

粒子物理与核物理实验中的 数据分析

陈少敏
清华大学

第十讲：矩方法

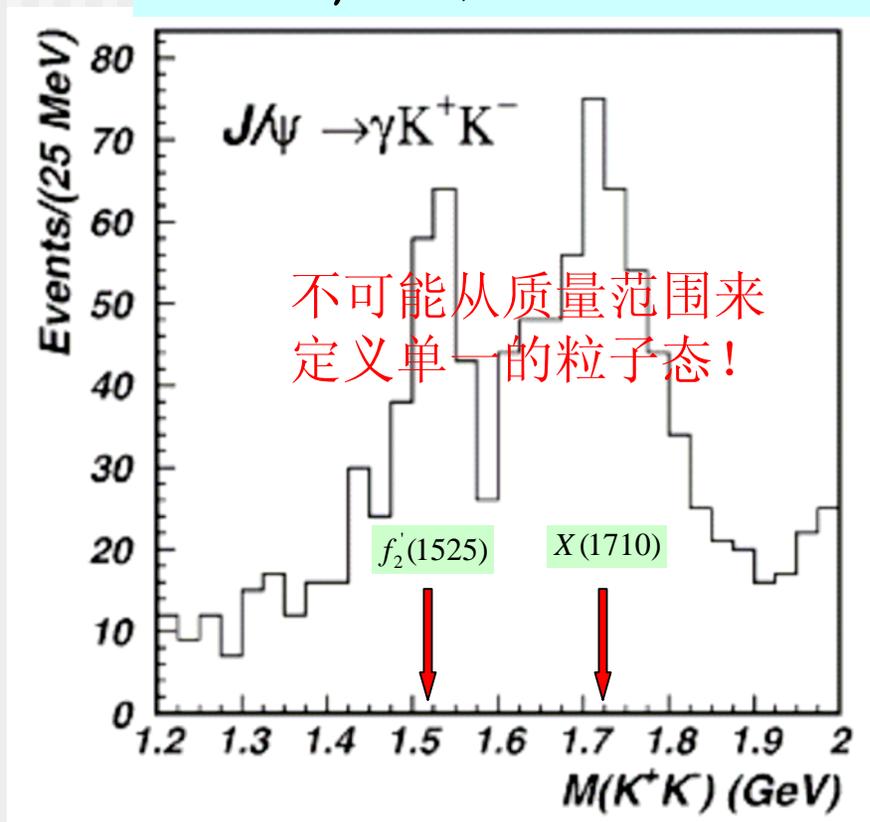
本讲要点

- 矩的几种定义
- 简单的矩方法
- 一般的矩方法
- 矩方法与最大似然法和最小二乘法的比较

自旋宇称在实验中的确定问题

实验上通常观测的是几种态的叠加，需要用几方面的实验观测来区分各种可能的自旋取值以及相应分布范围，例如

$$J/\Psi \rightarrow \gamma + X, \quad X \rightarrow K^+ + K^-$$



注意： X 可以包含有几个粒子 (如左图)。

已知： 与自旋宇称有关的概率密度函数包含三个实验观测角度和一个不变质量。

问题： 能否存在一种简单而又不失精确度的方法确定自旋？



矩方法。

概率分布中的矩定义

考虑一个服从概率密度函数 $P(x)$ 的连续随机变量 x 。定义围绕一固定值 x_0 的第 k 阶矩或简单矩为

$$\mu_k = \int (x - x_0)^k P(x) dx \xrightarrow{\text{代数矩}} \mu'_k = \int x^k P(x) dx$$

如果令 $x_0=0$ ，则一阶矩就是随机变量 x 的期望值定义(也称作一阶代数矩)

$$E[x] = \int xP(x)dx = \mu = \mu'_1$$

如果令 $x_0=E[x]$ ，随机变量 x 围绕期望值的二阶矩就是随机变量 x 的方差定义(也称作二阶中心矩)

