

## 第6章 多个测量值的合并估计

粒子物理实验中, 对于同一个物理参数 (如某个粒子衰变到特定末态的分支比), 通常会有不同的实验进行测量, 或者同一实验通过不同的过程进行测量. 将这些不同测量的结果进行统计意义上的合并, 通常会获得比单个测量结果更高的精度和可信度.

但是, 这些实验中有些只报道参数的结果及误差, 而获得结果所用的拟合方法 (极大似然法或最小二乘法) 及相关的概率分布的信息却难以获得; 参数的误差可能是对称的或不对称的; 参数的结果可能是 68.27% 中心置信区间, 也可能是单侧置信限 (如上限). 有些实验虽然给出了相关的概率分布的信息, 但可能用了不同的拟合方法; 信号事例的数据可能是直方图数据或散点图数据, 或者是子区间划分不同的直方图数据. 测量同一个物理参数的多个实验的结果有可能是相互独立的, 但也可能是 (部分) 关联的. 因此, 同一参数多个测量值的合并估计并不是一件简单的工作. 本章中我们阐述合并估计的几种典型的处理方法, 它们中的每一种只是在特定的条件下适用. 对于情况复杂的实际问题, 可能需要利用它们的适当组合才能处理.

鉴于信号的拟合通常利用极大似然法和最小二乘法, 所以 6.1 节和 6.2 节分别讨论极大似然合并估计和最小二乘合并估计. 6.3 节则专门讨论粒子衰变分支比的合并估计, 因为分支比的测量是粒子物理实验测量的重要课题.

### 6.1 多个测量值的极大似然合并估计

#### 6.1.1 各实验似然函数已知时的合并估计

##### 1. 一般方法

设两个相互独立的实验中各自的一组观测分别用  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  和  $\mathbf{y} = \{y_1, y_2, \dots, y_m\}$  表示, 观测量  $x$  和  $y$  是相互独立的随机变量, 它们的总体分布分别是  $f_1(x|\theta)$  和  $f_2(y|\theta)$ , 依赖同一个未知参数  $\theta$ , 即实验所要确定的物理量, 相应的似然函数分别是  $L(\mathbf{x}|\theta)$  和  $L(\mathbf{y}|\theta)$ . 这样, 对于这两个实验, 所有观测值  $\mathbf{x}$  和  $\mathbf{y}$  的联合似然函数可表示为

$$L(\mathbf{x}, \mathbf{y}|\theta) = \prod_{i=1}^n f_1(x_i|\theta) \prod_{j=1}^m f_2(y_j|\theta) = L(\mathbf{x}|\theta) \cdot L(\mathbf{y}|\theta). \quad (6.1.1)$$

利用该似然函数, 可通过似然方程求得  $\theta$  的极大似然估计  $\hat{\theta}$ :

$$\frac{\partial}{\partial \theta} \ln L(x, y|\theta) = \frac{\partial}{\partial \theta} \left[ \sum_{i=1}^n \ln f_x(x_i|\theta) + \sum_{j=1}^m \ln f_y(y_j|\theta) \right]_{\theta=\hat{\theta}} = 0,$$

该  $\hat{\theta}$  即两个实验对参数  $\theta$  的合并估计. 利用式 (6.1.1) 的似然函数, 可以根据第 5 章中阐述的方法进行区间估计确定给定置信水平的置信区间, 特别是确定  $\hat{\theta}$  的标准误差.

以上步骤可以直接推广到多个相互独立的实验测量的情况.

## 2. 正态型似然函数

当两个实验的总体分布是独立的正态分布时, 联合似然函数为两个正态函数的乘积, 两个相互独立的实验结果的合并特别简单.

两个独立的实验  $X$  和  $Y$  测量同一物理量  $\mu$ , 可视为两个总体期望值 (未知量) 相等, 两个实验中变量  $x$  和  $y$  的总体方差  $\sigma_x^2, \sigma_y^2$  不等但为已知. 令  $x_i$  ( $i = 1, 2, \dots, n$ ) 和  $y_j$  ( $j = 1, 2, \dots, m$ ) 是两个实验的测量值,  $n, m$  是两个实验中测量值个数 (子样容量), 故联合似然函数式 (6.1.1) 可写成

$$L(x, y|\mu) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_x} \exp \left[ -\frac{1}{2} \left( \frac{x_i - \mu}{\sigma_x} \right)^2 \right] \times \prod_{j=1}^m \frac{1}{\sqrt{2\pi}\sigma_y} \exp \left[ -\frac{1}{2} \left( \frac{y_j - \mu}{\sigma_y} \right)^2 \right], \quad (6.1.2)$$

似然方程为

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{(x_i - \mu)}{\sigma_x^2} + \sum_{j=1}^m \frac{(y_j - \mu)}{\sigma_y^2}.$$

由此求得两个实验对  $\mu$  的合并极大似然估计

$$\hat{\mu} = \frac{\frac{n}{\sigma_x^2} \bar{x} + \frac{m}{\sigma_y^2} \bar{y}}{\frac{n}{\sigma_x^2} + \frac{m}{\sigma_y^2}}, \quad (6.1.3)$$

其中

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{m} \sum_{j=1}^m y_j$$

是两个实验各自对期望值  $\mu$  的极大似然估计.  $\hat{\mu}$  的方差是

$$V(\hat{\mu}) = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_x^2} + \sum_{j=1}^m \frac{1}{\sigma_y^2}} = \frac{1}{\frac{n}{\sigma_x^2} + \frac{m}{\sigma_y^2}}. \quad (6.1.4)$$

用  $\hat{\mu}_{\bar{x}}, \hat{\sigma}_{\bar{x}}$  表示整个实验  $X$  ( $n$  次测量) 对于物理量  $\mu$  的测量值及其标准误差 (对于正态分布, 标准差等于标准误差), 则有

$$\hat{\mu}_{\bar{x}} = \bar{x}, \quad \hat{\sigma}_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}. \quad (6.1.5)$$

类似地, 对于实验  $Y$ , 有

$$\hat{\mu}_{\bar{y}} = \bar{y}, \quad \hat{\sigma}_{\bar{y}} = \frac{\sigma_y}{\sqrt{m}}. \quad (6.1.6)$$

于是式 (6.1.3) 和式 (6.1.4) 可改写为

$$\hat{\mu} = \frac{\frac{1}{\hat{\sigma}_{\bar{x}}^2} \hat{\mu}_{\bar{x}} + \frac{1}{\hat{\sigma}_{\bar{y}}^2} \hat{\mu}_{\bar{y}}}{\frac{1}{\hat{\sigma}_{\bar{x}}^2} + \frac{1}{\hat{\sigma}_{\bar{y}}^2}}, \quad V(\hat{\mu}) = \frac{1}{\frac{1}{\hat{\sigma}_{\bar{x}}^2} + \frac{1}{\hat{\sigma}_{\bar{y}}^2}}. \quad (6.1.7)$$

更为一般的情况是, 两个独立的实验测量总体分布都是 (或近似地) 正态分布, 测量同一个物理量 (总体期望值), 但测量误差不同而且未知. 虽然可根据式 (6.1.2) 的似然函数对  $\mu, \sigma_x, \sigma_y$  求偏导数写出似然方程, 但不能得到  $\mu, \sigma_x, \sigma_y$  的解析解. 我们可采用如下方法求出合并的极大似然估计  $\hat{\mu}$  及其方差  $V(\hat{\mu})$ . 首先求出两个实验中各自随机变量  $x$  和  $y$  的期望值和方差的同时估计

$$\begin{aligned} \hat{\mu}_x &= \bar{x}, \quad \hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \\ \hat{\mu}_y &= \bar{y}, \quad \hat{\sigma}_y^2 = \frac{1}{m} \sum_{j=1}^m (y_j - \bar{y})^2. \end{aligned} \quad (6.1.8)$$

这样, 合并的极大似然估计  $\hat{\mu}$  及其方差  $V(\hat{\mu})$  与式 (6.1.3) 和式 (6.1.4) 有类似的形式:

$$\hat{\mu} = \frac{\frac{n}{\hat{\sigma}_x^2} \bar{x} + \frac{m}{\hat{\sigma}_y^2} \bar{y}}{\frac{n}{\hat{\sigma}_x^2} + \frac{m}{\hat{\sigma}_y^2}}, \quad V(\hat{\mu}) = \frac{1}{\frac{n}{\hat{\sigma}_x^2} + \frac{m}{\hat{\sigma}_y^2}}.$$

同样可利用式 (6.1.5) 和式 (6.1.6) 定义两个实验对于  $\mu$  的测量值及其误差 ( $\sigma_x, \sigma_y$  分别用  $\hat{\sigma}_x, \hat{\sigma}_y$  代替), 则两个实验的合并估计也可用式 (6.1.7) 表示.

对于多个实验测定同一个物理量, 可按上述原则作类似的推导, 得出合并估计值及其误差  $\hat{\mu}, \sigma_{\hat{\mu}}$ :

$$\hat{\mu} = \frac{\sum_{i=1}^n \frac{\mu_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}, \quad \sigma_{\hat{\mu}}^2 \equiv V(\hat{\mu}) = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}, \quad (6.1.9)$$

式中,  $\mu_i, \sigma_i$  是第  $i$  个实验对于物理量  $\mu$  的测量值及其测量误差.

