粒子物理与核物理实验中的数据分析

陈少敏
清华大学

第十四讲：系统误差

本讲要点

- 系统误差的定义
- 系统不确定性与系统错误
- 系统误差的检查与结果并合
- 系统误差分析举例

统计误差与系统误差

统计误差

- 如果我们进行重复实验，结果的涨落会有多大?
- 暗示一些用来定义测量结果可能性的假定。
- 通常在拟合后，根据似然函数的变化得到统计误差的大小。

系统误差

- 由于采用的假设存在不确定性，那么它对结果造成的影响是什么?

模型或理论的不确定性；
测量装置的模型化带来的影响。

- 误差源不会随着实验的重复而发生变化；
- 通常情况下，结果会受到诸如刻度常数，效率，等等此类数值的不确定影响。

注意：对系统误差曾经出现过两种定义。

系统不确定性与错误

定义一：系统效应是包括了诸如本底，选择的偏向性，扫描效率，能量分辨率，角度分辨率，计数器效率随束流与能量的变化，等等。在估计这些系统效应带来的不确定性称为系统误差。

定义二：系统误差是由实验仪器、刻度、实验技术等等的过失造成的，可重复产生的精度不确定性。

例子一：

- 量能器能量从电信号 D 转为物理量 E：$E=(\alpha \pm\Delta\alpha)D$；
- 从观测的衰变事例数 N 计算衰变比率 B：$B=N / [N_T(\varepsilon \pm\Delta\varepsilon)]$ 。