$$V[x] = \int (x - E[x])^2 P(x) dx = \sigma^2 = \mu_2$$

代数矩与中心矩的关系

代数矩

$$\mu'_0 = 1$$

$$\mu'_1 = \mu$$

$$\mu'_2 = \sigma^2 + \mu^2$$

低阶矩之间的关系

$$\mu_2 = \mu'_2 - \mu^2$$

中心矩

$$\mu_0 = 1$$

$$\mu_1 = 0$$

$$\mu_2 = \sigma^2$$

一般情况下，它们的关系可以有如下表示

$$\mu_k = \sum_{l=0}^k \binom{k}{l} \mu'_{k-l} (-\mu'_1)^l$$

$$\mu'_k = \sum_{l=0}^k \binom{k}{l} \mu_{k-l} (\mu'_1)^l$$

□ 高阶矩对研究概率密度函数在 $|x-\mu|$ 大值区间的行为很有帮助。

□ 对称分布的所有奇数中心矩为零。

统计样本矩的定义

假设观测值 $x=(x_1, \dots, x_n)$ ，则 k 阶样本代数矩定义为

$$m'_k = \frac{1}{n} \sum_{i=1}^n x_i^k$$



是 μ'_k 的无偏估计量

相应的一阶样本代数矩与二阶样本中心矩就是通常定义的样本平均值与无偏的样本方差

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \hat{\mu}$$

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

而 k 阶样本中心矩定义为

$$m_k = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^k$$



若 μ 已知

$$\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

用简单矩方法测定待定参数

考虑采用计算随机变量 x 自身代数矩的简单方法，对含有 r 个参数 $\vec{\lambda} = (\lambda_1, \dots, \lambda_r)$ 的概率密度函数情况，可以通过使 r 个不同阶代数矩与对应的样本矩相等，建立方程组

$$\begin{aligned} \mu'_1(\vec{\lambda}) &= m'_1 \\ &\vdots \\ \mu'_r(\vec{\lambda}) &= m'_r \end{aligned}$$

r 个方程

 r 个未知数

解出待定参数 $\vec{\lambda}$

可以证明样本矩的协方差矩阵为

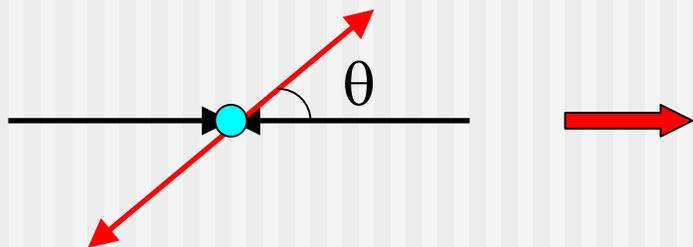
$$V_{kl} = \text{cov}(m'_k, m'_l) \cong \frac{1}{n-1} (\overline{x^{k+l}} - \overline{x^k} \cdot \overline{x^l})$$



参数的误差

角分布理论的简单验证

在实验 $e^+e^- \rightarrow \mu^+\mu^-$ 中，理论预言角分布为



$$\frac{dn}{d \cos \theta} = n(1 + \cos^2 \theta)$$

$n =$ 事例数

将角分布归一化变为 $\cos \theta$ 的概率密度函数，则其二阶代数矩期待值

$$E[\cos^2 \theta] = \frac{\int_{-1}^{+1} \cos^2 \theta \cdot n(1 + \cos^2 \theta) d \cos \theta}{\int_{-1}^{+1} n(1 + \cos^2 \theta) d \cos \theta} = 0.4$$

为了验证理论，我们计算 $\cos \theta$ 二阶代数样本矩平均值

$$\overline{\cos^2 \theta} = \frac{1}{n} \sum_{i=1}^n \cos^2 \theta_i$$

假设的统计检验可以通过简单比较二阶代数矩的期待值与样本矩平均值来完成。

简单验证中的误差估计

在前面例子中对于不含参数的简单情形 $\cos\theta$ 二阶代数矩平均值的误差估计可以按下列方法进行

已知真值 $E[\cos^2 \theta] = 0.4$

样本矩的方差为 $S^2 = \frac{1}{n} \sum_{i=1}^n (\cos^2 \theta_i - E[\cos^2 \theta])^2$

样本矩平均值的方差可以证明为 $V[\overline{\cos^2 \theta}] = S^2 / n$

假设实验观测1000次, $\overline{\cos^2 \theta} = 0.39$, 并计算出 $S^2 = 0.15$, 则实验结果报告为

0.39 ± 0.01 观测值在一个标准误差范围内与理论预期相符。

含参数情况举例

在上例中，假设已知理论中包含一未知参数 α ，例如

$$\frac{dn}{d \cos \theta} = n(1 + \alpha \cos^2 \theta)$$

和前例一样，计算出 $\cos \theta$ 二阶代数矩的理论期待值

$$E[\cos^2 \theta] = \frac{\int_{-1}^{+1} \cos^2 \theta \cdot n(1 + \alpha \cos^2 \theta) d \cos \theta}{\int_{-1}^{+1} n(1 + \alpha \cos^2 \theta) d \cos \theta} = \frac{5 + 3\alpha}{5(3 + \alpha)}$$

则参数 α 与二阶代数矩的关系为

$$\alpha = \frac{5(3E[\cos^2 \theta] - 1)}{3 - 5E[\cos^2 \theta]} \quad \longrightarrow \quad \hat{\alpha} = \frac{5(3\overline{\cos^2 \theta} - 1)}{3 - 5\overline{\cos^2 \theta}}$$

只要函数是可积的，采用矩方法原则上就可以测定参数。

简单矩方法应用的其它问题

非物理解问题：利用矩方法测定参数，可能会出现非物理结果。例如前例的二阶代数矩中，如果

$$\overline{\cos^2 \theta} \rightarrow 0.6 \quad \Rightarrow \quad \alpha \rightarrow \infty$$

