算得的置信密度分布 $p(M|\hat{p},\widehat{\beta\gamma})$,就是根据动量和速度的测量结果 $\hat{p}$ 和 $\widehat{\beta\gamma}$，对于粒子质量M的一个完整的概率估计.

由(4.5)，任意质量区间 $[M_1,M_2]$ 的置信水平，即区间 $[M_1,M_2]$ 内包含质量真值M的概率为

$$\begin{aligned}\Pr(M_1\leqslant M\leqslant M_2)&=\int_{M_1}^{M_2}p_c(M|\hat{p},\widehat{\beta\gamma})dM\\&=\int_{-\infty}^{\infty}p_c(\beta\gamma|\widehat{\beta\gamma})\left[\int_{M_1}^{M_2}p_c(p=M\beta\gamma|\hat{p})\beta\gamma dM\right]d(\beta\gamma)\\&=\int_{-\infty}^{\infty}p_c(\beta\gamma|\widehat{\beta\gamma})\left[\int_{p_1}^{p_2}p_c(p|\hat{p})dp\right]d(\beta\gamma);\end{aligned}$$

$$p_1 = M_1\beta\gamma,$$
$$p_2 = M_2\beta\gamma. \tag{5.5}$$

动量 p 和速度 $\beta\gamma$ 都不是直接观测量，它们的置信密度 $p_c(p|\hat{p})$ 和 $p_c(\beta\gamma|\widehat{\beta\gamma})$ 还必须从直接观测量的分布导出.

2. 动量估计

若荷电粒子在磁场 $H(G)$ 中的曲率为 C (cm^{-1})，则粒子动量

$$p = 300H\frac{1}{C}\,(\mathrm{eV}/c).$$

扣除系统误差后，粒子径迹的曲率观测值 $\hat{C}$ 服从正态分布

$$p(\hat{C};\ C) = n(\hat{C};\ C,\ \sigma_c^2).$$

C 是曲率真值，σ_c 是曲率观测的标准误差，由测量无磁场径迹的曲率分布定出. 前面已经讲过，正态分布 (4.12) 期待值的置信密度是正态密度 (4.14)，所以曲率 C 的置信密度是

$$p_c(C|\hat{C}) = n(C;\ \hat{C},\ \sigma_c^2).$$

对于动量估计值 $\hat{p} = 300H\frac{1}{\hat{C}}$，由 (1.2) 可得

$$\begin{aligned}
p_c(p|\hat{p})dp &= p_c(C|\hat{C})dC = n(C;\ \hat{C},\ \sigma_c^2)dC \\
&= \frac{1}{\sqrt{2\pi}\sigma_c}\exp\left[-\frac{1}{2}\left(\frac{C-\hat{C}}{\sigma_c}\right)^2\right]dC \\
&= n(x;\ 0,\ 1)dx; \\
x &= \frac{\hat{C}-C}{\sigma_c} = \frac{p_{max}}{\hat{p}} - \frac{p_{max}}{p}, \\
p_{max} &= 300H\frac{1}{\sigma_c}.
\end{aligned} \tag{5.6}$$

利用上式，可以对粒子动量作统计推断：在动量区间 $[p_1, p_2]$ 中包含粒子动量真值 p 的概率

$$\begin{aligned}
\Pr(p_1 \leqslant p \leqslant p_2) &= \int_{p_1}^{p_2} p_c(p|\hat{p})dp = \int_{x_1}^{x_2} n(x;\ 0,\ 1)dx \\
&= N(x_2;\ 0,\ 1) - N(x_1;\ 0,\ 1); \\
x_1 &= \frac{p_{max}}{\hat{p}} - \frac{p_{max}}{p_1}, \\
x_2 &= \frac{p_{max}}{\hat{p}} - \frac{p_{max}}{p_2}; \\
N(x;\ 0,\ 1) &= \int_{-\infty}^{x} n(x;\ 0,\ 1)dx.
\end{aligned} \tag{5.7}$$

标准正态分布函数 $N(x;\ 0,\ 1)$ 的数值可以在数理统计用表中查出.

3. 速度估计

荷电粒子速度观测值 $\widehat{\beta\gamma}$ 系由一定长度上云室径迹水珠数目观测值 $\hat{m}$ 导出. 水珠数目的期待值 m 和速度绝对值 $|\beta\gamma|$ 的函数关系，即所谓“游离曲线”

$$m = m(|\beta\gamma|) \tag{5.8}$$

与云室气体成份和工作条件有关，由拟合大量已知速度粒子的水珠数目观测值得到．对于一定速度的粒子，若它的水珠数目期待值为 m，且 $m \gg 1$，则实际观测到的水珠数 $\hat{m}$ 近似服从正态分布，分布期待值为 m，方差正比于 m：