例子二：

- 忘记在测量中考虑温度的影响；
- 在计算过程中对数值取整造成精度上的误差。

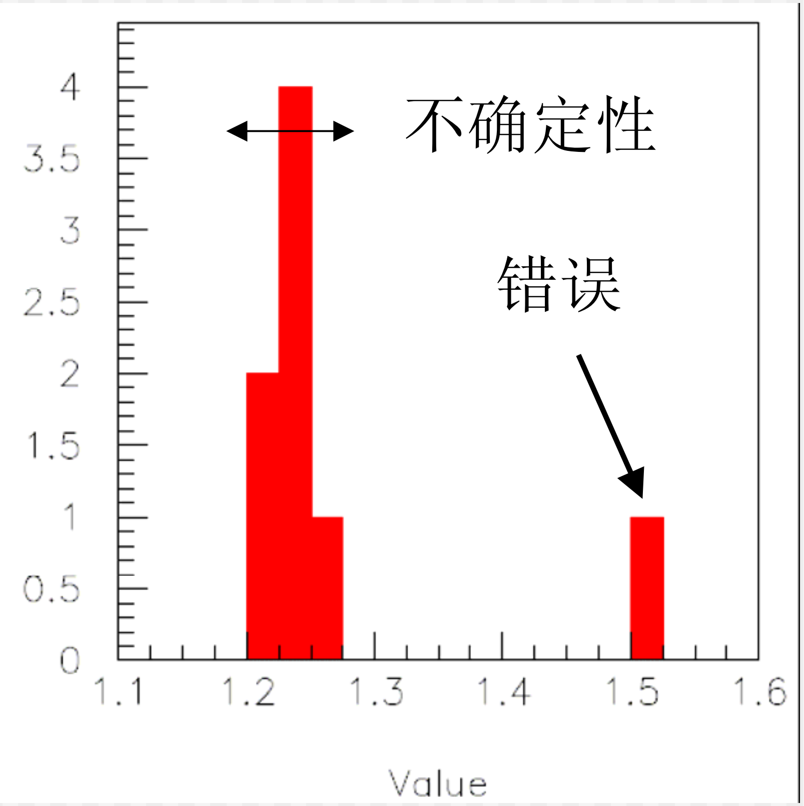
定义的不同表明了处理方式将会有不同。

随机不确定性与错误

在同一测量量给出的几个读数中

1.23, 1.25, 1.24, 1.25,

1.21, 1.52, 1.22, 1.27



可以看出哪些是由不确定性引起的，哪些是由于错误引起的。

- 统计分析提供了用以鉴别和确定不确定性大小的工具。例如通过计算均方差(RMS)的方法估计不确定性。

- 统计分析还提供了如何鉴别一个错误的方法，但它不能告诉我们下一步该如何做，因为它无法告诉我们错误的根源在哪里。

从语义学上定义系统误差

- 物理学家通常将随机(统计)误差定义为随机不确定性而不是随机的错误
- 为了与上述定义保持一致，应该将系统误差定义为系统不确定性而不是系统的错误

systematic error = systematic uncertainty ≠ systematic mistake

⇒ 与定义一相符，而与定义二不符

⇒ 必须把错误结果从所谓的不确定性效应中的误差区分开来

- 系统的错误应始终保持其应有的清晰定义
- 从名称上给出恰当的定义，可以澄清一个问题，那就是统计学并不提供任何工具告诉我们该如何处理系统误差。因此，在所有统计理论的各种参考书中，均没有如何确定系统误差的描述。

系统误差与偏向性

- 历史上有不少实验文章把系统误差与偏向性作为等效处理
- 但是这种处理方法在实际问题显得上不够充分。因为在讨论偏向性时，还必须考虑以下几种情况：

➢我们知道系统有偏向性，然后设法将其消除掉，即可处理完毕；

➢我们没有认识到系统有偏向性，也没有采取任何措施加以处理，这是一种错误；

➢我们知道系统有偏向性，但是不知道偏离的方向和大小。

例如，用一把钢尺测量物体的长度，如何保证结果的准确性...

例子：用钢尺测量物体长度

- 如果伸缩系数精确已知，由于实际测量环境的温度与在对钢尺进行标度时候的温度可能有差异，测量结果可能包含系统偏向性。根据对温度差异的测量，可以对结果进行修正，存在于长度测量过程中的系统偏向性因此得到精确估计。结果修正以后，不存在系统误差。

- 如果温度效应对长度测量的影响被忽略，结果会有错误。要想找到该错误的原因，可以通过一致性检验，利用统计原理揭示可能的结果不一致性，以便研究人员根据常识、经验或直觉来寻找影响的根源。

- 如果温度效应对长度测量可以预测，但是在实验过程中并没有记录对温度的测量值。可以估计实验过程中温度变化的大小，并将此看作是上述系统效应的一种系统不确定性，给出可以接受的系统误差。

系统误差可以是贝叶斯的

- 随机不确定性符合频率论中概率的定义。多次测量的情况下，结果各自有不同。通过概率可以表述结果出现某种极端情况的可能性。
- 但是如果测量含有系统不确定性，根据定义每次观测的结果并不发生改变。这种雷同的结果不能用于表述任何概率的含义，即不符合频率论的定义。

例如：在正负电子对撞实验中，计算有多少反应发生(亮度估计)

Bhabha 事例：$e^+e^- \to e^+e^-$

$$N_{ee} = \int L \bullet \sigma_{ee} dt = \sigma_{ee} \bullet \int L dt \quad \Rightarrow \quad \int L dt = N_{ee} / \sigma_{ee}$$

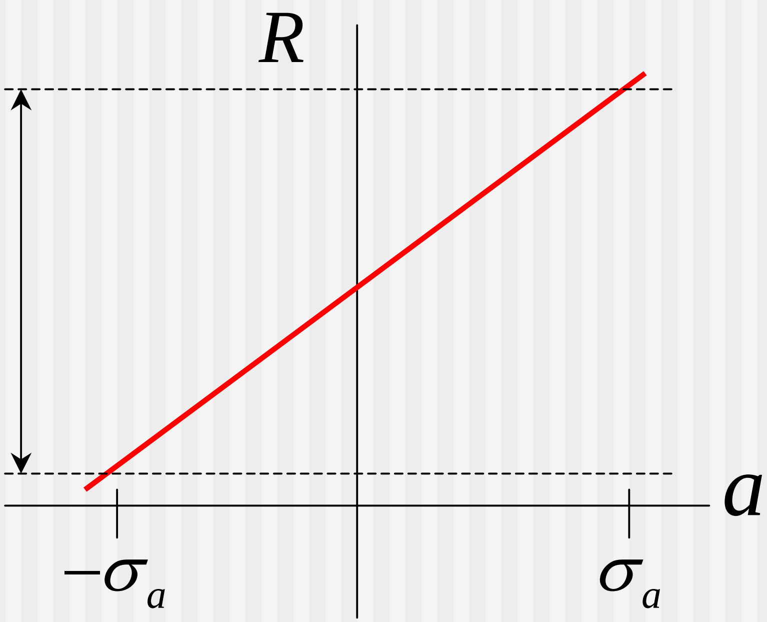
如果理论计算精度只到第三阶 $\Delta\sigma_{ee} \sim O(\alpha^3)$