### 6.1.2 各实验似然函数未知时的合并估计

在许多情形下, 实验测量中的概率密度往往并不确切地知道. 对同一个待估计参数  $\theta$ , 不同的实验只是报道其测量结果  $\mu_{\sigma^-}^{+\sigma^+}$ . 应当认为对于该实验,  $\mu$  是  $\theta$  的最佳估计, 而  $(\mu - \sigma^-, \mu + \sigma^+)$  是可信度为 68.3% 的似然区间. 因此, 怎样从几个实验的测量值  $\mu_i, \sigma_i^+, \sigma_i^-$  ( $i = 1, 2, \dots, n$ ) 求得参数  $\theta$  的正确的合并估计  $\mu, \sigma^+, \sigma^-$ , 是一个需要解决的问题.

R. Barlow<sup>[113]</sup> 建议利用实验测量值  $\mu_i, \sigma_i^+, \sigma_i^-$  构造参数化的似然函数来逼近实验中的实际似然函数. 尝试了多种形式的参数化似然函数之后, 发现宽度可变的正态函数是比较好的选择, 即真值为  $\mu$ , 而测量值为  $\mu_i, \sigma_i^+, \sigma_i^-$  的似然函数可以用下述形式的正态函数来逼近:

$$\ln L(\mu_i|\mu) = -\frac{(\mu - \mu_i)^2}{2V_i(\mu)}. \quad (6.1.10)$$

对于线性离差的正态函数方案

$$V_i(\mu) = [\sigma_i(\mu)]^2, \quad \sigma_i(\mu) = \sigma_i + \sigma'_i(\mu - \mu_i), \quad (6.1.11)$$

$$\sigma_i = \frac{2\sigma_i^+\sigma_i^-}{\sigma_i^+ + \sigma_i^-}, \quad \sigma'_i = \frac{\sigma_i^+ - \sigma_i^-}{\sigma_i^+ + \sigma_i^-}, \quad (6.1.12)$$

即标准离差  $\sigma_i(\mu)$  在真值  $\mu$  附近是线性变化的. 当  $\sigma_i^+ = \sigma_i^-$  时正、负误差对称,  $\sigma'_i = 0$ ,  $\sigma_i(\mu) = \sigma_i^+ = \sigma_i^-$ , 回复到通常的正态似然函数的情形.

对于线性方差正态函数方案

$$V_i(\mu) = V_i + V'_i(\mu - \mu_i), \quad (6.1.13)$$

$$V_i = \sigma_i^+\sigma_i^-, \quad V'_i = \sigma_i^+ - \sigma_i^-, \quad (6.1.14)$$

即方差  $V_i(\mu)$  在真值  $\mu$  附近是线性变化的. 同样在正、负误差对称的情况下,  $\sigma_i^+ = \sigma_i^-$ , 则  $V_i = (\sigma_i^+)^2 = (\sigma_i^-)^2$ ,  $V'_i = 0$ , 回复到正态似然函数的情形.

#### 1. 同一物理量多次测量的合并估计

对同一物理量  $\mu$  的  $n$  个不同测量值  $\mu_i, \sigma_i^+, \sigma_i^-$  ( $i = 1, 2, \dots, n$ ), 其联合似然函数为

$$\ln L(\mu) = -\frac{1}{2} \sum_i \frac{(\mu - \mu_i)^2}{V_i(\mu)}. \quad (6.1.15)$$

物理量  $\mu$  的最佳估计  $\hat{\mu}$  由该似然函数的极大值位置决定. 对于线性离差正态似然函数方案,  $\hat{\mu}$  的解为

$$\hat{\mu} = \frac{\sum_i \omega_i \mu_i}{\sum_i \omega_i}, \quad (6.1.16)$$

$$\omega_i = \frac{\sigma_i}{[\sigma_i + \sigma'_i(\hat{\mu} - \mu_i)]^3}. \quad (6.1.17)$$

我们注意到  $\hat{\mu}$  的表达式与  $n$  次独立测量时正态总体期望值  $\hat{\mu}$  的极大似然估计形式相同, 只不过权因子稍有不同而已 (见式 (6.1.9)). 对于线性方差正态似然函数方案,  $\hat{\mu}$  的解为

$$\hat{\mu} = \sum_i \omega_i \left[ \mu_i - \frac{V'_i}{2V_i} (\hat{\mu} - \mu_i)^2 \right] / \sum_i \omega_i, \quad (6.1.18)$$

$$\omega_i = V_i / [V_i + V'_i(\hat{\mu} - \mu_i)]^2. \quad (6.1.19)$$

作为一个例子, 图 6.1 给出了我们构造的参数化似然函数与泊松分布似然函数 (期望值  $\mu=5$ ) 的对比, 图 6.2 给出了与对数正态分布似然函数 ( $x = \ln y$ ,  $y$  是  $\mu = 8$ ,  $\sigma = 3$  的正态变量) 的对比. 由图可见, 离差可变或者方差可变的正态似然函数所确定的可信度 68.3% 的正、负误差, 与这两种原分布似然函数确定的正、负误差, 数值是极为接近的. 由于对数正态分布, 特别是泊松分布对于描述子样容量较小的实验的似然函数往往是相当好的近似, 这就说明了参数化似然函数确定的正、负误差一般是相当精确的. 但是由图也可以看到, 对于高的可信度, 如 95.4% (相应于  $\Delta \ln L = -2$ ), 参数化似然函数方案确定的正、负误差与实际值就有比较明显的差别, 这也表明了这种方法适用范围的局限.

式 (6.1.16)、式 (6.1.17) 和式 (6.1.18)、式 (6.1.19) 各是一组非线性方程组,  $\hat{\mu}$  需要用迭代法求得数值解.  $\hat{\mu}$  的初值  $\hat{\mu}(0)$  可取为  $\frac{1}{n} \sum_i \mu_i$ , 代入  $\omega_i$  表达式右端中的  $\hat{\mu}$  计算  $\omega_i^{(0)}$ , 再代入  $\hat{\mu}$  的表达式右边计算  $\hat{\mu}(1)$ . 较之  $\hat{\mu}(0)$ ,  $\hat{\mu}(1)$  应当是  $\hat{\mu}$  的更好的近似. 经过  $k$  次迭代后, 如满足

$$|\hat{\mu}(k+1) - \hat{\mu}(k)| < 10^{-6} L,$$

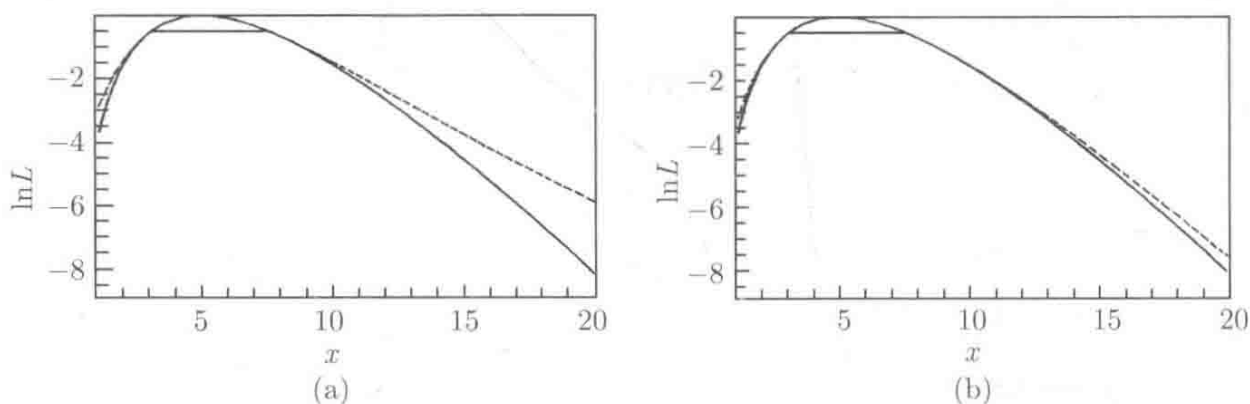


图 6.1 泊松分布似然函数 (期望值  $\mu = 5$ , 用实线表示) 与参数化方法构造的似然函数 (虚线) 的对比

$\Delta \ln L = -1/2$  的横线确定可信度 68.3% 的似然区间. (a) 是线性离差似然函数方案, (b) 是线性方差似然函数方案

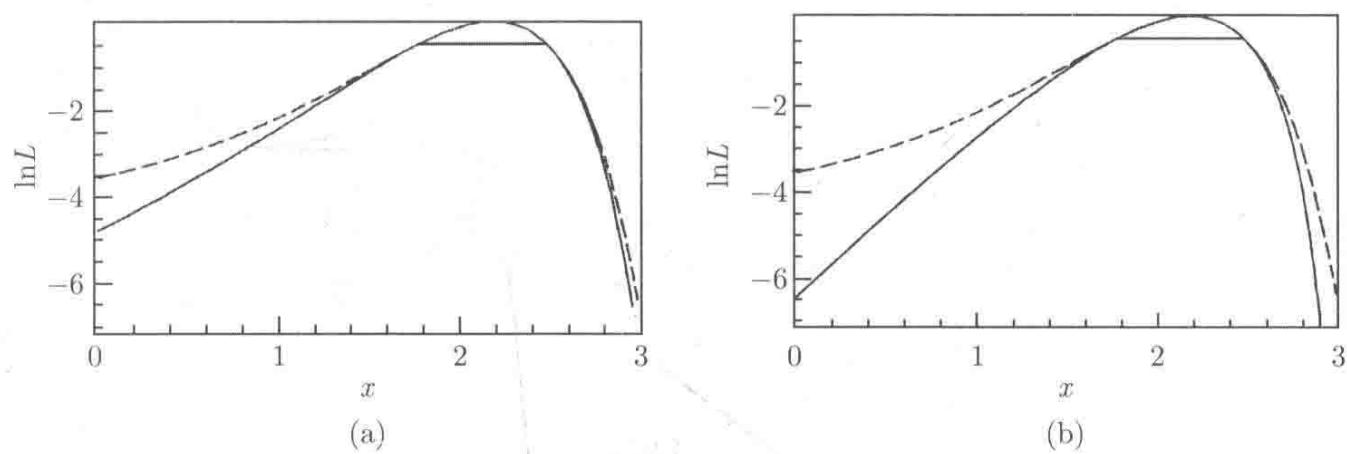


图 6.2 对数正态分布似然函数 ( $x = \ln y$ ,  $y$  是  $\mu = 8$ ,  $\sigma = 3$  的正态变量, 用实线表示) 与参数化方法构造的似然函数 (虚线) 的对比  
(a) 是线性离差正态似然函数方案, (b) 是线性方差似然函数方案

即可认为达到收敛,  $\hat{\mu}(k+1)$  可取为  $\hat{\mu}$  的解. 这里  $L$  定义为区间  $(\mu_l - 3\sigma_l^-, \mu_u + 3\sigma_u^+)$  的长度,  $\mu_l$  是  $n$  个测量值  $\mu_i$  中的最小值, 而  $\mu_u$  是最大值.

