

3  
0-65

23

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

R - 292

Д. Орир

ЗАМЕТКИ О СТАТИСТИКЕ ДЛЯ ФИЗИКОВ

Дубна 1959 год

B-7-71-43

R-292

Д. Орир

ЗАМЕТКИ О СТАТИСТИКЕ ДЛЯ ФИЗИКОВ

Объединенный институт  
ядерных исследований  
БИБЛИОТЕКА

/перевод Ю.А.Шербакова, редакция С.Н.Соколова/

## О Г Л А В Л Е Н И Е

	Стр.
Предисловие . . . . .	3
I. Прямая вероятность . . . . .	4
2. Обратная вероятность . . . . .	4
3. Соотношения правдоподобия . . . . .	6
4. Принцип максимума правдоподобия . . . . .	8
5. Распределения Гаусса . . . . .	10
6. Магическая формула: Ошибка в результатах, полученных из метода максимума правдоподобия; один параметр . . . . .	12
7. Ошибка в результатах, полученных из метода максимума правдоподобия, М-параметров, коррелированные ошибки . . . . .	16
8. Перенос ошибок ( propagation of errors ), матрица ошибок . . . . .	20
9. Систематические ошибки . . . . .	22
10. Единственность решения в принципе максимума правдоподобия . . . . .	23
II. Доверительные интервалы и их зависимость от выбора параметра. . . . .	25
12. $S$ -функция Бартлета . . . . .	26
13. Биномиальное распределение . . . . .	28
14. Распределение Пуассона . . . . .	30
15. Обобщенный метод максимума правдоподобия . . . . .	32
16. Метод наименьших квадратов . . . . .	36
17. Точность соответствия, $\chi^2$ -распределение . . . . .	41
Приложение: I. Предсказание из соотношений правдоподобия . . . . .	45
2. Распределение суммы наименьших квадратов . . . . .	47

## П р е д и с л о в и е

Эти заметки базируются на цикле лекций, прочитанном в Радиационной лаборатории летом 1958 года. Я хочу четко указать на отсутствие моего знакомства с математической литературой и соответствующего отсутствия поэтому математической строгости в данном изложении. Первоисточником для основного материала и подхода, изложенного здесь, был Энрико Ферми.

Первая разработка большей части излагаемого материала была предпринята в обсуждениях с Э.Ферми, Ф.Сольмицем и Г.Бакусом в Чикагском университете осенью 1953 г. Я благодарен др.Ф.Сольмицу за многие полезные дискуссии, и я много позаимствовал из его доклада "Заметки о методах наименьших квадратов и максимального правдоподобия"<sup>1)</sup>. Другими использованными материалами были работы Анниса, Честона и Примакова<sup>2)</sup>, М.С.Бартлета<sup>3)</sup> и Н.Крамера<sup>4)</sup>. Главная цель изложения будет состоять в изучении распределения Гаусса, биномиального распределения, распределения Пуассона и метода наименьших квадратов в том объеме, как это нужно в использовании метода максимума правдоподобия.

- 
1. F.Solnitz, Notes on the Least Squares and Maximum Likelihood Methods, Institute for Nuclear Studies Report, University of Chicago.
  2. M.Annis, W.Cheston, and H.Primakoff, On Statistical Estimation in Physics, Revs. Modern. Phys. 25, 818 (Oct. 1953)
  3. M.S.Bartlett, On the Statistical Estimation of Mean Lifetimes, Phil. Mag., 44, 249 (1953).
  4. H.Cramer, Mathematical Methods of Statistics (Princeton University Press, 1946).

