试验数据处理措施
第三部分：统计学措施

第十二章 最大似然法
(Maximum Likelihood method)

第十二章 最大似然法
(Maximum Likelihood Method)

点估计的措施之一，是参数估计中常用的措施，具有下列的特点：

1. 在一定的条件下，ML估计式满足一致性、无偏性、有效性等要求；
2. 当样本容量n→∞时，ML估计式满足正态分布→方差轻易计算；
3. 用ML措施可较轻易地得到参数的估计式；

本章内容：

1. 最大似然原理；
2. 用ML措施求解参数估计问题的环节；
3. ML估计式的特征；
4. 怎样计算ML估计值的方差；
5. 利用似然函数进行区间估计

第十二章 最大似然法
(Maximum Likelyhood Method)

12.1 最大似然原理

12.1 最大似然原理

(一) 似然函数的定义

p.d.f： f(x|θ)

测量量： $\underline{x}$ = {x1, x2, …, xn }

$$L(\underline{x}\,|\,\theta) = \prod_{i=1}^{n} f(x_i\,|\,\theta)$$

$$\int_{\Omega} L(\underline{x}\,|\,\theta)\,d\underline{x} = 1$$

(二) 最大似然原理

未知参数θ的最佳估计值$\hat{\theta}$应满足如下的条件：

i. $\hat{\theta}$位于θ的允许取值范围；

ii. 对于给定的一组测量值，$\hat{\theta}$使L取极大值：

$$L(\underline{x}\,|\,\hat{\theta}) \geq L(\underline{x}\,|\,\theta)$$

12.1 最大似然原理

(三)估计值 $\hat{\theta}$的求法

似然方程： $\frac{\partial L(\underline{x}|\theta)}{\partial\theta}=\frac{\partial}{\partial\theta}\prod_{i=1}^{n}f(x_i,\theta)=0$

极大值条件： $\left.\frac{\partial^2 L(\underline{x}|\theta)}{\partial\theta^2}\right|_{\theta=\hat{\theta}}<0$

因为lnL是L的单调上升函数，lnL和L具有相同的极大值点，所以，L→lnL， 求和运算比乘积运算轻易处理

似然方程： $\frac{\partial\ln L(\underline{x}|\theta)}{\partial\theta}=\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\ln f(x_i,\theta)=0$

极大值条件： $\left.\frac{\partial^2\ln L(\underline{x}|\theta)}{\partial\theta^2}\right|_{\theta=\hat{\theta}}<0$

假如有k个位置参数，$\underline{\theta}=\{\theta_1,\theta_2,\ldots,\theta_k\}$

→k阶似然方程

$$\frac{\partial\ln L(\underline{x}|\underline{\theta})}{\partial\theta_j}=\frac{\partial}{\partial\theta_j}\sum_{i=1}^{n}\ln f(x_i,\underline{\theta})=0\quad j=1,2,\mathrm{L}\ ,k$$

估计值： $\hat{\underline{\theta}}=\{\hat{\theta}_1,\hat{\theta}_2,\mathrm{L}\ ,\hat{\theta}_k\}$

12.1 最大似然原理

极大值条件：二次矩阵$U(\hat{\theta})$是负定的(Negative definite)

$$U_{ij}(\hat{\theta}) = \frac{\partial^2 \ln L(\underline{x} \mid \underline{\theta})}{\partial \theta_i \partial \theta_j} \Big|_{\underline{\theta}=\underline{\hat{\theta}}}$$

第十二章 最大似然法
(Maximum Likelyhood Method)

12.2 用ML措施进行参数估计的环节

12.2 用ML措施进行参数估计的环节

1) 构造概率密度函数；
2) 构造似然函数；
3) 求似然函数的极大值。

12.2 用ML措施进行参数估计的环节

（一）构造概率密度函数

物理系统的特征：某些量的理论概率分布函数
试验的条件：辨别率、探测效率

例：不变质量谱分析：$e^+e^- \rightarrow J/\psi \rightarrow \gamma K^+K^-$

- 经过测量K^+K^-的动量，可得到K^+K^-的不变质量分布，对该分布进行统计分析，可得到衰变过程中产生的共振态的信息；
- 描述不变质量m的分布的p.d.f应包括对该分布有贡献的物理过程

12.2 用ML措施进行参数估计的环节

1. 信号事例：

$$J/\psi \to \gamma X$$
$$\qquad\qquad \hookrightarrow K^+K^-$$

在不变质量为m_0处出现共振态X的弹性散射振幅可用Breit-Wigner公式描述：

$$BW = \frac{\sqrt{\Gamma}}{m - m_0 + i\Gamma/2}$$

Γ：X的宽度，m_0：X的静质量，m：K+K-的不变质量

（1）假如Γ较小

$$|BW| = \frac{\Gamma^2}{(m-m_0)^2 + \Gamma^2/4}$$

试验成果包括质量辨别率σ和探测效率的影响， Γ~ σ，故必须对理论公式进行修正

$$|BW|^2 \to \int |BW|^2 \eta(m) R(m, m') dm'$$

12.2 用ML措施进行参数估计的环节

其中：

η(m)：效率函数，因η(m)随m的变化较小，故η(m)~常数
R(m,m′)：辨别率函数，真值为m时，取得测量值m′的概率

$$R(m,m')=\frac{1}{\sqrt{2\pi\sigma}}\exp[-\tfrac{1}{2}(m-m')^2/\sigma^2]$$

σ：质量辨别率

所以，窄共振峰的p.d.f为

$$\int\left|BW\right|^2 R(m,m')dm'=\frac{1}{\sqrt{2\pi\sigma}}\mathrm{Re}(w(z))$$