求得了  $\mu$  的估计值  $\hat{\mu}$  后, 其正、负误差  $\sigma^+$ 、 $\sigma^-$  仍需数值地求解. 受正态总体期望值误差式 (6.1.9) 的启发, 正、负误差的初值可由下式决定:

$$\sigma^+(0) = \left[ \sum_i \frac{1}{(\sigma_i^+)^2} \right]^{-\frac{1}{2}}, \quad \sigma^-(0) = \left[ \sum_i \frac{1}{(\sigma_i^-)^2} \right]^{-\frac{1}{2}}. \tag{6.1.20}$$

计算  $\Delta \ln L(\hat{\mu}(0))^{+/-} = \ln L(\hat{\mu}(0))^{+/-} - \ln L(\hat{\mu})$  (其中  $\hat{\mu}(0)^+ = \hat{\mu} + \sigma^+(0)$ ,  $\hat{\mu}(0)^- = \hat{\mu} - \sigma^-(0)$ ), 看它们与  $-1/2$  相差多大, 再调节  $\hat{\mu}(1)^{+/-}$  的值, 如此迭代, 使  $\Delta \ln L(\hat{\mu}(k))^{+/-}$  与  $-1/2$  的差别小于一个给定小量 (如  $0.5 \times 10^{-7}$ ), 即可认为结果收敛,  $\hat{\mu}(k)^{+/-}$  对应的  $\sigma^{+/-}(k)$  即是  $\sigma^{+/-}$  的正确估计.

将  $\mu$  的估计值  $\hat{\mu}$  代入式 (6.1.15) 求得  $\ln L(\hat{\mu})$ , 它代表了不同测量结果用同一个理论模型描述时的差异程度, 即等同于皮尔逊拟合优度检验中的  $\chi^2$  变量观测值, 可以用自由度  $n - 1$  的  $\chi^2$  分布作拟合优度检验.

**示例 同一物理量多次测量的合并估计.**

设对同一物理量有 3 个测量结果:  $1.9^{+0.7}_{-0.5}$ ,  $2.4^{+0.8}_{-0.6}$ ,  $3.1^{+0.5}_{-0.4}$ , 求该物理量期望值和不对称误差的合并估计.

利用可变宽度参数化正态函数逼近的似然函数 (线性离差方案), 3 个测量结果可用图 6.3 中的 3 条实线来表示. 图中上下两条水平线的  $\ln L$  值差别为  $-0.5$ . 可以看到每条似然函数曲线与下直线的两个交点正确地反映了每个测量结果的正、负误差. 虚线是 3 条实线的合并结果, 其极大值对应于合并期望值, 与下水平线的两个交点表示合并结果的正、负误差, 数字结果是  $2.76^{+0.29}_{-0.27}$ .

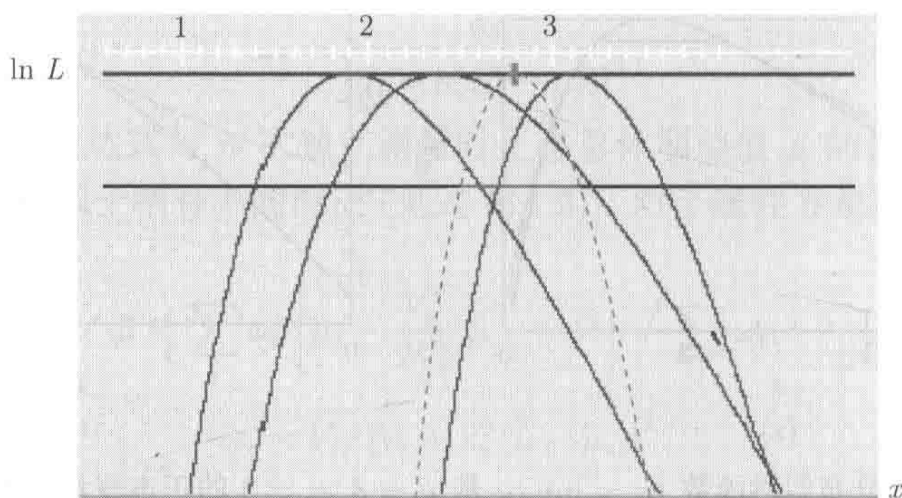


图 6.3 同一量的 3 个测量值 (具有不对称误差) 的合并估计

作为比较, 如果用正态近似多次测量的加权平均公式

$$\hat{\mu} = \frac{\sum_{i=1}^n \frac{\mu_i}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} = \frac{\sum_{i=1}^n w_i \mu_i}{\sum_{i=1}^n w_i}, \quad V(\hat{\mu}) = \frac{1}{\sum_{i=1}^n \frac{1}{\sigma_i^2}} = \frac{1}{\sum_{i=1}^n w_i}$$

(计算  $\hat{\mu}$  时,  $\sigma_i = (\sigma_i^+ + \sigma_i^-) / 2$ ;  $V^{+/-}(\hat{\mu})$  则用  $\sigma_i^+$  和  $\sigma_i^-$  分别计算) 计算, 所得结果是  $2.61_{-0.28}^{+0.34}$ , 与正确结果  $2.76_{-0.27}^{+0.29}$  差别较大.

## 2. 多个物理量不对称误差的合并估计

现在再来讨论另外一种经常遇到的问题——两个 (或多个) 物理量不对称误差的合并估计问题. 假设待测的物理量需要从实验测定的某种分布来确定, 该分布中的事例 (子样) 除了待测物理量的信号之外, 还包含其他过程的污染 (称为本底) 事例, 而且本底的来源不止一种. 我们只知道若干种本底各自的测量值及其不对称误差, 怎样求得所有本底的总误差. 因为这一信息对于确定信号 (即待测物理量) 的大小及误差是至关重要的. 例如, 每项本底事例数  $N_i$  的测量值及其不对称误差为  $\mu_i, \sigma_i^+, \sigma_i^-$  ( $i = 1, 2, \dots, n$ ), 要求总的本底事例数  $N = \sum_{i=1}^n N_i$  的期望值  $\mu$  及其误差.

显然期望值用  $\mu = \sum_{i=1}^n \mu_i$  来估计, 误差则通过以下方法求得.

对于  $n$  项本底来源, 用  $x_i$  ( $i = 1, 2, \dots, n$ ) 表示本底测量值与其期望值间的差别, 即离差. 与式 (6.1.15) 类似, 我们可构造这  $n$  次测量的联合似然函数

$$\ln L(x) = -\frac{1}{2} \sum_{i=1}^n \left( \frac{x_i}{\sigma_i + \sigma_i' x_i} \right)^2 \quad (6.1.21)$$

或

$$\ln L(x) = -\frac{1}{2} \sum_{i=1}^n \frac{x_i^2}{V_i + V'_i x_i}, \tag{6.1.22}$$

其中,  $\sigma_i, \sigma'_i, V_i$  和  $V'_i$  由式 (6.1.12) 和式 (6.1.14) 确定,  $n$  项测量离差  $x_i$  的总和表示为

$$u = \sum_{i=1}^n x_i. \tag{6.1.23}$$

为了找到  $u$  的似然函数  $L(u)$ , 对似然函数  $\ln L(x)$  在约束方程 (6.1.23) 条件下求极大值, 利用不定乘子法得到的解是

$$x_i = \frac{u\omega_i}{\sum_{i=1}^n \omega_i}, \tag{6.1.24}$$

$$\omega_i = \frac{(\sigma_i + \sigma'_i x_i)^3}{2\sigma_i} \quad \text{或} \quad \omega_i = \frac{(V_i + V'_i x_i)^2}{2V_i + V'_i x_i}. \tag{6.1.25}$$

这是一组非线性方程,  $L(u)$  可用以下方法逐点地计算出来. 以  $u_0 = 0$  作为起点, 这时  $x_i = 0$  ( $i = 1, 2, \cdots, n$ ). 由式 (6.1.21) 和式 (6.1.22) 知  $\ln L(u_0 = 0) = 0$ . 将  $u$  的值增加一个小量变为  $u_1$ , 利用式 (6.1.24) 和式 (6.1.25) 迭代逐次逼近, 直到计算出  $u_1$  相应的  $x_i$  值, 于是代入式 (6.1.21) 和式 (6.1.22) 计算出  $u_1$  处的  $\ln L(u_1)$  值. 如此重复便可计算出整条  $L(u)$ - $u$  曲线, 再利用  $\Delta \ln(u) = -\frac{1}{2}$  的关系就可求得  $n$  项本底的正负总误差.