$$p(\hat{m};\, m) \simeq n(\hat{m};\, m,\, km). \tag{5.9}$$

参数 k 在游离曲线拟合时一并定出[6,7]．

对于水珠数目观测值（即其估计值）$\hat{m}$，期待值 m 的置信密度，可利用求置信密度的公式（4.8）计算：

$$\begin{aligned} p_c(m \mid \hat{m}) &= -\int_{-\infty}^{\hat{m}} \frac{\partial}{\partial m} p(\hat{m};\, m)\, d\hat{m} \\ &\simeq -\int_{-\infty}^{\hat{m}} \frac{\partial}{\partial m} \left\{ \frac{1}{\sqrt{2\pi}\,\sqrt{km}} \exp\left[-\frac{1}{2}\left(\frac{\hat{m}-m}{\sqrt{km}} \right)^2 \right] \right\} d\hat{m} \\ &= \frac{1}{2}\left(1+\frac{\hat{m}}{m}\right) \frac{1}{\sqrt{2\pi}\,\sqrt{km}} \exp\left[-\frac{1}{2}\left(\frac{m-\hat{m}}{\sqrt{km}} \right)^2 \right]. \end{aligned} \tag{5.10}$$

4. 质量推断

利用（5.10）和（5.8），可将求质量区间 $[\mathrm{M}_1, \mathrm{M}_2]$ 置信水平的公式（5.5）写为

$$\begin{aligned} \Pr(\mathrm{M}_1 \leqslant \mathrm{M} \leqslant \mathrm{M}_2) &= \int_{\mathrm{M}_1}^{\mathrm{M}_2} p_c(\mathrm{M} \mid \hat{\mathrm{p}},\, \widehat{\beta\gamma})\, d\mathrm{M} \\ &= \int_0^{\infty} p_c(m \mid \hat{m}) \Pr(\mathrm{p}_1 \leqslant |\mathrm{p}| \leqslant \mathrm{p}_2)\, dm; \\ \Pr(\mathrm{p}_1 \leqslant |\mathrm{p}| \leqslant \mathrm{p}_2) &= \int_{\mathrm{p}_1}^{\mathrm{p}_2} p_c(\mathrm{p} \mid \hat{\mathrm{p}})\, d\,\mathrm{p} + \int_{-\mathrm{p}}^{-\mathrm{p}_1} p_c(\mathrm{p} \mid \hat{\mathrm{p}})\, d\,\mathrm{p}, \\ \mathrm{p}_1 &= \mathrm{M}_1 \cdot |\beta\gamma(m)|, \\ \mathrm{p}_2 &= \mathrm{M}_2 \cdot |\beta\gamma(m)|. \end{aligned} \tag{5.11}$$

式中，$\beta\gamma(m)$ 表示与 m 值对应满足函数关系（5.8）的 $\beta\gamma$ 值．

将（5.6）、（5.10）代入（5.11），对 m 积分（可以用数值积分方法完成），可以计算出任意质量区间 $[\mathrm{M}_1, \mathrm{M}_2]$ 的置信水平．

若由（5.11）计算得到区间 $[\mathrm{M}_1, \mathrm{M}_2]$ 的置信水平为 ξ，即

$$\Pr(\mathrm{M}_1 \leqslant \mathrm{M} \leqslant \mathrm{M}_2) = \int_{\mathrm{M}}^{\mathrm{M}_2} p_c(\mathrm{M} \mid \hat{\mathrm{p}},\, \widehat{\beta\gamma})\, d\mathrm{M} = \xi. \tag{5.12}$$

上述结果的统计意义是：设想测量很多粒子径迹，其动量和速度观测值分别为 $(\hat{\mathrm{p}}', \widehat{\beta\gamma'})$，$(\hat{\mathrm{p}}'', \widehat{\beta\gamma''})$，$\cdots$．对于每组观测值，用上述方法分别找出满足（5.12）的区间 $[\mathrm{M}_1', \mathrm{M}_2']$，$[\mathrm{M}_1'', \mathrm{M}_2'']\cdots$，则其中平均有份额为 ξ 的区间包含了被测粒子的质量真值．无论这些被测量的粒子质量相同还是质量不同，上述断言都是正确的．

志谢：作者在与吴枚、徐春娴、马宇蒨等同志的多次讨论中得到不少启发．

参　考　资　料

[1] 李惕碚，《实验的数学处理》，科学出版社（将出版）．

[2] M. 费史，《概率论及数理统计》，王福保译，上海科学技术出版社（1962）．

[3] A. Mood and F. Graybill, *《Introduction to the Theory of Statistics》*, McGraw-Hill Book Company, Inc., NewYork (1963).

[4] 中国科学院原子能研究所云南站,物理, 1 (1972), 57.
[5] 李惕碚,16580 事例 c 粒子质量的统计推断, (1972),未发表.
[6] 李惕碚,游离测量的数学拟合,(1973),未发表.
[7] 中国科学院高能物理研究所三室,μ 介子在氦氩混合气体中的游离能损曲线,(1974),未发表.

CONFIDENCE DENSITY METHOD FOR PARAMETER ESTIMATION

LI TI-PEI

(*Institute of High Energy Physics, Academia Sinica*)

ABSTRACT

Based on an analysis of the Bayesian posteriori density method and the Non-Bayesian confidence interval method for parameter estimation, the concept and a computation formula for the confidence density are introduced from the confidence interval. It is shown that the confidence density distribution is a complete description of the result of the Non-Bayesian parameter estimation and possesses a definite probability meaning. As an example to show its application, a statistical inference to the mass of a charged particle is made.