## 1. Прямая вероятность

По определению понятия вероятности написано много книг. Мы же просто заметим два свойства: (а) статистическая независимость (события должны быть совершенно несвязанными) и (б) закон больших чисел. Это означает, что если  $N_i$  из  $N$  событий принадлежат классу  $I$ , то мы имеем

$$\lim_{N \rightarrow \infty} \left[ \frac{N_i}{N} \right] = P_I.$$

Обычным примером прямой вероятности в физике является такой случай, когда точно известна волновая функция конечного состояния (или плотность вероятности). Одним из таких случаев, будет также случай, когда в некотором эксперименте по рассеянию мы знаем заранее угловое распределение  $f(x)$ , где  $x = \cos \theta$ . В таком эксперименте можно предсказать с определенностью, что число частиц, которое испускается под определенным углом  $x_1$ , в интервал  $\Delta x_1$ , будет  $N f(x_1) \Delta x_1$ , где  $N$  (полное число рассеянных частиц) является очень большим числом. Следует заметить, что функция

$f(x_1)$  нормирована на единицу:

$$\int_{-1}^1 f(x) dx = 1.$$

Как физики, мы называем такие функции функциями распределения. Математики называют их функциями плотности вероятности. Заметьте, что элемент вероятности  $d_p$ , будет

$$d_p = f(x) dx.$$

## 2. Обратная вероятность

Обычная задача, с которой встречается физик, состоит в определении из экспериментальных измерений волновой функции конечного состояния.

Например, рассмотрим распад частицы со спином  $1/2$  —  $\eta$  — мезона, в котором не сохраняется четность. Из закона сохранения углового момента нам было заранее известно, что

$$f(x) = \frac{1 + \alpha x}{2}.$$

Однако численная величина  $\alpha$  является некоторой универсальной физической константой, которую еще следует найти. Мы будем всегда использовать индекс 0 для обозначения истинной физической величины изучаемого параметра. Обычно физик делает эксперимент и получает результат  $\alpha = \alpha^* \pm \Delta\alpha$ . Большая часть данной статьи посвящена вопросам, что мы подразумеваем под  $\alpha^*$  и  $\Delta\alpha$  и каков "наилучший" путь для вычисления  $\alpha^*$  и  $\Delta\alpha$ ? Эти вопросы являются исключительно важными для физиков.

Грубо говоря,  $\Delta\alpha$  представляет из себя стандартное отклонение<sup>5/</sup> и, обычно физики под этим подразумевают, что "вероятность" нахождения  $\alpha_0$  в интервале

$$(\alpha^* - \Delta\alpha) < \alpha_0 < (\alpha^* + \Delta\alpha)$$

составляет 68,3% (площадь под кривой Гаусса в пределах одного стандартного отклонения). Использование термина "вероятность" в предыдущей фразе будет шокировать математика. Он бы сказал, что вероятность получения  $(\alpha^* - \Delta\alpha) < \alpha_0 < (\alpha^* + \Delta\alpha)$  составляет или 0 или 1.

---

<sup>5/</sup> Некоторые физики используют вероятную ошибку чаще чем стандартное отклонение. Кроме того, некоторые физики перед публикацией своих результатов преднамеренно умножают полученные стандартные отклонения на фактор "безопасности" /такой как  $\sqrt{2}$ /. Такая практика запутывает других физиков, которые в процессе своей работы должны комбинировать, сопоставлять, интерпретировать или преобразовывать экспериментальные результаты.

Тот вид вероятности, которую подразумевает здесь физик, называется обратной вероятностью в противоположность прямой вероятности, используемой математиком. Большинство физиков используют то же самое слово "вероятность" для обозначения двух совершенно различных понятий: прямой вероятности и обратной вероятности<sup>х/</sup>. В остальной части этой статьи мы приспособимся к этому небрежному "физическому" использованию слова "вероятность".

### 3. Соотношения правдоподобия

Предположим, нам известно, что должна быть справедлива или гипотеза А или гипотеза В. Известно также, что если гипотеза А является справедливой, то экспериментальное распределение переменной  $x$  должно быть  $f_A(x)$ , и если справедлива гипотеза

В, то распределение должно быть  $f_B(x)$ . Например, если гипотеза А состоит в том, что  $\tau^\pm$  мезон имеет спин нуль, а гипотеза В - что он имеет спин 1, то  $f_A(x)=1$

и  $f_B(x)=2x$ , где  $x$  - кинетическая энергия  $\Pi^-$ -мезона от распада, деленная на свою максимальную величину.