**示例** 多个物理量不对称误差的合并估计.

设对两个物理量的测量值为  $4^{+2.346}_{-1.682}, 5^{+2.581}_{-1.916}$ , 求两者之和的合并估计.

显然两个物理量之和的中心值为 9, 用上面的方法 (线性离差方案) 求得正、负误差, 最后结果是  $9^{+3.333}_{-2.668}$ .

事实上,  $4^{+2.346}_{-1.682}$  和  $5^{+2.581}_{-1.916}$  分别是对于期望值为 4 和 5 的泊松变量利用参数化似然函数方法确定的正负误差, 两者之和应服从期望值为 9 的泊松分布, 其精确结果是  $9^{+3.342}_{-2.676}$ . 与上面的合并结果比较, 差别小于 0.3%.

表 6.1 给出不同的测量值之和用参数化似然函数方法确定的正、负误差, 表中的输入值只列出了中心值, 它们对应的正、负误差是利用参数化似然函数方法确定的正负误差, 如输入值  $4+5$  对应于  $4^{+2.346}_{-1.682} \oplus 5^{+2.581}_{-1.916}$ . 用参数化似然函数方法确定的正、负误差与精确值  $\sigma_- = 2.676, \sigma_+ = 3.342$  的差别都很小. 在两次测量结果合并的情形下, 相对误差  $< 1\%$ .

本节讨论的利用参数化似然函数方法进行多次测量的不对称误差合并计算已编为程序包, 读者可从网页 <http://www.slac.stanford.edu/~barlow/statistics.html><sup>[114]</sup> 下载.

表 6.1 线性离差和线性方差正态近似确定的正负误差  $\sigma_+$ 、 $\sigma_-$

输入值	线性离差方案		线性方差方案	
	$\sigma_-$	$\sigma_+$	$\sigma_-$	$\sigma_+$
4+5	2.653	3.310	2.668	3.333
3+6	2.653	3.310	2.668	3.333
2+7	2.653	3.310	2.668	3.333
1+8	2.654	3.313	2.668	3.333
3+3+3	2.630	3.278	2.659	3.323
1+1+1+1+1+1+1+1	2.500	3.098	2.610	3.270

注: 期望值为 9 的泊松分布, 置信概率 68.27% 的区间是  $9^{+3.342}_{-2.676}$ .

6.2 多个测量值的最小二乘合并估计

若实验在  $N$  个观测点  $x_1, x_2, \dots, x_N$  得到测量值  $y_1, y_2, \dots, y_N$ , 相应的测量值真值  $\eta_1, \eta_2, \dots, \eta_N$  为未知. 假定理论模型

$$\eta_i = f(\theta_1, \theta_2, \dots, \theta_L; x_i), \quad L \leq N$$

描述了真值  $\eta_i$  与  $x_i$  的函数关系, 该函数与待估计的未知参数  $\theta = \{\theta_1, \theta_2, \dots, \theta_L\}$  相关, 则最小二乘原理告诉我们,  $\theta$  的最优估计值是使量

$$Q^2 = \sum_{i=1}^N \sum_{j=1}^N (y_i - \eta_i)(V^{-1}(\mathbf{y}))_{ij}(y_j - \eta_j) \tag{6.2.1}$$

达到极小的参数值  $\hat{\theta}$ . 其中,  $V(\mathbf{y})$  是  $N$  个观测值  $y_1, y_2, \dots, y_N$  的协方差矩阵.

但是这种解法建立在协方差矩阵  $V(\mathbf{y})$  已知的基础之上. 然而实际情况中不乏只知道  $\mathbf{y}$  的数值并知道不同的测量值  $y_i$  之间存在相互关联, 但其协方差矩阵  $V(\mathbf{y})$  不确切知道或无法定量确定的情况. 这样, 此前叙述的方法不能用来求得未知参数  $\theta$  的估计及其误差.

本节将讨论测量值协方差矩阵  $V(\mathbf{y})$  未知情形下估计未知参数及其误差的方法. 我们来讨论一种比较简单的情形. 假定  $N$  个实验对同一个物理量进行测量得到了  $N$  个测量值  $y_i \pm \sigma_i$  ( $i = 1, 2, \dots, N$ ). 按照惯常的理解,  $y_i$  是第  $i$  个实验对物理量  $\eta$  的最优估计,  $y_i \pm \sigma_i$  确定了  $\eta$  的 68.3% 置信度的区间. 我们的问题是怎样从这  $N$  个测量结果求得物理量及其误差的合并估计.

在这一问题中, 观测值真值  $\eta$  是一个数值而不是矢量, 而且它本身就是待估计的未知参数. 由上述的一般最小二乘原理可知, 这种情形下  $\eta$  的最优估计是使量

$$Q^2 = \sum_{i=1}^N \sum_{j=1}^N (y_i - \eta)(V^{-1})_{ij}(y_j - \eta) \quad (6.2.2)$$

达到极小来求得, 其解为

$$\hat{\eta} = \left[ \sum_{i,j=1}^N (V^{-1})_{ij} \right]^{-1} \left[ \sum_{i,j=1}^N (V^{-1})_{ij} y_j \right], \quad (6.2.3)$$

而  $\hat{\eta}$  的方差则为

$$\sigma_{\hat{\eta}}^2 = \left[ \sum_{i,j=1}^N (V^{-1})_{ij} \right]^{-1}, \quad (6.2.4)$$

(见式 (6.2.1), 注意此时仅有一个未知参量  $\eta$ , 即式中的  $\eta_i$  需用  $\eta$  代替). 若协方差矩阵  $V(Y)$  已知, 据此即可求得物理量及其方差的最优估计  $\hat{\eta}$ ,  $\sigma_{\hat{\eta}}^2$ .

当这  $N$  次实验测量结果是相互独立、不相关联的, 协方差矩阵仅对角元不为 0, 式 (6.2.2) 简化为

$$Q^2(\eta) = \sum_{i=1}^N \left( \frac{y_i - \eta}{\sigma_i} \right)^2. \quad (6.2.5)$$

$\eta$  的最优估计及其方差则为

$$\hat{\eta} = \frac{\sum_{i=1}^N \omega_i y_i}{\sum_{i=1}^N \omega_i}, \quad (6.2.6)$$

$$\sigma_{\hat{\eta}}^2 = \frac{1}{\sum_{i=1}^N \omega_i}, \quad (6.2.7)$$

$$\omega_i = \frac{1}{\sigma_i^2}. \quad (6.2.8)$$

可以看到, 这与利用极大似然法对同一物理量作  $N$  次独立观测求期望值及其方差的式 (6.1.9) 完全相同.

假定测量值  $y_i$  是期望值  $\eta$ , 方差  $\sigma_{\eta}^2$  的正态变量, 式 (6.2.5) 表示的  $Q^2$  是  $N-1$  个独立的标准正态变量的平方和 (有一个待定参数  $\eta$ ), 因而是自由度  $N-1$  的  $\chi^2$  变量,  $Q^2(\hat{\eta})$  的期望值为  $N-1$ . 可见, 在测量值服从正态分布的假设下, 量  $Q^2(\hat{\eta})$

的值与自由度  $N - 1$  的差异可以反映不同测量值  $Y_i$  之间的关联程度. 当  $Q^2(\hat{\eta})$  接近  $N - 1$  时, 各测量值之间是相互近似独立的. 反之, 若  $Q^2(\hat{\eta})$  与  $N - 1$  差别明显, 则可能是各次测量报道的误差  $\sigma_i$  不精确, 或者各次测量之间存在不可忽略的相互关联. 下面我们来讨论后一种情况.

即使在多次测量存在关联的情形下, 式 (6.2.6) 求得的参数估计值  $\hat{\eta}$  虽然不一定是最优估计, 但只要每个测量值  $y_i$  是  $\eta$  的无偏估计, 则  $\hat{\eta}$  也是  $\eta$  的无偏估计. 由于我们现在处理的是协方差矩阵未知的情形, 无法用式 (6.2.3) 求得  $\eta$  的精确估计, 所以我们仍然用式 (6.2.6) 计算  $\eta$  的估计值  $\hat{\eta}$ . 以下的讨论集中在如何处理多次测量存在关联, 但协方差矩阵未知条件下方差  $\sigma_{\hat{\eta}}^2$  的估计问题.

(1)  $Q^2(\hat{\eta}) > N - 1$  的情形.

这种情形对应于多次测量值之间存在负关联, 即协方差矩阵的非对角矩阵元联系系数为负值. 当多次测量相互独立, 这种情形对应于各次测量的误差  $\sigma_i$  可能过低估计了, 于是利用名义  $\sigma_i$  值和式 (6.2.7) 求得的  $\sigma_{\hat{\eta}}^2$  值将小于真实方差. 文献 [115] 建议的处理方法是定义标度因子  $f$

$$f = Q^2(\hat{\eta}) / (N - 1), \quad (6.2.9)$$

将式 (6.2.8) 中的  $\sigma_i^2$  用  $f\sigma_i^2$  代替来求得  $\sigma_{\hat{\eta}}^2$ , 这相当于将原来的方差增大了  $f$  倍. 这样得到的方差可能是偏大的, 因而是保守的, 因为在负关联的情形下, 式 (6.2.7) 求得的方差值已经大于真实的方差了. 这种保守的处理是为了保证实验结果的稳健性.