亮度计算结果总是给出同样的不准确性。

可以猜测这种不准确性(例如第四阶的几倍)，这种带有假设性的估计因此是带有主观性的(或贝叶斯的)概率。

由误差的不确定性估计误差

假设实验结果 R 取决于某些参数 a，对这些参数了解有限，存在某种不确定性 σ_a 。而且它们的不确定性对最终结果的影响不能通过误差传递的代数计算得到。



引用结果为 R(a)，系统不确定性σ_a 。通过采用蒙特卡罗方法计算 R(a-σ_a) 与R(a+σ_a) 。为了更为准确，也可以取几倍的 σ_a ，目的是得到斜率

➡ $R' = dR/da$

➡ 由于a的不确定性导致 $\Delta R = \sigma_a R'$也包含不确定性

问题：如何将这种不确定性较好地反映到R的误差上？

由误差的不确定性估计误差(续)

考虑结果为 R，系统效应为 a。由于对R’的估计本身也可能含有误差 $\sigma_{R'}$

同时 a 也有误差 $a_0 \pm \sigma_a$

它们如何影响R的系统不确定性？三种方案：

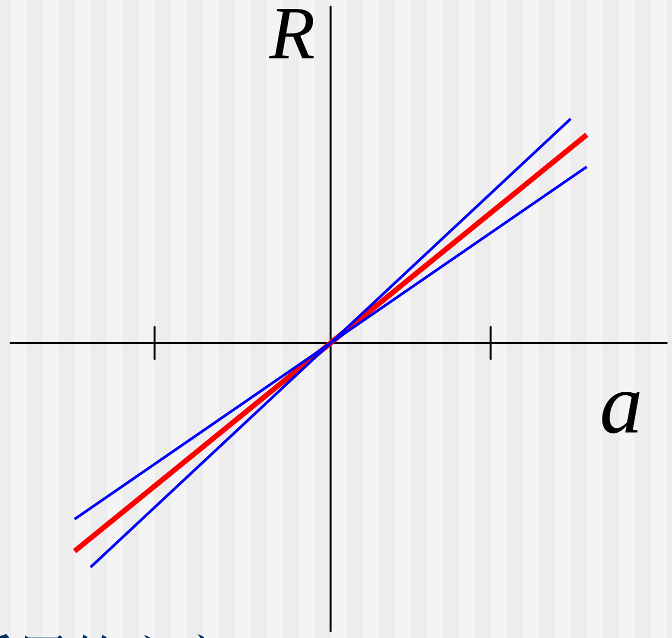
方案一：$\sigma^2 = (R'\sigma_a)^2 + (\sigma_{R'}\sigma_a)^2$

平方和作为R’的不确定性也是另一个不确定性，因此应当给予考虑。

方案二：$\sigma^2 = (R'\sigma_a)^2 - (\sigma_{R'}\sigma_a)^2$

R’已经被改动到偏离真值，因此必须进行相关的补偿。

方案三：$\sigma^2 = (R'\sigma_a)^2$ 我一直采用的方案



如果对结果不会带来太大影响，否则应考虑前两种方案。

如何找出系统错误

忽略系统效应是一个错误。需要研究所有可能的因素(包括各种隐含的因素在内)对结果产生的影响，检查可能出现的系统错误。

- 把数据分成子样本；
- 改变选择条件；
- 改变直方图的区间大小；
- 改变参数形式，包括多项式的次序；
- 改变拟合方法；
- 寻找不可能发生的情况
- 盲分析方法
- 采用两种不同的分析方法

例如在测量电荷与宇称不守恒实验中，有关系统误差研究的描述：
"...一致性检验，包括按衰变类型区分数据，不同标记的种类...，我们还对无电荷宇称不守恒的衰变模式进行了参数拟合，没有发现明显的不对称现象..."