在矩方法中，我们无法加上限制条件使得参数的测定值保持在物理允许的范围。

假设检验问题：利用矩方法测定参数，由于只比较积分值并解方程得到参数估计值，信息含量不足，因此无法判断所得到的参数是否合理。实际应用中需要辅之以其它方法来检验。

适用范围问题：矩方法虽然简单，但在处理多参数问题中，由于涉及更高阶的积分，使研究变得复杂。在这种情况下，可以考虑采用所谓的“一般的矩方法”。

一般的矩方法

简单矩定义

$$\mu'_k = \int x^k P(x) dx$$

一般的矩定义

$$\gamma_k = \int f_k(x) P(x) dx$$

把对随机变量求平均推广到对随机变量的函数求平均。

一般的矩方法(续一)

对于有 k 个参数情形, 可以选用一组自变量为 x 的 k 个线性独立函数来构造矩, 其函数矩的期待值与估计值为

$$\gamma_r(\vec{\lambda}) = \int f_r(x) P(x | \vec{\lambda}) dx, \quad r = 1, 2, \dots, k$$

$$\hat{\gamma}_r(\vec{\lambda}) = \overline{f_r(x)} = \frac{1}{n} \sum_{i=1}^n f_r(x_i), \quad r = 1, 2, \dots, k$$

估计值为函数的样本平均值, 相应的协方差矩阵元为

$$\begin{aligned} V_{rs}[\hat{\gamma}] &\cong \frac{1}{n(n-1)} \sum_{i=1}^n \left(f_r(x_i) - \overline{f_r(x)} \right) \left(f_s(x_i) - \overline{f_s(x)} \right) \\ &= \frac{1}{n(n-1)} \left\{ \sum_{i=1}^n f_r(x_i) f_s(x_i) - n \cdot \overline{f_r(x)} \cdot \overline{f_s(x)} \right\} \end{aligned}$$

一般的矩方法(续二)

例如, 假设一角分布为

$$P(\cos \theta, \phi) = \frac{3}{4\pi} \left(\frac{1}{2}(1 - \rho_{00}) + \frac{1}{2}(3\rho_{00} - 1)\cos^2 \theta - \rho_{1-1}\sin^2 \theta \cos 2\phi - \sqrt{2} \operatorname{Re} \rho_{10} \sin 2\theta \cos \phi \right)$$

这里 $-1 \leq \cos \theta \leq +1$, $0 \leq \phi \leq 2\pi$, ρ_{00} , ρ_{1-1} , $\operatorname{Re} \rho_{10}$ 为待定参数。

如果定义与角度有关的三个函数

$$f_1(\cos \theta, \phi) = \cos^2 \theta, \quad f_2(\cos \theta, \phi) = \sin^2 \theta \cos 2\phi, \quad f_3(\cos \theta, \phi) = \sin 2\theta \cos \phi$$

可以计算出相应函数的样本平均值

$$\bar{f}_1 = \frac{1}{5}(1 + 2\rho_{00}) = \frac{1}{n} \sum_{i=1}^n \cos^2 \theta_i$$

$$\bar{f}_2 = -\frac{4}{5}\rho_{1-1} = \frac{1}{n} \sum_{i=1}^n \sin^2 \theta_i \cos 2\phi_i$$

$$\bar{f}_3 = -\frac{4}{5}\sqrt{2} \operatorname{Re} \rho_{10} = \frac{1}{n} \sum_{i=1}^n \sin 2\theta_i \cos \phi_i$$

简记为



$$\hat{\rho} \equiv c + S\bar{f}$$

三个参数可以从三个不同的样本矩平均值中直接得到。

一般的矩方法(续三)

其中

$$\hat{\rho} = \{\rho_{00}, \rho_{1-1}, \text{Re } \rho_{10}\}$$

$$\bar{f} = \{\bar{f}_1, \bar{f}_2, \bar{f}_3\}$$

$$c = \{-1/2, 0, 0\}$$

$$S = \begin{pmatrix} 5/2 & 0 & 0 \\ 0 & -5/4 & 0 \\ 0 & 0 & -5/4\sqrt{2} \end{pmatrix}$$

由于矩的估计值就是函数的样本平均值，因此满足

$$V_{rs}(\hat{\gamma}) = V_{rs}(\bar{f})$$

可以证明参数估计值的方差矩阵可以确定为

$$V(\hat{\rho}) = SV(\bar{f})S^T = \begin{pmatrix} 8V_{11} & -4V_{12} & -2\sqrt{2}V_{13} \\ -4V_{12} & 2V_{22} & \sqrt{2}V_{23} \\ -2\sqrt{2}V_{13} & \sqrt{2}V_{23} & V_{33} \end{pmatrix}$$

一般的矩方法(续四)

如果概率密度函数可以表示成

$$P(x | \vec{\lambda}) = F + \sum_{r=1}^k \lambda_r g_r(x),$$

这里 F 为归一化常数, $g_r(x)$ 代表一组 k 个正交函数, 满足

$$\int_{\Omega} g_r(x) g_s(x) dx = \delta_{rs}, \quad \int_{\Omega} g_r(x) dx = 0, \quad s = 1, 2, \dots, k$$