(2)  $Q^2(\hat{\eta}) < N - 1$  的情形.

这种情形对应于多次测量值之间存在正关联, 即协方差矩阵的非对角矩阵元联系系数为正值. 这时用式 (6.2.7) 求得的方差  $\sigma_{\hat{\eta}}^2$  可能比真实的方差小. 文献 [116] 对这种情况建议的处理方法是建立一个等效的协方差矩阵  $C$ , 其矩阵元为

$$C_{ii} = \sigma_i^2, \quad C_{ij} = f\sigma_i\sigma_j, \quad i \neq j; \quad i, j = 1, 2, \dots, N, \quad (6.2.10)$$

即认为不同测量之间的关联系数同为正常数  $f$ ,  $f$  由下式求得:

$$\chi^2(f) = \sum_{i,j=1}^N (y_i - \hat{\eta})(y_j - \hat{\eta})(C^{-1})_{ij} = N - 1. \quad (6.2.11)$$

由此可求得估计量  $\hat{\eta}$  的方差

$$\sigma_{\hat{\eta}}^2 = \frac{\sum_{i,j=1}^N \omega_i \omega_j C_{ij}}{\left( \sum_{i=1}^N \omega_i \right)^2}. \quad (6.2.12)$$

当  $f = 0, C_{ij} = 0(i \neq j), \sum_{i,j=1}^N \omega_i \omega_j C_{ij} = \sum_{i=1}^N \omega_i^2 C_{ii} = \sum_{i=1}^N \omega_i$  时, 式 (6.2.12) 回复到  $N$  次独立测量情况下  $\sigma_{\hat{\eta}}^2$  的表达式 (6.2.7).

式 (6.2.12) 可表示为

$$\sigma_{\hat{\eta}}^2 = \frac{\sum_{i,j=1}^N \omega_i \omega_j C_{ij}}{\left(\sum_{i=1}^N \omega_i\right)^2} = \frac{\sum_{i=1}^N \omega_i + 2 \sum_{j>i}^N \omega_i \omega_j C_{ij}}{\left(\sum_{i=1}^N \omega_i\right)^2} = \frac{1}{\sum_{i=1}^N \omega_i} + \frac{2 \sum_{j>i}^N \omega_i \omega_j C_{ij}}{\left(\sum_{i=1}^N \omega_i\right)^2}.$$

其中, 右边第一项正是式 (6.2.7), 即不考虑关联情形下的方差估计, 而第二项当  $f$  大于 0 时恒为正. 故可知, 考虑各测量值之间存在正关联情形下求得的误差恒大于各测量值独立情形下的误差.

**示例** 强作用常数  $\alpha_S$  的合并估计 [117].

设在不同的质心系能量下测定的强作用常数在  $Z^0$  粒子质量处的值  $\alpha_S(M_Z)$  的数据如表 6.2 所示, 不同的质心系能量下测定的  $\alpha_S(M_Z)$  是根据  $\alpha_S$  质心系能量的依赖关系推定的, 因此  $\alpha_S(M_Z)$  的误差都包含了理论 (模型) 误差, 即测量误差是相关的, 试确定它们的合并估计.

表 6.2 强作用常数  $\alpha_S(M_Z)$  的测量数据

$E_{\text{cm}}/\text{GeV}$	$\alpha_S(M_Z)$	$\sigma$
183	0.121	0.006
189	0.121	0.005
195	0.122	0.006
201	0.124	0.006
206	0.124	0.006

**解** 利用协方差未知情形下的解法求解. 先根据多次独立测量的式 (6.2.6) 求得合并估计

$$\hat{\alpha}_S(M_Z) \equiv \hat{\eta} = \frac{\sum_{i=1}^N \omega_i y_i}{\sum_{i=1}^N \omega_i} = 122.3 \times 10^{-3}, \quad \omega_i = \frac{1}{\sigma_i^2}.$$

若用独立测量的式 (6.2.5) 计算  $Q^2(\hat{\eta})$ , 则有

$$Q^2(\hat{\eta}) = 0.28 < N - 1 = 4,$$

这种情形对应于多次测量之间存在正关联, 用独立测量的式 (6.2.7) 计算求得的  $\alpha_S(M_Z)$  的合并估计误差

$$\sigma_{\hat{\eta}} = \frac{1}{\sqrt{\sum_{i=1}^N \omega_i}} = 2.6 \times 10^{-3}$$

可能比真实的方差小.

现在按式 (6.2.10) 建立等效协方差矩阵  $C$  (为计算方便起见, 下面的计算中暂时忽略数据中的因子  $10^{-3}$ , 到最后再加上):

$$C = \begin{pmatrix} 36 & 30f & 36f & 36f & 36f \\ 30f & 25 & 30f & 30f & 30f \\ 36f & 30f & 36 & 36f & 36f \\ 36f & 30f & 36f & 36 & 36f \\ 36f & 30f & 36f & 36f & 36 \end{pmatrix},$$

用软件包 Mathematica 求逆矩阵得

$$C^{-1} = \begin{pmatrix} a & b & d & d & d \\ b & e & b & b & b \\ d & b & a & d & d \\ d & b & d & a & d \\ d & b & d & d & a \end{pmatrix},$$

其中

$$a = \frac{1+3f}{36(1+3f-4f^2)}, \quad b = \frac{-f}{30(1+3f-4f^2)},$$

$$d = \frac{-f}{36(1+3f-4f^2)}, \quad e = \frac{1+3f}{25(1+3f-4f^2)}.$$

根据式 (6.2.11) 有

$$\chi^2(f) = \sum_{i,j=1}^N (y_i - \hat{\eta})(y_j - \hat{\eta})(C^{-1})_{ij} = N - 1 = 4$$

$$= \begin{pmatrix} -1.3 & -1.3 & -0.3 & 1.7 & 1.7 \end{pmatrix} C^{-1} \begin{pmatrix} -1.3 \\ -1.3 \\ -0.3 \\ 1.7 \\ 1.7 \end{pmatrix},$$

由此可求得  $(f - 0.930679)(f + 0.249979) = 0$ , 即  $f$  有两个解,  $f = -0.2500$  和  $f = 0.9307$ . 前面已述,  $Q^2(\hat{\eta}) = 0.28 < N - 1 = 4$  这种情形对应于多次测量之间存在正关联, 所以应取  $f = 0.9307$ .

根据式 (6.2.12) 求得估计量  $\hat{\eta}$  的方差

$$\sigma_{\hat{\eta}}^2 = \frac{\sum_{i,j=1}^N \omega_i \omega_j C_{ij}}{\left(\sum_{i=1}^N \omega_i\right)^2} = \frac{1}{\sum_{i=1}^N \omega_i} + \frac{2 \sum_{j>i}^N \omega_i \omega_j C_{ij}}{\left(\sum_{i=1}^N \omega_i\right)^2} = 2.6^2 + 24.45 = 31.21,$$

于是有 (现在加上因子  $10^{-3}$ )

$$\sigma_{\hat{\eta}} = 5.59 \times 10^{-3}.$$

它比用独立测量的式 (6.2.7) 计算求  $\alpha_S(M_Z)$  的合并估计误差  $\sigma_{\hat{\eta}} = 2.6 \times 10^{-3}$  明显要大. 最终得到  $\alpha_S(M_Z)$  的合并估计及误差为

$$\hat{\alpha}_S(M_Z) = 0.1223 \pm 0.0056.$$

我们注意到, 5 次测量值的合并估计量  $\hat{\eta}$  的误差  $\sigma_{\hat{\eta}} = 0.0056$  与单次测量值的误差 (0.006) 几乎没有什么差别, 这一事实可以这样理解: 因为每一单次测量中包含公共的误差, 这一公共误差占了误差的绝大部分 (这由各测量值之间存在着  $f = 0.9307$  的正关联可以断定), 而测量次数的增加不可能减小这种公共误差.

## 6.3 分支比多个测量值的合并估计

粒子物理实验中, 对于某一短寿命粒子的某一特定末态的衰变分支比, 通常会有不同的实验进行测量, 或者同一实验通过不同的衰变道进行测量. 将这些不同的测量结果进行统计意义上的合并, 通常会获得比单个测量结果更高的精度.

对同一物理量作  $I$  次独立测量, 获得的观测值为  $x_i \pm \sigma_i$  ( $i = 1, 2, \dots, I$ ), 假定测量服从正态分布, 则这  $I$  次独立测量的合并估计为  $\mu \pm \sigma$ , 其中

$$\mu = \frac{\sum_{i=1}^I \frac{x_i}{\sigma_i^2}}{\sum_{i=1}^I \frac{1}{\sigma_i^2}}, \quad \sigma = \frac{1}{\sqrt{\sum_{i=1}^I \frac{1}{\sigma_i^2}}}. \quad (6.3.1)$$

一般情形下可以利用式 (6.3.1) 求得分支比多次测量值的合并估计. 但是在某些情形下, 某些实验由于测量到的特定末态事例数过少, 只能给出一定置信水平下的分支比上限. 这时就无法利用该公式求得不同实验的合并估计.

我们讨论一种特定的但在粒子物理实验中会经常遇到的情形. 在分支比测量的各单个实验中, 利用某个特定的选择判据, 从实验原始数据中选择出待测分支比相对应的信号道的候选信号事例, 后者包含真正的信号事例和本底事例两部分. 在一个适当的运动学变量 (特征变量) 的实验观测谱的信号区内, 信号事例的谱型和本底事例的谱型具有显著的差异, 并且通常可以利用 MC 模拟或实验的控制样本数据求得. 这样, 可以通过特征变量谱的拟合来求得信号事例数, 从而得到待测的分支比.