在检验中如何判定显著区别？

标准分析给出 $a_1 \pm \sigma_1$。

检查：不同方法给出 $a_2 \pm \sigma_2$

几乎可以确定 $a_1 \neq a_2$! 通过计算 $\Delta = a_1 - a_2$ ，并且与相关误差 σ_Δ^2 相比较来判断有没有显著的差别。由于两种方法有数据重叠，因此误差为

$$\sigma_\Delta^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$$

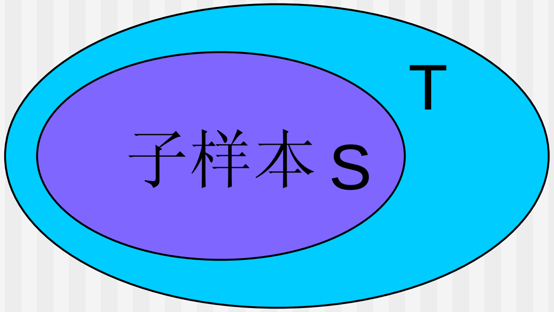
观察 Δ 偏离多少个 σ_Δ

假设估计值是平均值，可以用子样本检验结果

$$a_1 = \frac{1}{N_T}\sum\nolimits_T x_i, \quad a_2 = \frac{1}{N_S}\sum\nolimits_S x_i,$$

$$\sigma_1 = \frac{\sigma}{\sqrt{N_T}}, \quad \sigma_2 = \frac{\sigma}{\sqrt{N_S}}, \quad \mathrm{Cov}(a_1, a_2) = N_S \frac{1}{N_T}\frac{1}{N_S}\sigma^2,$$

$$\rho = \sigma_1/\sigma_2, \quad \Rightarrow \sigma_\Delta^2 = \sigma_2^2 - \sigma_1^2$$



一般情况下检查区别

如果两种分析样本不完全重叠 $\sigma_{\Delta}^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$

引入加权平均 $a(w) = wa_1 + (1-w)a_2$

方差 $\sigma_{a(w)}^2 = w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2w(1-w)\rho\sigma_1\sigma_2$

选择 w 使得方差最小，即 $\dfrac{\partial \sigma_{a(w)}^2}{\partial w} = 0 \quad \rightarrow \quad w$

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

$$\sigma_{\min}^2 = \frac{\sigma_1^2\sigma_2^2(1-\rho^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

对ρ只能采用估计的方法

还可以将两个样本之和构成一个新的样本，求出对应的方差，然后利用第九讲练习中的结果给出方差估计值。

通常研究系统误差的操作

数据分析所进行的步骤：

- 设计选择条件，得到结果；
- 利用似然函数或泊松统计得到随机误差；
- 制作一张大表；
- 将选择条件做随意幅度的改变，结果填入表中；
- 重复上述步骤直到时间/金钱/导师的耐心耗尽；
- 将变化进行平方相加；
- 把得到的结果当作"系统误差"；
- 如果误差太大，把它描述为"保守估计"。

这种将检查错误的误差合并成误差估计是一种不恰当的方式。

➡ 应该把通过检验与不通过检验的结果分别对待。

结果正误判断

- 结果有问题吗?

所下的结论取决于分歧的大小，检查的次数以及基本的合理性。

- 如果全都通过，就不要改动结果。

注意：不要加上(小)的分歧到系统误差上。

- 如果没有通过

1) 检查验证的结果，找出并改正错误；
2) 检查分析，找出并改正错误；
3) 担心某种效应会出现，该检查则变为一种估计；
4) 担心它只是冰山中的一角，应从另一侧面进行检验知道确认无误。

- 最后是将所有的系统误差并合。

一些建议

- 不要把“系统效应”或“系统错误”与系统误差混为一谈；
- 应时刻清楚你所作的是检查错误还是估计不确定性；
- 不要把成功通过检查的结果并入总的系统误差而为将来的结果验证造成障碍；
- 不要急于将失败的检验结果并入总的系统误差中，除非你确实知道失败的原因并且采取了对结果进行了修正的步骤；
- 要把所有检验的工作都准确反映到结果的论述上。