Если справедлива гипотеза А, то суммарная вероятность получения в результате отдельного опыта набора величин  $X_1, X_2, \dots, X_N$  будет

$$d\rho_A = \prod_{i=1}^N f_A(x_i) dx_i.$$

Соотношение правдоподобия  $R$  будет

$$R = \prod_{i=1}^N \frac{f_A(x_i)}{f_B(x_i)} \quad (I)$$

<sup>х/</sup> Так как нас всегда, по существу, интересует распределение разности  $Q_0 - Q^*$ , в которую  $Q_0$  и  $Q^*$  входят вполне равноправно, то скрупулезное различие между обсуждаемой автором "прямой" и "обратной" вероятностью не имеет практического значения. /Примеч.ред./.

Это представляет собой вероятность того, что отдельный экспериментальный результат из  $N$  событий окажется таким, как это должно быть, если справедлива гипотеза  $A$ , деленную на вероятность того, что экспериментальные данные будут такими, какими они должны быть, если бы оказалась справедливой гипотеза  $B$ . Вышеприведенная длинная фраза является точным выражением, использующим понятие прямой вероятности. У физиков имеется более короткий путь, для того чтобы сказать то же самое, используя понятие обратной вероятности. Они говорят, что уравнение (I) дает для  $A$  лучшие шансы, чем для  $B$ .

Формализм обратной вероятности вводит обратные вероятности, отношение которых представляет собой отношение правдоподобия в том случае, если не существует априорных вероятностей в пользу  $A$  или  $B$ . Весь нижеизлагаемый материал в данной работе основывается только на этом главном принципе. Модификации, которые следует ввести в том случае, когда существуют априорные сведения, рассматриваются в разделе IO.

Для физика, планирующего новый эксперимент, очень важна работа, состоящая в предварительном выяснении того, как много событий необходимо для подтверждения гипотезы. Предположим, что для  $\tau$  мезона желательно показать, что имеется  $10^4$  шансов против 1, что спин его не равен 1.

Сколько событий необходимо для этого?

Эта проблема и общий метод, применяемый в этом случае, обсуждается в приложении I "Предсказание из соотношений максимума правдоподобия."

#### 4. Метод максимума правдоподобия

Предыдущий раздел был посвящен случаю, когда имеют дело с дискретным набором гипотез, между которыми необходимо сделать выбор. Чаще же в физике имеют дело с бесконечным набором гипотез; т.е. с параметром, который является непрерывной переменной. Например, в распределении  $\mu$ -e распада

$$f(a; x) = \frac{1+ax}{2}$$

возможные значения для  $a_0$  принадлежат скорее к непрерывному, чем дискретному набору. В этом случае, как и прежде, мы привлечем тот же самый основной принцип, который гласит, что относительная вероятность любых двух различных величин  $a$  представляет собой отношение вероятностей получения отдельных экспериментальных результатов  $X_i$ , если принять сначала одну, а затем другую величину  $a$  за истинную. Эта вероятностная функция от  $a$  называется функцией правдоподобия  $\mathcal{L}(a)$

$$\mathcal{L}(a) = \prod_{i=1}^N f(a, x_i). \quad (2)$$

Функция правдоподобия  $\mathcal{L}(a)$  есть суммарная плотность вероятности получения отдельного экспериментального результата  $X_1, \dots, X_n$ , если принять, что  $f(a, x)$  является истинной нормированной функцией распределения:

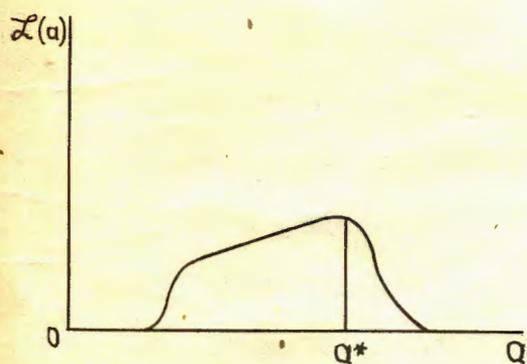
$$\int f(a; x) dx = 1$$

Относительные вероятности для  $a$  могут быть представлены в виде графика дающего зависимость  $\mathcal{L}(a)$  от  $a$ . Наиболее вероятная