为了便于说明问题, 我们仍然以 2.2.5 节中叙述过的 BES 实验  $\psi(2S) \rightarrow \eta J/\psi$  分支比测量为例<sup>[24]</sup>. 实验测量的末态是  $\gamma\gamma e^+e^-$  和  $\gamma\gamma\mu^+\mu^-$  事例, 将轻子对的不变质量约束为  $J/\psi$  质量后, 可得到候选信号事例的特征量——两光子不变质量  $M_{\gamma\gamma}$  的分布, 如图 2.9 所示. 图中 548MeV 处的峰对应于  $\psi(2S) \rightarrow \eta J/\psi$  信号, 而平滑的宽分布则对应于本底. 如果信号函数和本底函数已知, 则通过实验观测到的两光子不变质量  $M_{\gamma\gamma}$  谱的拟合可以确定信号事例数, 从而计算出分支比. 分支比的计算公式是

$$B(R \rightarrow X) = \frac{N_s(R \rightarrow X \rightarrow Y)}{N_R \varepsilon(R \rightarrow X \rightarrow Y) \cdot BR(X \rightarrow Y)}. \quad (6.3.2)$$

式中,  $R$  标记衰变粒子  $\psi(2S)$ ,  $X$  标记  $\eta J/\psi$ ,  $Y$  标记  $\gamma\gamma e^+e^-$  或  $\gamma\gamma\mu^+\mu^-$ ;  $N_R$  是原始数据中包含的粒子  $R$  的全部衰变事例数,  $\varepsilon$  是信号事例的探测效率, 假定  $BR(X \rightarrow Y)$ ,  $N_R$  和  $\varepsilon$  通过其他测量为已知;  $N_s$  是本实验中的观测信号事例数. 于是分支比公式可写为 (符号已作了简化)

$$B_i = \frac{N_{is}}{N_{Ri}\varepsilon_i BR_i} \equiv \frac{N_{is}}{A_i}, \quad i = 1(\gamma\gamma e^+e^-), 2(\gamma\gamma\mu^+\mu^-). \quad (6.3.3)$$

式中, 带下标  $i$  的符号标记第  $i$  个实验测量的值, 式 (6.3.3) 的形式很容易推广到多个测量的情形. 该实验中通过  $J/\psi \rightarrow e^+e^-$  和  $J/\psi \rightarrow \mu^+\mu^-$  两种衰变方式求得各自的  $\psi(2S) \rightarrow \eta J/\psi$  衰变分支比为  $2.91 \pm 0.12$  和  $3.06 \pm 0.14$  (误差均为统计误差). 假定这两个测量是统计意义上独立的, 利用式 (6.3.1) 可求得两者的合并估计为  $2.98 \pm 0.09$  (误差为统计误差). 但是, 如果其中的一个结果是分支比上限, 则无法进行合并估计.

本节讨论同一分支比的多个测量结果 (包含上限) 的合并估计. 我们给出的方法对于单个测量结果是否用上限来表示没有限制. 利用所有单个实验的实验测量谱数据构建相应的联合最小二乘 ( $\chi^2$ ) 函数或联合似然函数, 通过联合  $\chi^2$  函数的极小化或联合似然函数的极大化求得分支比的合并估计.

联合  $\chi^2$  函数或联合似然函数的构建方式取决于各个实验所提供的特征变量观测谱形式. 下面根据各个实验特征变量观测谱的不同形式阐述合并估计的不同方法.

### 6.3.1 合并估计的最小二乘法

设  $I$  个实验测量同一衰变分支比,  $I$  个实验谱为特征变量  $m_i$  ( $i = 1, 2, \dots, I$ ) 的直方图数据, 每个实验谱的特征变量  $m_i$  可以各不相同 (或相同), 各直方图的信号区间和子区间划分也可以不同 (或相同). 设实验  $i$  信号区间内总事例数为  $N_i$ , 直方图子区间数为  $J_i$ , 子区间  $j_i$  ( $i = 1, 2, \dots, J_i$ ) 中的观测事例数为  $n_{ij_i}$ , 联合最小二乘函数  $\chi^2$  定义为

$$\chi^2 = \sum_{i=1}^I \chi_i^2 = \sum_{i=1}^I \sum_{j_i=1}^{J_i} \frac{(n_{ij_i} - N_i p_{ij_i})^2}{N_i p_{ij_i}}. \quad (6.3.4)$$

其中,  $p_{ij_i}$  是一个事例出现在子区间  $j_i$  中的概率:

$$p_{ij_i} = p_{ij_i}(\theta_i) = \int_{\Delta m_{j_i}} f_i(m_i | \theta_i) dm_i, \quad (6.3.5)$$

$f_i(m_i | \theta_i)$  为实验  $i$  中信号区间内信号候选事例特征变量  $m_i$  的概率密度:

$$f_i(m_i | \theta_i) = w_{is} f_{is}(m_i | \theta_{is}) + (1 - w_{is}) f_{ib}(m_i | \theta_{ib}). \quad (6.3.6)$$

式中,  $\theta_{is}$  和  $\theta_{ib}$  分别是描述信号区间内信号和本底的概率密度  $f_{is}$  和  $f_{ib}$  的参数 (可以是多个参数);  $w_{is}$  表示信号区间内信号事例占全部事例的比例, 即  $N_{is} = w_{is} N_i$ .  $f_{is}$  和  $f_{ib}$  的形式在实验  $i$  的数据分析中应当已经知道.

由式 (6.3.3) 可知实验  $i$  中信号区间内的信号事例数  $N_{is}$  为

$$N_{is} = A_i B_i, \quad A_i \equiv N_{Ri} \varepsilon_i B R_i, \quad i = 1, 2, \dots, I. \quad (6.3.7)$$

当进行合并估计时,  $B = B_i$ . 由  $N_{is} = w_{is} N_i$  知  $w_{is} = A_i B / N_i$ , 代入式 (6.3.6) 得

$$f_i(m_i | \theta_i) = \frac{A_i}{N_i} B \cdot f_{is}(m_i | \theta_{is}) + \left(1 - \frac{A_i}{N_i} B\right) f_{ib}(m_i | \theta_{ib}). \quad (6.3.8)$$

这时, 式 (6.3.4) 所示的  $\chi^2$  表达式中的未知参数  $\theta$  包括  $\theta_s$ ,  $\theta_b$  和  $B$ :

$$\theta = \{B, \theta_s, \theta_b\}, \quad \theta_s = \{\theta_{1s}, \theta_{2s}, \dots, \theta_{Is}\}, \quad \theta_b = \{\theta_{1b}, \theta_{2b}, \dots, \theta_{Ib}\}. \quad (6.3.9)$$

利用最优化程序包对函数  $\chi^2$  求极小可求得未知参数  $\theta$  的估计值  $\hat{\theta}$  (包括分支比合并估计  $\hat{B}$ ) 及其统计 (拟合) 误差 (包括  $B$  的统计误差  $\sigma_{B,st}$ ).

如果  $I$  个实验谱为特征变量  $m_i$  ( $i = 1, 2, \dots, I$ ) 的散点图数据, 则首先要变换成直方图数据, 再通过上述步骤求得未知参数  $\theta$  的估计值  $\hat{\theta}$  及其统计 (拟合) 误差, 这一散点图-直方图变换会损失部分测量信息从而增加估计值的误差.

### 6.3.2 合并估计的极大似然法

#### 1. 常规极大似然法

首先讨论  $I$  个实验谱为直方图数据的情形.

设  $I$  个实验测量同一个分支比,  $I$  个实验谱为特征变量  $m_i$  ( $i = 1, 2, \dots, I$ ) 的直方图数据, 每个实验谱的特征变量  $m_i$  可以各不相同 (或相同), 各信号区间和直方图的子区间划分也可以不同 (或相同). 设实验  $i$  信号区间内的总事例数为  $N_i$ , 直方图子区间数为  $J_i$ , 子区间  $j_i (= 1, 2, \dots, J_i)$  中的观测事例数为  $n_{ij_i}$ . 第  $j_i$  个子区间内观测到  $n_{ij_i}$  个事例 ( $j_i = 1, 2, \dots, J_i$ ) 的联合概率为多项分布函数:

$$L_i(n_{i1}, n_{i2}, \dots, n_{iJ_i}) = N_i! \prod_{j_i=1}^{J_i} \frac{p_{ij_i}(m_i|\theta_i)^{n_{ij_i}}}{n_{ij_i}!}, \quad i = 1, 2, \dots, I, \quad (6.3.10)$$

$p_{ij_i}(m_i|\theta_i)$  由信号区间内的概率密度  $f_i(m_i|\theta_i)$  在第  $j_i$  子区间中的积分求出:

$$p_{ij_i}(m_i|\theta_i) = \int_{\Delta m_{j_i}} f_i(m_i|\theta_i) dm_i, \quad (6.3.11)$$

$f_i(m_i|\theta_i)$  为实验  $i$  中信号区间内特征变量  $m_i$  的概率密度, 仍由式 (6.3.8) 描述.