那么从 $g_r(x)$ 的样本估计值可直接给出参数估计值与方差

$$\hat{\lambda}_r = \overline{g_r(x)} = \frac{1}{n} \sum_{i=1}^n g_r(x_i), \quad r = 1, 2, \dots, k$$

$$V_{rr}(\vec{\lambda}) = \frac{1}{n-1} \left(\frac{1}{n} \sum_{i=1}^n g_r^2(x_i) - \hat{\lambda}_r^2 \right) \cong \frac{1}{n-1} (1 - \hat{\lambda}_r^2), \quad r = 1, 2, \dots, k$$

一般的矩方法(续五)

在粒子物理与核物理研究中，所涉及的末态粒子角分布大都可以表示成球谐函数的形式。

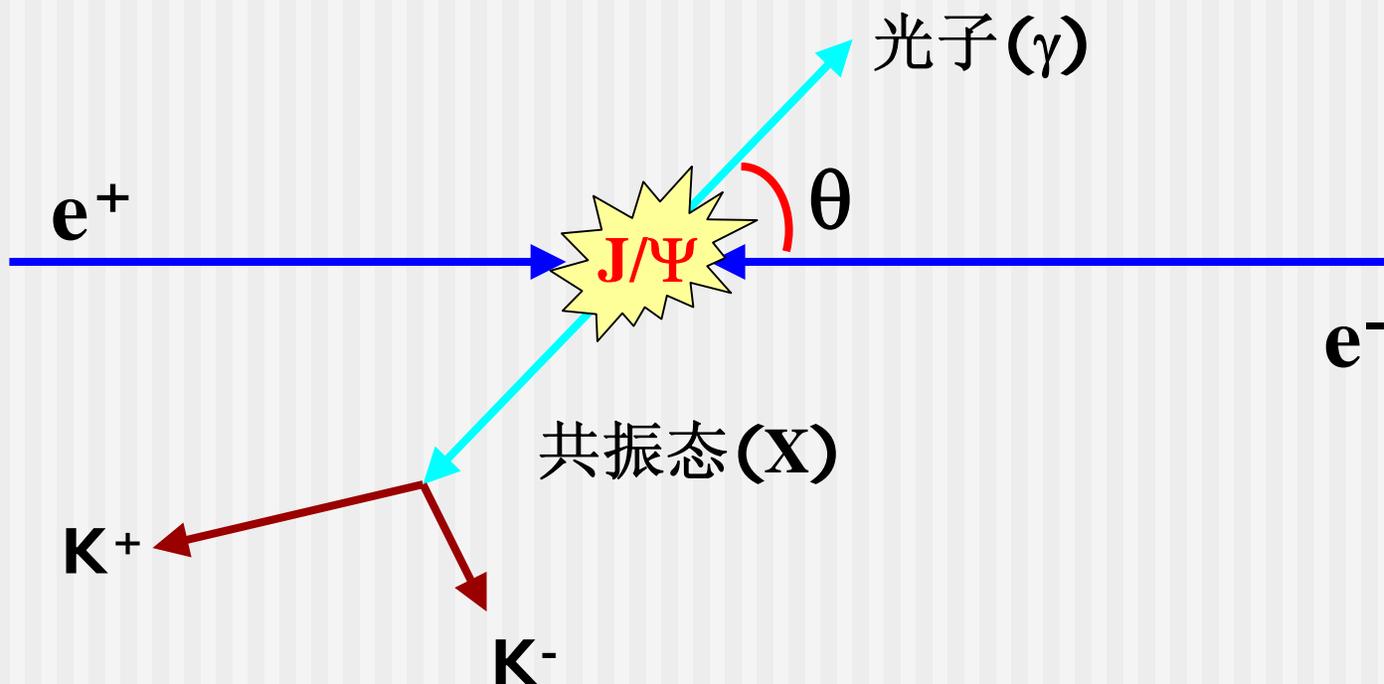
$$\int_{4\pi} Y_j^m(\Omega) Y_{j'}^{m'}(\Omega)^* d\Omega = \delta_{jj'} \delta_{mm'}$$
$$\int_{4\pi} Y_j^m(\Omega) d\Omega = 0$$

因此，常常采用矩方法来从实验上测定粒子的自旋。

为了减少因矩方法经过积分后，可能因信息不全而带来的影响，通常计算比待定参数的还要多的矩，比较对应的样本矩来构造最小二乘函数，得到反映实验数据与理论模型之间拟合优度的定量表述，给出参数的估计值。