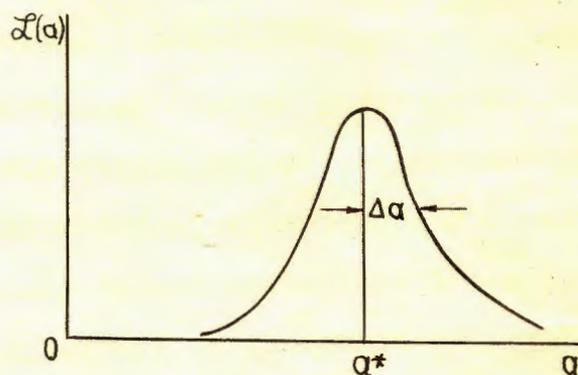
величина  $\alpha$  называется максимально-правдоподобным решением  $\alpha^*$ . Среднеквадратичное распределение  $\alpha$  около  $\alpha^*$  является обычной мерой точности определения  $\alpha = \alpha^*$ . Мы назовем ее  $\Delta \alpha$

$$\Delta \alpha = \sqrt{\left[ \frac{\int (\alpha - \alpha^*)^2 \mathcal{L} d\alpha}{\int \mathcal{L} d\alpha} \right]}. \quad (3)$$

В общем случае функция правдоподобия будет близка к гауссовской (может быть показано, что она стремится к гауссовскому распределению при  $N \rightarrow \infty$ ), и будет выглядеть подобно графику, представленному на фиг.2.



Фиг. 1.



Фиг. 2.

График на фиг.1 представляет собой картину, которую мы имеем в случае малой статистики. В таком случае лучше давать графики для  $\mathcal{L}(\alpha)$ , чем просто приводить  $\alpha^*$  и  $\Delta \alpha$ . Непосредственные методы нахождения  $\Delta \alpha$  излагаются в разделах 6 и 7.

Правильность такого подхода с использованием понятия обратной вероятности подтверждается теоремой максимума правдоподобия, которая была доказана Крамером<sup>4</sup> путем использования понятия прямой вероятности. Теорема утверждает, что в пределе  $N \rightarrow \infty$

$a^* \rightarrow a_0$ , и более того, что не имеется какого-либо другого метода оценки, который бы являлся более точным. В общем случае, когда имеется  $M$ -параметров  $a_1 \dots a_M$ , которые необходимо найти, процедуры получения максимально правдоподобного решения состоят в решении системы из  $M$  уравнений,

$$\boxed{\frac{\partial w}{\partial a_i} \Big|_{a_i = a_i^*} = 0} \quad \text{где } w = \ln \mathcal{L}(a_1, \dots, a_M) \quad (4)$$

### 5. Распределения Гаусса

В качестве первого приложения метода максимального правдоподобия мы рассмотрим пример измерения физического параметра  $a_0$ , где  $x$  является результатом измерения обычного типа, в котором, как известно, имеется ошибка измерения  $\sigma$ . Тогда, если  $x$  распределено по Гауссу<sup>6)</sup>, то функция распределения будет иметь вид

$$f(a_0; x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x-a_0)^2}{2\sigma^2} \right].$$

---

6) Вывод распределения Гаусса и его связь с распределениями Пуассона и биномиальным дается в главе II Физической статистики Р.Б.Линдсея (Wiley, New.York 1941)

Для набора из  $N$  измерений  $x_i$ , каждое со своей собственной ошибкой измерения  $\sigma_i$ , функция правдоподобия может быть записана как

$$\mathcal{L}(\alpha) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp[-(x_i - \alpha)^2 / 2\sigma_i^2].$$