于是,  $I$  个实验的联合似然函数为

$$L(\theta) = \prod_{i=1}^I L_i = \prod_{i=1}^I \left[ N_i! \prod_{j_i=1}^{J_i} \frac{p_{ij_i}(m_i|\theta_i)^{n_{ij_i}}}{n_{ij_i}!} \right]. \quad (6.3.12)$$

不考虑与待拟合参数  $\theta$  无关的项, 有

$$\ln L = \sum_{i=1}^I \sum_{j_i=1}^{J_i} n_{ij_i} \ln p_{ij_i}(m_i|\theta_i). \quad (6.3.13)$$

联合似然函数中的未知参数  $\theta$  包括  $\theta_s$ ,  $\theta_b$  和  $B$ :

$$\theta = \{B, \theta_s, \theta_b\}, \quad \theta_s = \{\theta_{1s}, \theta_{2s}, \dots, \theta_{Is}\}, \quad \theta_b = \{\theta_{1b}, \theta_{2b}, \dots, \theta_{Ib}\}. \quad (6.3.14)$$

利用最优化程序包对变量  $\ln L$  求极大可求得未知参数  $\theta$  的估计值  $\hat{\theta}$  (包括分支比合并估计  $\hat{B}$ ) 及其统计 (拟合) 误差 (包括  $B$  的统计误差  $\sigma_{B,st}$ ). 在最优化计算中,  $\theta_s$ ,  $\theta_b$  的初值可取为各个实验中确定的值, 而  $B$  的初值可利用各个实验确定的  $B_i$  的某种加权平均.

在各实验  $i$  信号区间内的总事例数  $N_i$  充分大, 且每个子区间内特征变量  $m_i$  的概率密度  $f_i(m_i|\theta_i)$  的变化不显著的情形下, 这类常规极大似然法对于  $I$  个实验谱为直方图数据的情形可给出相当可靠的分支比合并估计  $\hat{B}$  及其统计误差  $\sigma_{B,st}$ .

现在来讨论  $I$  个实验谱为散点图数据的情形.

设  $I$  个实验测量同一分支比给出的观测谱为不同 (或相同) 特征变量  $m_i$  在不同 (或相同) 信号区间内的散点图数据. 令实验  $i$  散点图的事例数为  $N_i$ , 这  $N_i$  个事例出现在  $m_{i1}, m_{i2}, \dots, m_{iN_i}$  ( $i = 1, 2, \dots, I$ ) 处. 其信号区间内的概率密度仍然如式 (6.3.8) 所描述.  $I$  个实验的观测谱共有  $N \equiv \sum_{i=1}^I N_i$  个事例. 定义  $N$  个事例的联合似然函数为

$$\begin{aligned} L &= L(m_{11}, m_{12}, \dots, m_{1N_1}; \dots; m_{I1}, m_{I2}, \dots, m_{IN_I} | B, \theta_s, \theta_b) \\ &= \prod_{i=1}^I L_i(m_{i1}, m_{i2}, \dots, m_{iN_i} | B, \theta_{is}, \theta_{ib}) \\ &= \prod_{i=1}^I \prod_{j=1}^{N_i} \left[ \frac{A_i}{N_i} B \cdot f_{is}(m_{ij} | \theta_{is}) + \left( 1 - \frac{A_i}{N_i} B \right) f_{ib}(m_{ij} | \theta_{ib}) \right], \end{aligned} \quad (6.3.15)$$

其对数似然函数为

$$\ln L = \sum_{i=1}^I \sum_{j=1}^{N_i} \ln \left[ \frac{A_i}{N_i} B \cdot f_{is}(m_{ij} | \theta_{is}) + \left( 1 - \frac{A_i}{N_i} B \right) f_{ib}(m_{ij} | \theta_{ib}) \right]. \quad (6.3.16)$$

联合似然函数中的未知参数  $\theta$  包括  $\theta_s$ ,  $\theta_b$  和  $B$ , 如式 (6.3.14) 所示. 利用最优化程序包对变量  $\ln L$  求极大可求得未知参数  $\theta$  的估计值  $\hat{\theta}$  (包括分支比合并估计  $\hat{B}$ ) 及其统计 (拟合) 误差 (包括  $B$  的统计误差  $\sigma_{B,st}$ ).

## 2. 广义极大似然法

首先讨论  $I$  个实验谱为直方图数据的情形.

设  $I$  个实验测量同一个分支比,  $I$  个实验谱为特征变量  $m_i$  ( $i = 1, 2, \dots, I$ ) 的直方图数据, 每个实验谱的特征变量  $m_i$  可以各不相同 (或相同), 各信号区间和直方图的子区间划分也可以不同 (或相同). 设实验  $i$  信号区间内的总事例数为  $N_i$ , 直方图子区间数为  $J_i$ , 子区间  $j_i$  ( $= 1, 2, \dots, J_i$ ) 中的观测事例数为  $n_{ij_i}$ . 在广义极大似然法中,  $n_{ij_i}$  是期望值  $\lambda_{ij_i}$  的泊松变量. 第  $j_i$  个子区间内观测到  $n_{ij_i}$  个事例 ( $j_i = 1, 2, \dots, J_i$ ) 的联合概率为

$$L_i(n_{i1}, n_{i2}, \dots, n_{iJ_i}) = \prod_{j_i=1}^{J_i} \frac{1}{n_{ij_i}!} \lambda_{ij_i}^{n_{ij_i}} e^{-\lambda_{ij_i}}, \quad (6.3.17)$$

$\lambda_{ij_i}$  由信号区间内的概率密度  $f_i(m_i | \theta_i)$  在第  $j_i$  子区间中的积分求出:

$$\lambda_{ij_i} = \lambda_i \int_{\Delta m_{j_i}} f_i(m_i | \theta_i) dm_i, \quad (6.3.18)$$

$\lambda_i$  是实验  $i$  直方图的总事例数  $N_i$  (泊松变量) 的期望值:

$$\lambda_i = \sum_{j_i=1}^{J_i} \lambda_{ij_i}, \quad (6.3.19)$$

$f_i(m_i|\theta_i)$  为实验  $i$  中信号区间内特征变量  $m_i$  的概率密度, 仍由式 (6.3.6) 描述.

在广义极大似然法中, 式 (6.3.7) 实验  $i$  中信号区间内的信号事例数  $N_{is}$  实际上应考虑为一个随机变量, 故应改为其期望值  $\lambda_{is}$ , 因此式 (6.3.7) 需用下式代替:

$$\lambda_{is} = A_i B_i, \quad A_i \equiv N_{Ri} \varepsilon_i B R_i, \quad i = 1, 2, \dots, I. \quad (6.3.20)$$

特征变量  $m_i$  的概率密度  $f_i(m_i|\theta_i)$  中的  $w_{is}$  现在为

$$w_{is} = \lambda_{is} / \lambda_i. \quad (6.3.21)$$

当进行合并估计时,  $B = B_i$ . 由  $\lambda_{is} = w_{is} \lambda_i$ , 知  $w_{is} = A_i B / \lambda_i$ , 代入式 (6.3.6) 得

$$f_i(m_i|\theta_i) = \frac{A_i}{\lambda_i} B \cdot f_{is}(m_i|\theta_{is}) + \left(1 - \frac{A_i}{\lambda_i} B\right) f_{ib}(m_i|\theta_{ib}). \quad (6.3.22)$$

$I$  个实验的联合似然函数定义为

$$L(\theta) = \prod_{i=1}^I L_i = \prod_{i=1}^I \prod_{j_i=1}^{J_i} \frac{\lambda_{ij_i} (m_i|\theta_i)^{n_{ij_i}}}{n_{ij_i}!} e^{-\lambda_{ij_i}}, \quad (6.3.23)$$

不考虑与待拟合参数  $\theta$  无关的项, 有

$$\ln L(\theta) = \sum_{i=1}^I \left[ \sum_{j_i=1}^{J_i} (n_{ij_i} \ln \lambda_{ij_i}) - \lambda_i \right]. \quad (6.3.24)$$

联合似然函数中的未知参数  $\theta$  包括  $\theta_s$ ,  $\theta_b$  和  $\lambda$ ,  $B$ :

$$\begin{aligned} \theta &= \{B, \lambda, \theta_s, \theta_b\}, \quad \lambda = \{\lambda_1, \lambda_2, \dots, \lambda_I\}, \\ \theta_s &= \{\theta_{1s}, \theta_{2s}, \dots, \theta_{Is}\}, \quad \theta_b = \{\theta_{1b}, \theta_{2b}, \dots, \theta_{Ib}\}. \end{aligned} \quad (6.3.25)$$

利用最优化程序包对变量  $\ln L$  求极大可求得未知参数  $\theta$  的估计值  $\hat{\theta}$  (包括分支比合并估计  $\hat{B}$ ) 及其统计 (拟合) 误差 (包括  $B$  的统计误差  $\sigma_{B, \text{st}}$ ). 在最优化计算中,  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_I\}$  的初值可取为  $\{N_1, N_2, \dots, N_I\}$ ,  $\theta_s$ ,  $\theta_b$  的初值可取为各个实验中确定的值, 而  $B$  的初值可利用各个实验确定的  $B_i$  的某种加权平均.