一般的矩方法确定粒子自旋

在北京正负电子对撞机中产生的一个物理过程过程

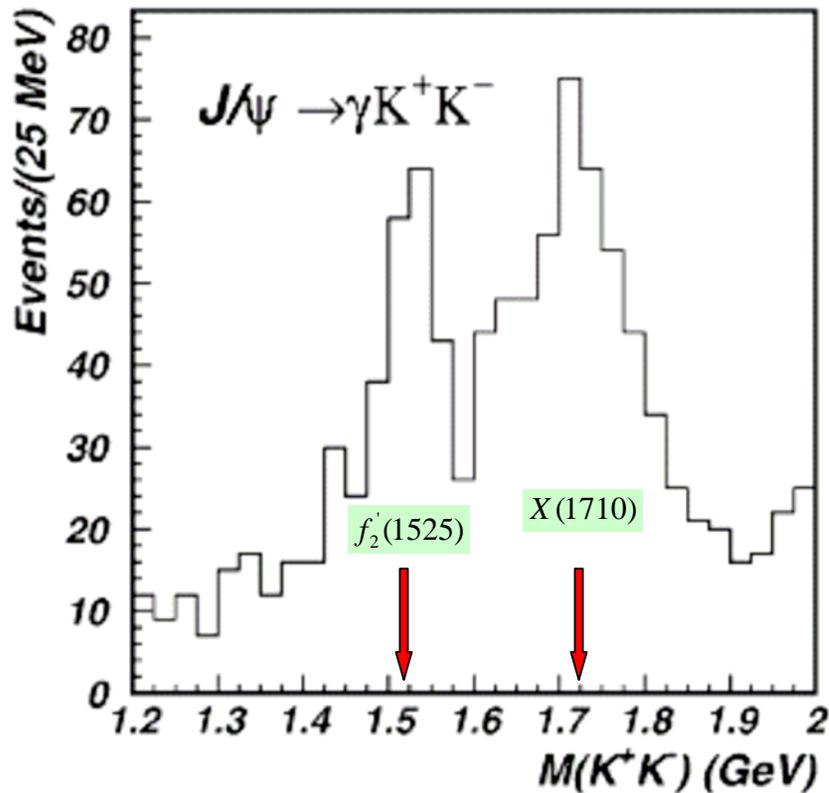


实验原理：根据角动量守恒，粒子X的自旋取值与其衰变的末态粒子角分布有关。通过计算不同自旋取值、系统轨道角动量和自旋在系统动量方向的投影所对应的函数矩，并与样本矩比较来确定粒子的自旋。

确定粒子的自旋(续一)

$$J/\Psi \rightarrow \gamma + X, \quad X \rightarrow K^+ + K^-$$

注意：**X**可以包含有几个粒子(如左图)。



实验观测量： γ ， K^+ 与 K^- 的能动量，因此

$$\cos \theta = \frac{p_z(\gamma)}{|p(\gamma)|},$$

$$E = E(K^+) + E(K^-),$$

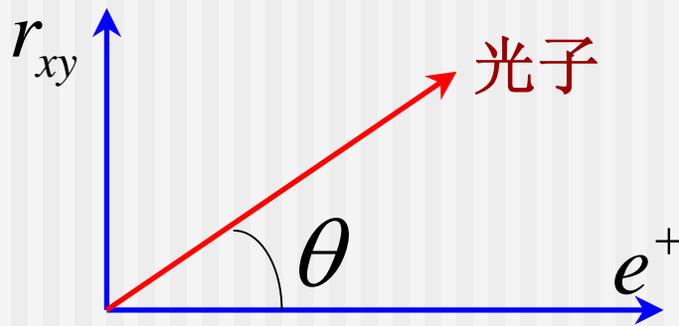
$$\vec{p} = \vec{p}(K^+) + \vec{p}(K^-)$$

$$m(K^+K^-) = \sqrt{E^2 - |\vec{p}|^2}$$

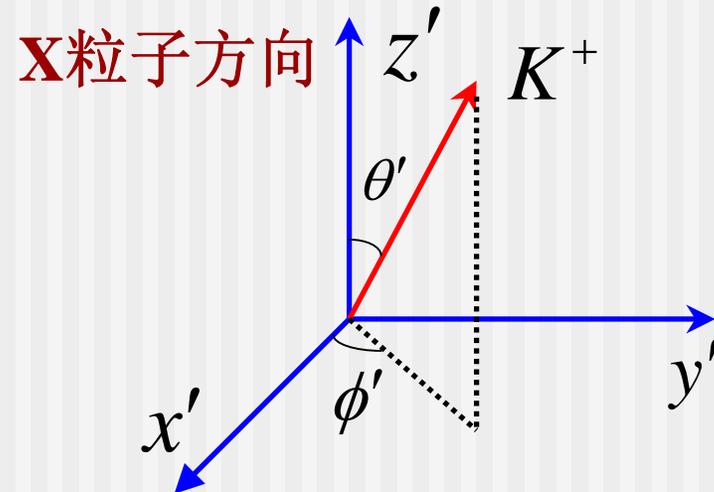
X， K^+ 的能动量可给出 K^+ 在**X**质心系的极角与方位角 $(\theta', \phi') = \Omega$