$$w(z)=e^{-z^2}erfc(-iz)$$

$$z=\frac{m-m_0}{\sqrt{2}\sigma}+i\frac{\Gamma}{2\sqrt{2}\sigma}$$

12.2 用ML措施进行参数估计的环节

（1）假如Γ较大，宽共振峰

因为Γ>> σ，所以R(m,m´)~ δ(m-m´)

假如在衰变过程中存在着多种宽共振，则可能存在仙湖干涉的现象，设有N_{amp}个相干的共振峰，则描述这些共振峰的p.d.f为

$$\sim \left| BW_1 + \sum_{k=2}^{N_{amp}} \sqrt{\delta_{k-1}} e^{-\varphi_{k-1}} BW_k \right|^2$$

$$BW_k = \frac{\sqrt{\Gamma_k}}{m - m_{0k} + i\Gamma/2}$$

φk-1：相位差

δk-1：第k个相干的共振峰事例数/第一种相干的共振峰的事例数

12.2 用ML措施进行参数估计的环节

2. 本底事例：相空间本底、粒子误判本底、其他衰变道本底等

$$f_{back}(m,\theta) \sim f_{ps}(m,\theta)\frac{1}{m_{\max}-m_{\min}}\left[1+\sum_{i=1}^{N_b} b_i P^i(x)\right]$$

$f_{ps}(m,\theta)$：相空间函数

$P^i(x)$：**i阶Legendre多项式**

$$x=-1+\frac{m-m_{\min}}{m_{\max}-m_{\min}}$$

bi：未知参数

12.2 用ML措施进行参数估计的环节

假如衰变过程中：N_{BW}个窄共振峰、N_{amp}个相干共振峰，则m的pdf

$$f(m|\theta)=\sum_{k=1}^{N_{bw}}\alpha_m\cdot\frac{1}{\sqrt{2\pi}\sigma_m}\mathrm{Re}(W(z))/C_{BW}$$

$$+\beta\cdot\left|BW_1+\sum_{k=2}^{N_{amp}}\sqrt{\delta_{k-1}}e^{-i\phi_{k-1}}BW_k\right|^2\Big/C_{amp}$$

$$+(1-\sum_{k-1}^{N_{BW}}\alpha_k-\beta)f_{back}\Big/C_{back}$$

其中：C_{BW}、C_{amp}、C_{back}为归一化常数，确保$\int f(m|\theta)dm=1$

α_k　：第k个窄共振峰事例数/总事例数

β　：Namp个相干共振峰事例数/总事例数

➔

（二）构造似然函数

12.2 用ML措施进行参数估计的环节

设对某物理系统进行了n次测量，x_1、x_2、…x_n

$$\ln L(\underline{x}|\underline{\theta}) = \sum_{i=1}^{n} \ln f(x_i|\underline{\theta})$$

根据需要可对 $L(\underline{x}|\underline{\theta})$ 进行变化：

1. 广义似然函数（Generalized Likelihood Function）

总事例数n也是随机变量，服从平均值为υ的泊松分布：

- 在试验条件一定的条件下，事例的产生率为常数，
- 在时间t内取得n个事例的概率为泊松分布。

观察到n个事例，且测量量为x_1、x_2、…x_n的联合概率为

$$L(n,\underline{x}|\underline{\theta}) = \frac{\nu^n e^{-\nu}}{n!} L(\underline{x}|\underline{\theta}) = \frac{\nu^n e^{-\nu}}{n!} \prod_{i=1}^{n} f(x_i|\underline{\theta})$$

➔广义似然函数， $\nu = \nu(\underline{\theta})$

优点：n对θ增长了附加的限制

条件：ν必须能够精确拟定

12.2 用ML措施进行参数估计的环节

2. 数据分类情况下的似然函数

对试验数据进行分间隔处理，（如作成直方图）然后用ML措施对分类后的数据进行处理。

优点：减小了数据量，使得对L 的计算速度加紧

缺陷：因为将原L 简化为少许的几种“平均”pdf的乘积，使得参数估计的精度下降。

设将x的变化范围提成了N个间隔

n_i ：第i个间隔内的事例数

$$\sum_{i=1}^{N} n_i = n$$

p_i ：某事例落入第i个间隔的概率

N个事例分布于N个间隔内，每个间隔内的事例数为n_1、n_2、...n_N的概率满足多项式分布：