例子:粒子寿命测量

- 寿命定义

$$P(L;\tau_0)=\frac{1}{\tau_0}\exp(-\frac{t}{\tau_0})=\frac{1}{c_{光速}\beta\gamma\tau_0}\exp(-\frac{L}{c_{光速}\beta\gamma\tau_0})=\frac{1}{<l>}\exp(-\frac{L}{<l>})$$

$$c_{光速}\beta\gamma=c_{光速}\sqrt{\frac{E_\tau^2}{m_\tau^2}-1}=25.459\times c_{光速}=7.6323\times10^{11}cm/s$$

- 实验中观测量的分布

$$F(L';<l>,k)=\alpha\bullet f(L';<l>,k)+(1-\alpha)BG_{本底}(L')$$

$$f(L';<l>,k)=\int_{-\infty}^{+\infty}P(L;<l>,k)\bullet Gauss(L-L',\sigma_L)dL$$

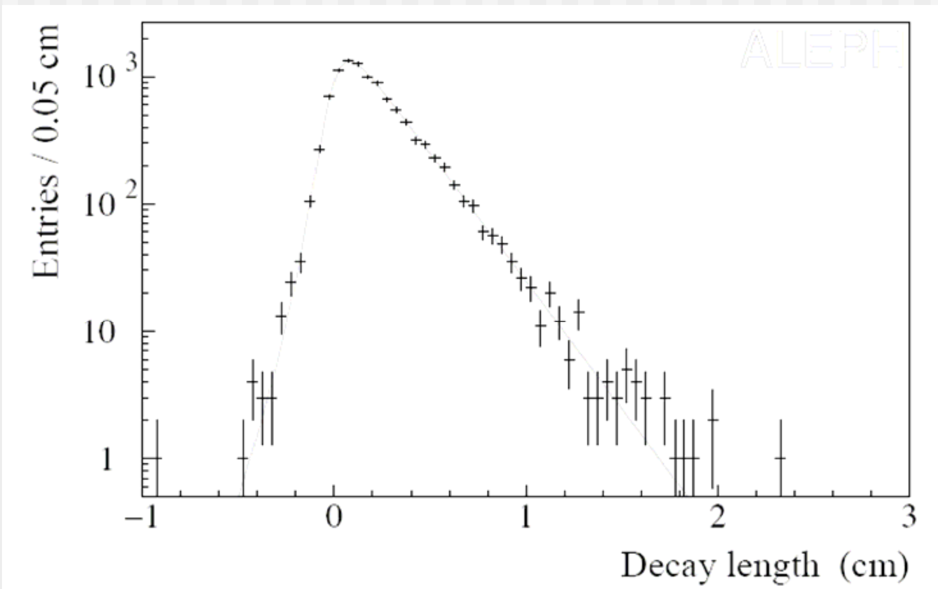
$\alpha=$信号所占的比率(由蒙特卡罗确定)