能否观察自旋取值随质量的变化关系？

确定粒子的自旋(续二)



$$\Omega = (\phi', \theta', 0)$$



在任一X不变质量区间与角度相关的概率密度函数与矩函数为

$$W(\theta, \Omega | \vec{\lambda})$$

⊗
构造函数矩

$$\text{Re}[D_{0,-m}^j(0, \theta, 0) D_{m,0}^{l*}(\Omega)]$$

$\vec{\lambda}$ 是表征动力学因子的参数

(j, l, m) :是与自旋、轨道角动量、轨道角动量第三分量有关的量。

$$T_{(j,l,m)} = \int W(\theta, \Omega | \vec{\lambda}) \text{Re}[D_{0,-m}^j(0, \theta, 0) D_{m,0}^{l*}(\Omega)] d \cos \theta d \Omega$$

确定粒子的自旋(续三)

各种可能情况下任意一个质量区间的矩与参数A与B(螺旋度振幅)有关

$$T_{(0,0,0)} \sim 8[(A_{10}^2 + A_{11}^2 + A_{12}^2) + B_{10}^2], \quad T_{(2,0,0)} \sim \frac{4}{5}[(A_{10}^2 - 2A_{11}^2 + A_{12}^2) + B_{10}^2],$$

$$T_{(0,2,0)} \sim 8\left[\frac{1}{7}(2A_{10}^2 + A_{11}^2 - 2A_{12}^2) + \frac{1}{\sqrt{5}}A_{10}B_{10}\alpha\right],$$

$$T_{(2,2,0)} \sim \frac{8}{5}\left[\frac{1}{7}(A_{10}^2 - A_{11}^2 - A_{12}^2) + \frac{1}{2\sqrt{5}}A_{10}B_{10}\alpha\right],$$

$$T_{(2,2,1)} \sim \frac{4}{5}\left[\frac{1}{7}(\sqrt{3}A_{10}A_{11} - 3\sqrt{2}A_{11}A_{12}) + \frac{\sqrt{3}}{2\sqrt{5}}A_{11}B_{10}\alpha\right],$$

$$T_{(2,2,2)} \sim \frac{4\sqrt{6}}{5}\left[-\frac{2}{7}A_{10}A_{12} + \frac{1}{2\sqrt{5}}A_{12}B_{10}\alpha\right],$$

$$T_{(0,4,0)} \sim \frac{8}{21}(6A_{10}^2 - 4A_{11}^2 + A_{12}^2), \quad T_{(2,4,0)} \sim \frac{4}{105}(6A_{10}^2 - 8A_{11}^2 + A_{12}^2),$$

$$T_{(2,4,1)} \sim \frac{4}{105}(3\sqrt{10}A_{10}A_{11} + \sqrt{15}A_{11}A_{12}), \quad T_{(2,4,2)} \sim \frac{4\sqrt{10}}{35}A_{10}A_{12}.$$

自旋=0⁺⁺:

B_{10}

自旋=2⁺⁺:

A_{10}, A_{11}, A_{12}

0⁺⁺与2⁺⁺干涉因子: α

原则上五个参数有五个矩方程即可。构造额外的矩只是为了尽可能包含数据的所有信息。

确定粒子的自旋(续四)

实验观测的样本矩为

$$N'^{(j,l,m)} = \sum_{\text{events}} \text{Re}[D_{0,-m}^j(0, \theta, 0) \cdot D_{m,0}^{l*}(\Omega)]$$

由于探测效率的缘故，必须对上述样本矩进行修正，使得

$$\text{样本矩: } N = C^{-1} N' \quad \xrightarrow{\text{对应于}} \quad \text{函数矩: } T$$

考虑到任一给定的不变质量和角度，经过探测器探测以后，其质量与角度值会发生改变。这种改变可通过效率矩阵来反映

$$C^{(j,l,m;j',l',m')} = \frac{1}{N_{MC}} \sum_{i=1}^{N_{Acc}} \text{Re}[D_{0,-m'}^{j'}(0, \theta, 0) D_{m',0}^{l'*}(\Omega)]$$

N_{MC} : 产生的事例数

N_{ACC} : 接收的事例数

$$\cdot (2 - \delta_{m0}) \text{Re}[D_{0,-m}^j(0, \theta, 0) D_{m,0}^{l*}(\Omega)]$$

确定粒子的自旋(续五)

该研究为了利用所有矩的信息，没有简单进行解方程求解参数值，而是，将所有十个含参数的矩与对应的样本矩进行比较，构造 χ^2 量，即

$$\chi^2 = (N - T)V^{-1}(N - T)$$

误差矩阵为

$$V = C^{-1}O(C^{-1})^+$$

与测量量相关的协方差矩阵可写为

$$O^{(j,l,m;j',l',k')} = \sum_{events} \text{Re}[D_{0,-m'}^{j'}(0, \theta, 0) D_{m',0}^{l'}{}^*(\Omega)] \\ \cdot \text{Re}[D_{0,-m}^j(0, \theta, 0) D_{m,0}^l{}^*(\Omega)]$$