Тогда

$$W = -\frac{1}{2} \sum_{i=1}^N \frac{(x_i - \alpha)^2}{\sigma_i^2} + \text{const}$$

$$\frac{\partial W}{\partial \alpha} = \sum \frac{x_i - \alpha}{\sigma_i^2} \quad (5)$$

$$\sum \frac{x_i}{\sigma_i^2} - \sum \frac{\alpha^*}{\sigma_i^2} = 0$$

$$\alpha^* = \frac{\sum \frac{1}{\sigma_i^2} x_i}{\sum \frac{1}{\sigma_i^2}}, \quad (6)$$

что будет максимально правдоподобным решением.

Заметьте, эти измерения должны быть взвешены (введены с весами) согласно обратной величине из квадрата их соответствующих ошибок. Когда все ошибки измерений одинаковы, то мы имеем

$$\alpha^* = \frac{\sum x_i}{N}.$$

Далее мы рассмотрим точность такого определения.

6. "Магическая формула": ошибка в результатах,  
полученных из метода максимума правдоподобия,  
один параметр

Можно показать, что для больших  $N$ ,  $\mathcal{L}(\alpha)$  стремится к гауссовскому распределению. Принимая такое допущение (фактически вышеприведенный пример всегда является гауссовым по  $\alpha$ ), мы имеем

$$\mathcal{L}(\alpha) \sim \exp[-(h/2)(\alpha - \alpha^*)^2],$$

где  $1/\sqrt{h}$  представляет собой среднеквадратичный разброс  $\alpha$  около  $\alpha^*$

$$W = -\frac{h}{2}(\alpha - \alpha^*)^2 + \text{const}$$

$$\frac{\partial W}{\partial \alpha} = -h(\alpha - \alpha^*)$$

$$\frac{\partial^2 W}{\partial \alpha^2} = -h.$$

Поскольку  $\Delta \alpha$ , как это определено в уравнении (3), равно  $1/\sqrt{h}$ , то мы имеем:

$$\Delta \alpha = \left[ -\frac{\partial^2 W}{\partial \alpha^2} \right]^{-\frac{1}{2}}$$

- магическая формула I (7)

Теперь дадим ошибку вышеприведенного определения. Уравнение (6) может быть легко найдено дифференцированием уравнения (5) по величине  $\alpha$ .

Ответ будет следующим:

$$\Delta \alpha = \left[ \sum \frac{1}{\sigma_i^2} \right]^{-\frac{1}{2}}$$

Эта формула обычно известна как закон сложения ошибок и относится к повторным измерениям той же самой величины, которая распределена по Гауссу с "ошибками"  $\sigma$ . Во многих конкретных задачах невозможно найти аналитически ни  $\alpha^*$  ни  $\Delta\alpha$ . В таких случаях кривая  $\mathcal{L}(\alpha)$  может быть найдена численно, путем испытания различных значений и использования уравнения (2) для получения соответствующих значений  $\mathcal{L}(\alpha)$ . Полная функция получается затем путем использования кривой Френча. Если  $\mathcal{L}(\alpha)$  похожа на кривую Гаусса, то  $\partial^2 W / \partial \alpha^2$  всюду одна и та же. Если это не так, то лучше использовать среднее

$$\overline{\frac{\partial^2 W}{\partial \alpha^2}} = \frac{\int (\partial^2 W / \partial \alpha^2) \mathcal{L}(\alpha) d\alpha}{\int \mathcal{L}(\alpha) d\alpha}.$$

Правдоподобный аргумент за использование приведенного выражения в качестве среднего следует из таких соображений: если спад  $\mathcal{L}(\alpha)$  более пологий, чем спад у гауссовского распределения, то  $\partial^2 W / \partial \alpha^2$  меньше чем  $\partial^2 W / \partial \alpha^2 |_{\alpha^*}$ .

Таким образом, исследование второй производной дает требуемую большую ошибку. Такая техника обсуждается далее в разделе I2.