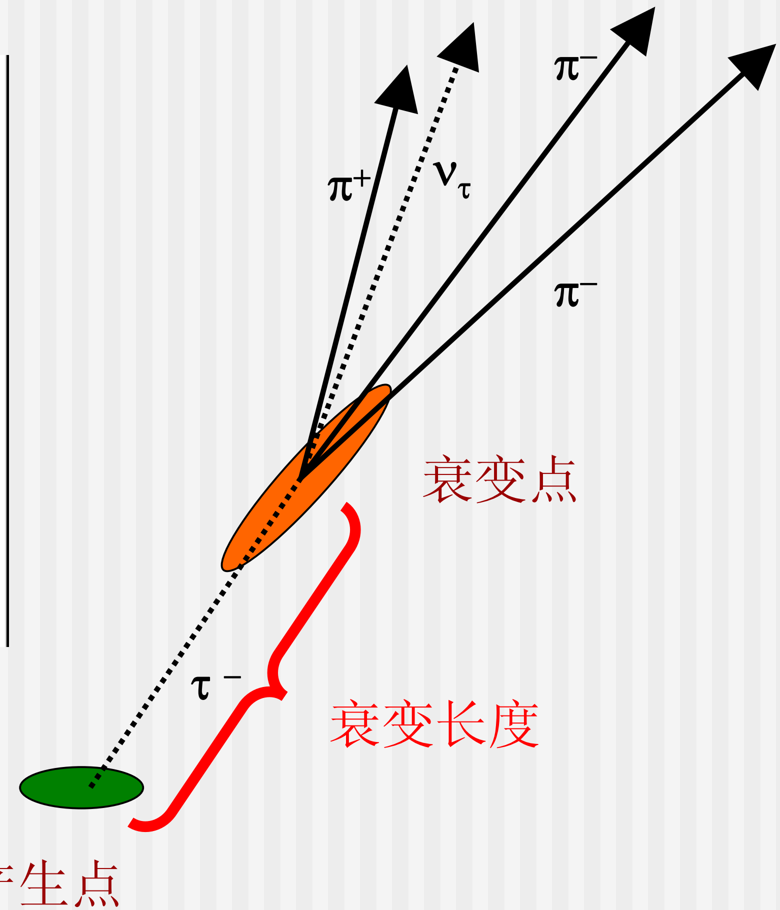
拟合参数:$<l>,k$

$\sigma_L=k\bullet\sigma$（蒙特卡罗确定σ，修正量k由拟合确定）

例子:粒子寿命测量(续一)

• 实验观测的衰变长度分布

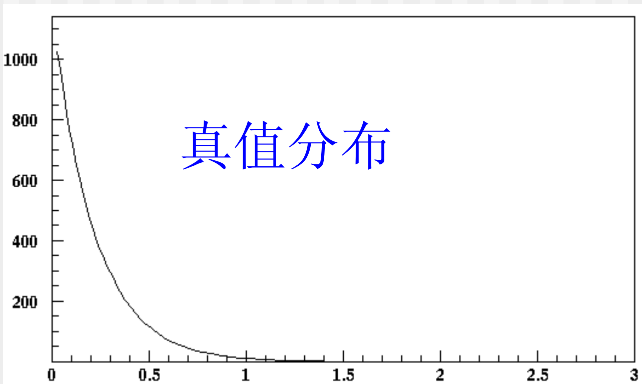




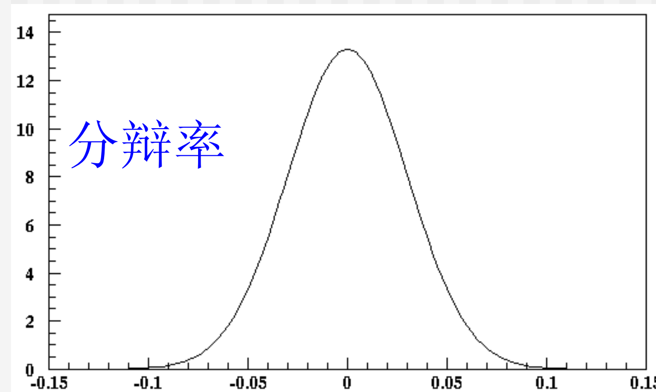
• 影响因素

a) 在衰变长度测量上的误差
b) 探测效率随衰变长度的变化关系
c) 本底对衰变长度分布的影响

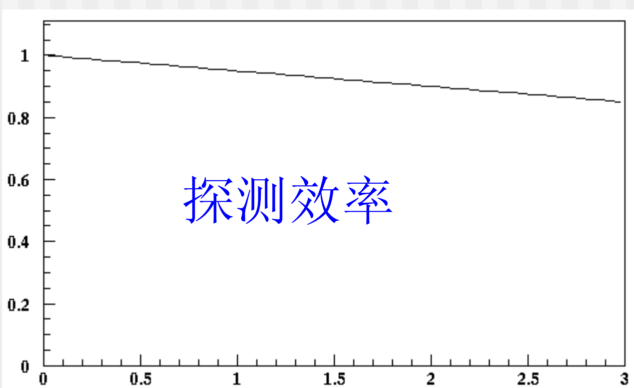
例子:粒子寿命测量(续二)