确定粒子的自旋(续六)

对十个矩进行参数拟合

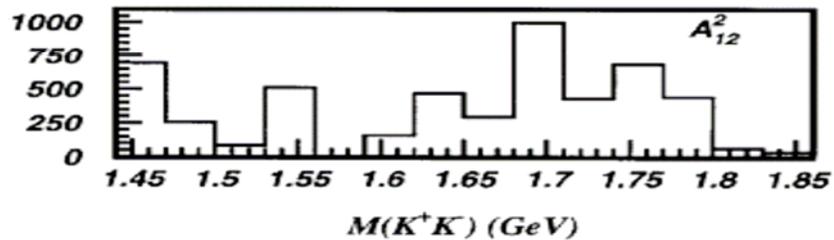
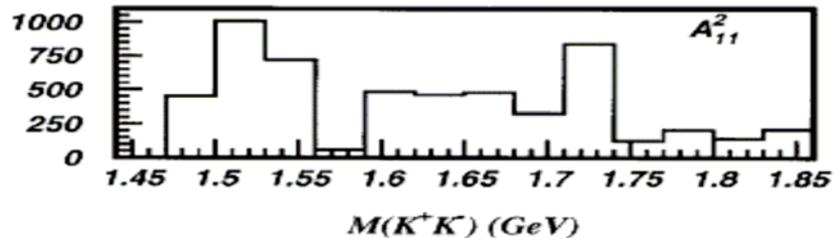
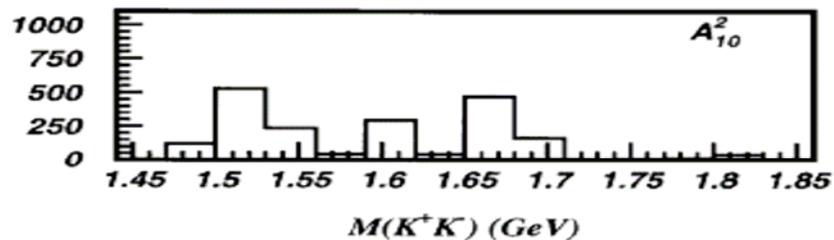
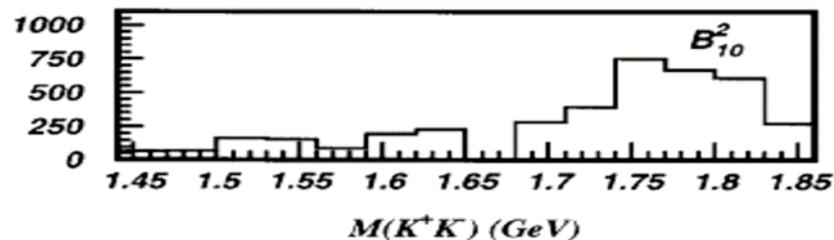
$$\chi^2 = \sum_{\mu, \nu=1}^{10} \left(\sum_{\sigma=1}^{10} C_{\mu\sigma}^{-1} N_{\sigma} - T_{\mu} \right) \cdot V^{-1} \left(\sum_{\sigma=1}^{10} C_{\nu\sigma}^{-1} N_{\sigma} - T_{\nu} \right)$$

按不变质量谱分成14个区间, 分别得到

$B_{10}, A_{10}, A_{11}, A_{12}, \alpha$

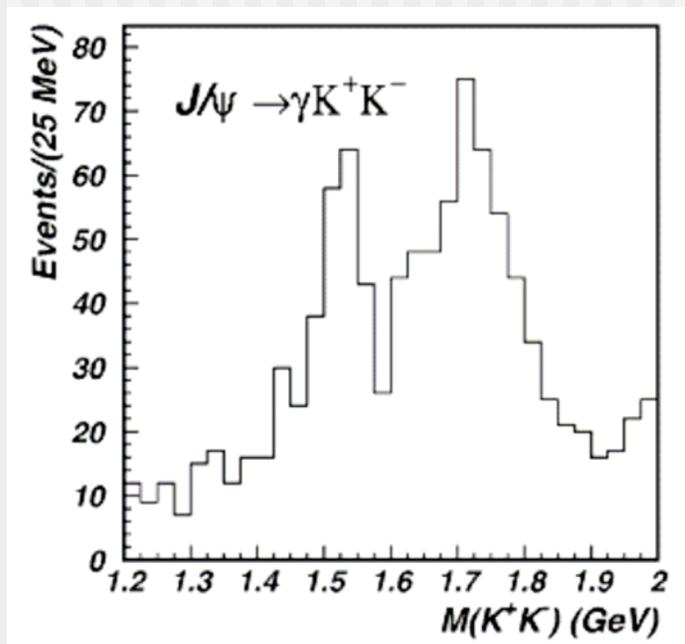


各螺旋度振幅随不变质量的变化关系(右图)。



确定粒子的自旋(续七)