对于  $I$  个实验谱为散点图数据的情形, 实验  $i$  的广义似然函数为

$$L_i(m_i|\theta_i, \lambda_i) = e^{-\lambda_i} \frac{\lambda_i^{N_i}}{N_i!} \prod_{j=1}^{N_i} f_i(m_{ij}|\theta_i),$$

将式 (6.3.22) 代入上式并略去不依赖于  $\theta_i$  的项, 可得

$$L_i(m_i|\theta_i, \lambda_i) = e^{-\lambda_i} \prod_{j=1}^{N_i} [A_i B \cdot f_{is}(m_{ij}|\theta_{is}) + (\lambda_i - A_i B) f_{ib}(m_{ij}|\theta_{ib})]. \tag{6.3.26}$$

因此  $I$  个实验谱的联合广义似然函数为

$$\begin{aligned} L(\theta) &= \prod_{i=1}^I L_i(m_i|\theta_i, \lambda_i) \\ &= \prod_{i=1}^I e^{-\lambda_i} \prod_{j=1}^{N_i} [A_i B \cdot f_{is}(m_{ij}|\theta_{is}) + (\lambda_i - A_i B) f_{ib}(m_{ij}|\theta_{ib})]. \end{aligned} \tag{6.3.27}$$

从而得

$$\ln L(\theta) = \sum_{i=1}^I \ln L_i = - \sum_{i=1}^I \lambda_i + \sum_{i=1}^I \sum_{j=1}^{N_i} \ln [A_i B \cdot f_{is}(m_{ij}|\theta_{is}) + (\lambda_i - A_i B) f_{ib}(m_{ij}|\theta_{ib})]. \tag{6.3.28}$$

联合似然函数中的未知参数  $\theta$  包括式 (6.3.25) 所示的  $\theta_s, \theta_b$  和  $\lambda, B$ . 利用最优化程序包对变量  $\ln L$  求极大可求得未知参数  $\theta$  的估计值  $\hat{\theta}$  (包括分支比合并估计  $\hat{B}$ ) 及其统计 (拟合) 误差 (包括  $B$  的统计误差  $\sigma_{B,st}$ ).

6.3.3 信度区间和上限的确定, 系统误差的考虑

1. 不考虑系统误差时信度区间和上限的确定

现在, 我们有了未知参数  $\theta$  的最小二乘估计或极大似然估计  $\hat{\theta}$  (包括分支比合并估计值  $\hat{B}$  及其统计误差  $\sigma_{B,st}$ ). 我们面临的问题是如何给出分支比的测量结果, 即实验报道合并估计值  $B$  的  $CL = 68.27\%$  区间还是  $CL = 90\%$  的上限.

我们利用 5.8.3 节叙述过的贝叶斯 HPD 区间来确定分支比  $B$  的  $CL = 68.27\%$  区间或  $CL = 90\%$  的上限. 因此, 需要利用贝叶斯后验密度作为 flip-flopping 策略的依据, 这就要用到似然函数. 当利用最小二乘法进行合并估计时, 需要额外构建似然函数, 这是最小二乘合并估计的一个缺陷.

根据 5.8.3 节的讨论, 现在分支比合并估计  $B$  的后验概率密度为

$$h(B/n) = \frac{L(n|B)}{\int_0^1 L(n|B) dB}, \tag{6.3.29}$$

式中,  $n$  表示  $I$  个实验的观测谱数据. 对于各实验谱为直方图和散点图这两种情形, 根据式 (6.3.10) 和式 (6.3.15) 可知有

$$\begin{aligned} n &= \{n_{11}, n_{12}, \cdots, n_{1J_1}; \cdots; n_{I1}, n_{I2}, \cdots, n_{IJ_I}\}, \\ n &= \{m_{11}, m_{12}, \cdots, m_{1N_1}; \cdots; m_{I1}, m_{I2}, \cdots, m_{IN_I}\}. \end{aligned} \tag{6.3.30}$$

式中,  $L(\mathbf{n}|B)$  表示根据式 (6.3.10)、式 (6.3.15)、式 (6.3.23)、式 (6.3.27) 计算似然函数时, 给定  $B$  值, 参数  $\theta$  中除  $B$  之外的参数都取  $\ln L$  达到极大时的估计值  $\hat{\theta}$ . 然后按照式 (5.8.24)~式 (5.8.26) 和式 (5.8.20) 可确定分支比合并估计  $B$  的 68.27% 信度区间或  $CL = 90\%$  的上限.

应当指出, 这样确定的信度区间或上限都没有包括  $B$  的系统误差的贡献.

## 2. 考虑系统误差时信度区间和上限的确定

为了求得  $B$  的合并估计的系统误差, 需要考虑  $I$  个实验测量的关联性.

如果  $B$  是  $I$  个实验独立测量的合并结果,  $B$  的系统误差可用多个独立测量的误差合并公式来估计:

$$\sigma_{B_i, \text{sys}}^2 = \left( \sum_{i=1}^I \sigma_{B_i, \text{sys}}^{-2} \right)^{-1}, \quad (6.3.31)$$

$$\frac{\sigma_{B_i, \text{sys}}^2}{B_i^2} = \frac{\sigma_{N_{Ri}}}{N_{Ri}^2} + \frac{\sigma_{BR_i}^2}{BR_i^2} + \frac{\sigma_{\varepsilon_i}^2}{\varepsilon_i^2} + \frac{\sigma_{N_{ib}}^2}{N_{ib}^2}. \quad (6.3.32)$$

其中,  $\sigma_{B_i, \text{sys}}$ ,  $\sigma_{\varepsilon_i}$ ,  $\sigma_{N_{ib}}$  是第  $i$  个实验对于  $B_i$ ,  $\varepsilon_i$ ,  $N_{ib}$  系统误差的估计;  $N_{ib}$  是第  $i$  个实验中的本底事例数,  $N_{ib} = N_i - A_i B_i$ . 式 (6.3.32) 等式右边的各个量在第  $i$  个实验的数据分析中应当已经知道. 就其性质, 第  $i$  个实验的  $\sigma_{N_{Ri}}$ ,  $\sigma_{BR_i}$ ,  $\sigma_{\varepsilon_i}$ ,  $\sigma_{N_{ib}}$  是相互独立的.

如果  $B$  的  $I$  个实验测量不独立, 在  $\sigma_{B_i, \text{sys}}^2$  中有相互独立的成分  $(\sigma_{B_i, \text{sys}}^2)_{\text{uncom}}$  和相同的成分  $(\sigma_{B, \text{sys}}^2)_{\text{com}} \equiv (\sigma_{B_i, \text{sys}}^2)_{\text{com}}$ , 则  $B$  的系统误差可由下式估计:

$$\sigma_{B, \text{sys}}^2 = (\sigma_{B, \text{sys}}^2)_{\text{uncom}} + (\sigma_{B, \text{sys}}^2)_{\text{com}}, \quad (6.3.33)$$

$$(\sigma_{B, \text{sys}}^2)_{\text{uncom}} = \left( \sum_{i=1}^I (\sigma_{B_i, \text{sys}}^2)_{\text{uncom}} \right)^{-1}. \quad (6.3.34)$$

例如, 在前面提到的 BES 实验  $\psi(2S) \rightarrow \eta J/\psi$  分支比测量中,  $\sigma_{N_{Ri}}^2$  是相同的成分, 而  $\sigma_{BR_i}$ ,  $\sigma_{\varepsilon_i}$ ,  $\sigma_{N_{ib}}$  则是相互独立的成分.

有了合并估计  $B$  的系统误差  $\sigma_{B, \text{sys}}$  的值, 根据 5.8.3 节的讨论, 可以利用后验密度  $h(B|\mathbf{n})$

$$h(B|\mathbf{n}) = \frac{\tilde{L}(\mathbf{n}|B)}{\int_0^1 \tilde{L}(\mathbf{n}|B) dB} \quad (6.3.35)$$

$$\tilde{L}(\mathbf{n}|B) = \int_0^1 \left[ L(\mathbf{n}|\nu) \cdot \frac{1}{\sqrt{2\pi}\sigma_{\nu, \text{sys}}} \exp \left( -\frac{(\nu - B)^2}{2\sigma_{\nu, \text{sys}}^2} \right) \right] d\nu \quad (6.3.36)$$

来确定分支比合并估计值  $B$  的 68.27%信度区间或  $CL=90\%$  的上限, 它们同时包含了统计误差和系统误差的贡献. 式中,  $\sigma_{\nu,\text{sys}}$  表示分支比合并估计  $B = \nu$  时的系统误差, 如可取  $\nu$  等于  $B$  的极大似然估计值  $\hat{B}$  时的系统误差  $\sigma_{\hat{B},\text{sys}}$ .

6.3.4 小结

粒子物理实验中同一分支比往往由不同的实验进行测量, 单个实验的测量结果可能是报道  $CL = 68.27\%$  中心区间, 也可能是  $CL = 90\%$  的上限. 多个实验结果的合并可以提高分支比测量的精度, 但是对仅给出上限的实验结果的合并是一个困难的问题. 本节提出的分支比多个测量的合并估计方法对于单个测量结果是否用上限来表示没有限制. 利用所有单个实验的实验测量谱数据构建相应的联合最小二乘 ( $\chi^2$ ) 函数或联合似然函数, 通过联合  $\chi^2$  函数的极小化或联合似然函数的极大化求得分支比的合并估计.

5.8.4 节中讨论了最小二乘拟合和极大似然拟合方法在分支比估计中的优劣比较, 直方图数据和散点图数据在分支比估计中的优劣比较. 这些讨论在同一分支比多个实验结果的合并估计中同样适用, 这里不再赘述. 结论同样是: 利用极大似然法, 特别是广义极大似然法进行分支比的合并估计能给出更为可靠的结果.

对于  $I$  个实验测量同一分支比给出的观测谱为同一特征变量  $m$  在相同信号区间内的子区间划分相同直方图数据的情形, 或者观测谱为  $m$  在相同信号区间内的一维散点图数据的情形, 这时可以构成  $I$  个实验的特征变量  $m$  的合并谱. 这种特殊情形下的分支比合并估计参见文献 [118].