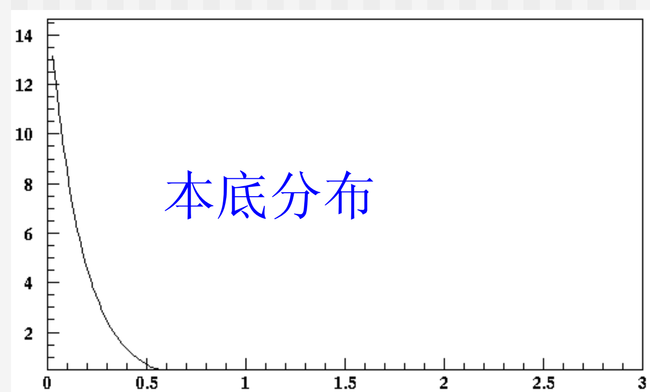
⊗



⊗



⊗



= 观测分布

例子:粒子寿命测量(续三)

似然函数

$$\log L = \sum_{i=1}^{\text{事例数}} F(L_i^{'}; <l>, k)$$

对数据的拟合结果

$$<l> = 0.2174 \pm 0.0023 \text{ cm}$$

$$k = 1.20 \pm 0.02$$

统计误差

注意：统计误差除了包括事例数统计涨落的贡献以外，还包括了因拟合分辨率调整因子 k 而造成的误差贡献。

例子:粒子寿命测量(续四)

- 在无本底样本中，用蒙特卡罗样本检验分析拟合程序，发现衰变长度与真值偏离的幅度为

$$\frac{\text{拟合值}-\text{真值}}{\text{真值}}=(-0.81\pm0.38)\%$$

- 在有本底样本中，用蒙特卡罗样本检验分析拟合程序，在扣除上述系统偏离后，仍然发现衰变长度与真值有一定的偏离

$$\frac{\text{拟合值}-\text{真值}}{\text{真值}}=(-0.26\pm0.09)\%$$

- 将重建衰变位置的拟合优度的限制条件改变，观察数据与蒙特卡罗样本中衰变长度的变化是否一致

数据：Δ_1 蒙特卡罗：Δ_2

$$\text{变化范围}:\frac{\Delta}{\text{真值}}=\frac{\Delta_1-\Delta_2}{\text{真值}}=\pm0.42\%$$

检查其它已知误差源的贡献，发现全部小于0.05%，故忽略不计

例子:粒子寿命测量(续五)

- 系统误差的合并

Source	Bias and uncertainty (%)
$\tau^+\tau^-$ Monte Carlo bias	-0.81 ± 0.38
Backgrounds	-0.26 ± 0.09
Pattern recognition errors	± 0.42
Total bias and uncertainty	-1.07 ± 0.57

- 结果修正

$$L_0 = \frac{<l>}{100\% + (-1.07 \pm 0.57)\%} = (0.21975 \pm 0.00233_{统计} \pm 0.00125_{系统})cm$$