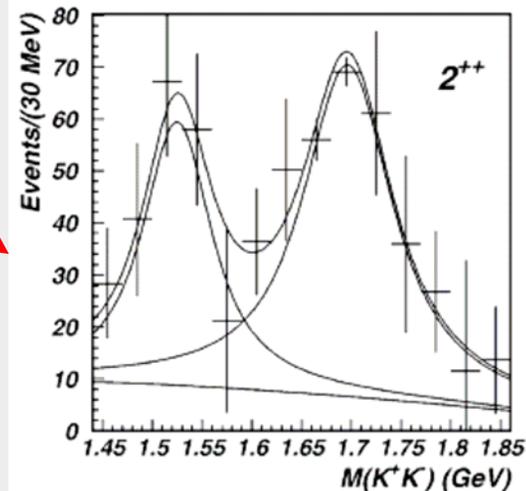
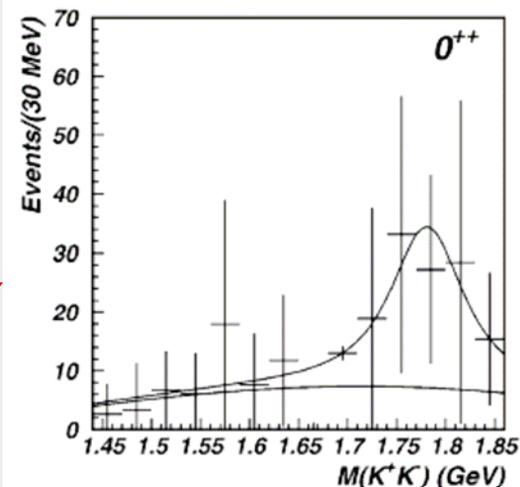
$$J/\Psi \rightarrow \gamma + X, \quad X \rightarrow K^+ + K^-$$



自旋为 0^{++}
的振幅影响

归一到与左图
相等的事例数

自旋为 2^{++}
的振幅影响



通过应用函数矩方法，可以较为方便地区分在不变质量分布上有重叠的不同自旋态。

关于矩方法的一点评论

原则上，利用最大似然法也能进行同样的研究，但是考虑到不同自旋态的联合概率密度函数必须已知，而通常情况下

$W(\theta, \Omega | \vec{\lambda})$ 是一个包含很多项的复杂函数

→
$$\ln L = \sum_{i=1}^{\text{事例数}} \ln \frac{W(\theta, \Omega | \vec{\lambda})}{\int W(\theta', \Omega' | \vec{\lambda}) p(\theta', \Omega' | \theta, \Omega) d \cos \theta' d \Omega'} + \text{与 } \vec{\lambda} \text{ 无关的项}$$

→ 拟合过程中每次参数变化都要重新计算，过程比较复杂

而在矩方法中，样本矩定义简单而且只需计算一次，参数的确定可简单通过比较样本矩与函数矩来完成，比较直观简便。

除非是对单个粒子态的不同自旋取值进行鉴别，可以采用最大似然方法，否则最好是利用函数矩方法，将问题简化来研究。

与最大似然法和最小二乘法比较

	矩方法	最大似然法	最小二乘法
数据输入	单个事例	单个事例	直方图
多维问题	最容易	归一化较复杂	较难
充分性	会有信息丢失	最具充分性	有时与区间大小有关
一致性	收敛于真值	收敛于真值	收敛于真值
有效性	不是最有效	通常最有效	基本上与似然法一样
无偏性	渐进无偏	渐进无偏	渐进无偏
拟合优度	较难评估	较难评估	很容易

充分性: 估计量应包含观测值对于未知参数的全部信息;

一致性: 样本容量增大时, 估计值收敛于真值;

有效性: 估计量的分布对其期望值具有最小方差;

无偏性: 无论样本容量多大, 估计值与真值无系统偏差。

小结

1. 矩的定义

简单矩，代数矩，中心矩

2. 简单矩方法

根据参数个数，建立不同阶代数矩方程并求解

3. 一般矩方法

把随机变量的函数看成新的随机变量，定义函数矩方程并求解

4. 与最大似然法和最小二乘法比较

对处理自旋宇称分析这类多维问题具有优势，但通常是结合最小二乘法来估计参数值。

习题

习题10.1 假设有一组观测量 $x=(x_1, \dots, x_n)$, 试证明其第 k 阶与第 l 阶代数矩协方差矩阵元可写为

$$V_{kl} = \text{cov}(m'_k, m'_l) \cong \frac{1}{n-1} (\overline{x^{k+l}} - \overline{x^k} \cdot \overline{x^l})$$

习题10.2 假设有一组观测量 $x=(\cos\theta_1, \dots, \cos\theta_n)$, 理论预期的角度分布满足

$$f(x; \alpha) \sim 1 + \alpha x^2, \quad -1 \leq x \leq +1$$

试写出简单矩方法中的二阶代数矩的方差, 并说明所得到的参数估计值正负误差可以是不对称的。