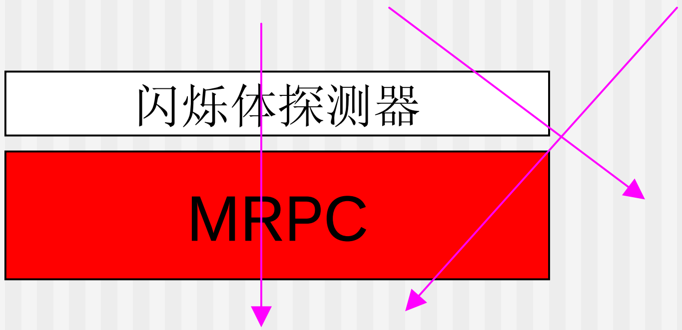
- 寿命测量最终结果

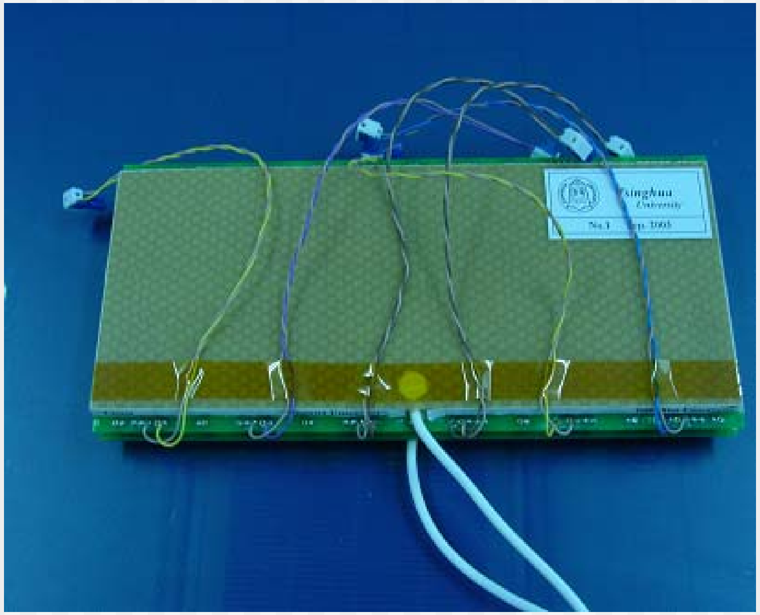
$$\tau_0 = \frac{L_0}{c_{光速}\beta\gamma} = (287.9 \pm 3.1_{统计} \pm 1.6_{系统}) \times 10^{-15} s$$

例子：系统误差的消除

多层阻性板(MRPC)探测器效率的测量

$$\text{效率} = \frac{\text{探测器探测到的事例数}}{\text{入射粒子数}}$$





- 用宇宙线来测量MRPC的探测效率；
- 入射的宇宙线粒子数由置于MRPC之上的闪烁体探测器来标定。

问题：如何判断入射粒子数中，闪烁体有击中而MRPC无击中的情况，以及在探测器探测到的事例数中，MRPC有击中而闪烁体无击中的情况。

例子：系统误差的消除(续一)

多层阻性板(MRPC)探测器效率的测量应改为

$$\text{效率}=\frac{\text{MRPC探测到的事例数}-\text{闪烁体无击中事例数}}{\text{入射粒子数}-\text{不进入MRPC的事例数}}$$

处理方案一

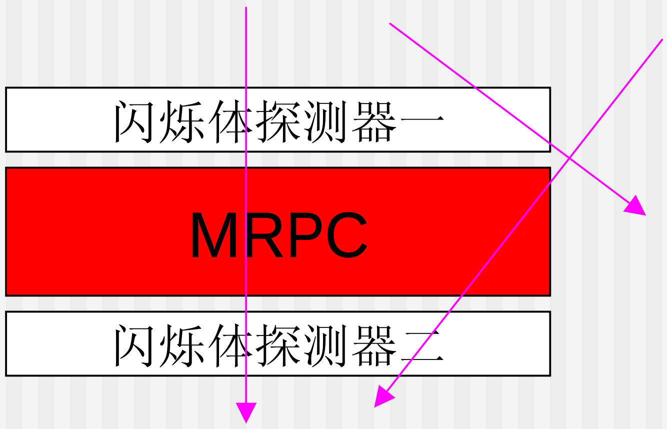
采用蒙特卡罗模拟宇宙线，闪烁体探测器，以及MRPC；找出闪烁体无击中的比率，以及宇宙线只击中闪烁体但并不进入MRPC的比率；然后对效率进行修正。

这种方法适用于实验已结束，不可能再进行改造与运行的情况。由于蒙特卡罗模拟采用的模型会存在不确定性，因此，会给被修正的结果引入新的系统误差。

例子：系统误差的消除(续二)

处理方案二

改进实验方法，在MRPC底部放置另一个闪烁体探测器。定义有粒子入射进入MRPC的条件为上下两个闪烁体在同一时间有击中。这种方法在硬件上消除了在定义入射粒子数与MRPC探测到的粒子数之间不明确的地方。



与处理方案一相比较，该方案不会给结果引入新的系统误差。而且可以最大限度的消除因定义引起的系统误差。

更完善的方法是在MRPC左右前后放置闪烁体探测器，要求在统计事例数时，它们均没有符合的粒子击中信息，从而消除探测器边缘效应对效率估计的影响。

小结

1. 系统误差的定义

   与随机(统计)误差对应，它只体现系统的不确定性

2. 系统不确定性与系统错误

   系统不确定性不可修正，但系统错误必须修正

3. 系统误差的检查与结果并合

   有多种检查方法，不要将通过检查的结果并入总误差

4. 系统误差分析举例

   给出如何分析与估计系统误差，以及减小误差的方法