Заметьте, что магическая формула I зависит от необходимости иметь экспериментальный результат, и только после этого может быть оценена ошибка. Однако, часто очень важно при постановке эксперимента иметь возможность оценивать заранее сколько данных необходимо для того, чтобы получить заданную степень точности.

Мы теперь выведем магическую формулу II, которой можно пользоваться в случае, когда известна  $f(a, x)$ . При таких обстоятельствах мы желаем найти  $\partial^2 W / \partial a^2$ , усредненную по большому числу повторных измерений,

7. Ошибки в результатах, полученных из метода максимума правдоподобия; М-параметров; коррелированные ошибки

Если из одного эксперимента состоящего из  $N$  событий должны быть определены  $M$ -параметров, то формулы ошибок, приведенные в предыдущем разделе, применимы только в тех редких случаях, когда ошибки не являются коррелированными.

Ошибки будут некоррелированными, только если  $(\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) = 0$  во всех случаях когда  $i \neq j$ .

В общем случае мы разлагаем  $w(\alpha)$  вокруг  $(\alpha^*)$  в ряд Тейлора.

$$W(\alpha) = W(\alpha^*) + \sum_{a=1}^M \frac{\partial w}{\partial \alpha_a} \Big|_{\alpha_a^*} \beta_a - \frac{1}{2} \sum_a \sum_b H_{ab} \beta_a \beta_b + \dots,$$

где  $\beta_i \equiv \alpha_i - \alpha_i^*$

и  $H_{ij} \equiv \frac{\partial^2 w}{\partial \alpha_i \partial \alpha_j} \Big|_{\alpha^*}$ . (9)

Второй член разложения исчезает, поскольку  $\frac{\partial w}{\partial \alpha_a} = 0$  являются уравнениями для  $\alpha_a^*$

$$\ln \mathcal{L}(\alpha) = W(\alpha^*) - \frac{1}{2} \sum_a \sum_b H_{ab} \beta_a \beta_b + \dots$$

Пренебрегая членами более высокого порядка мы имеем:

$$\mathcal{L}(\alpha) = C \exp\left(-\frac{1}{2} \sum_a \sum_b H_{ab} \beta_a \beta_b\right)$$

(на  $M$ -мерной гауссовой поверхности). Как и прежде наши формулы ошибок зависят от допущения, что  $\mathcal{L}(\alpha)$  подобна кривой Гаусса

в области  $\alpha_i \approx \alpha_i^*$ . Как упоминалось в разделе 4, если статистика так мала, что такое приближение является плохим, тогда следует лучше приводить просто график для  $\mathcal{L}(\alpha)$ .

Согласно уравнению (9)  $H$  является симметричной матрицей. Пусть  $\underline{U}$  будет унитарной матрицей, которая диагонализует  $H$ :

$$\underline{U} \cdot H \cdot \underline{U}^{-1} = \begin{pmatrix} h_1 & & 0 \\ & h_2 & \\ 0 & & h_m \end{pmatrix} \equiv \underline{h}, \text{ где } \underline{\tilde{U}} = \underline{U}^{-1}. \quad (10)$$

Пусть  $\underline{\beta} = (\beta_1, \beta_2, \dots, \beta_M)$  и  $\underline{\gamma} \equiv \underline{\beta} \cdot \underline{U}^{-1}$ . Тогда

элемент вероятности в  $\beta$ -пространстве будет

$$d^M p = C \exp \left[ -\frac{1}{2} (\underline{\gamma} \cdot \underline{U}) \cdot H (\underline{\gamma} \cdot \underline{U}) \right] d^M \beta.$$

Поскольку  $|\underline{U}| = 1$  является якобианом, связывающим элементы объема  $d^M \beta$  и  $d^M \gamma$ , то мы имеем

$$d^M p = C \exp \left[ -\left(\frac{1}{2}\right) \sum_a h_a \gamma_a^2 \right] d^M \gamma.$$

Теперь всю  $M$ -мерную гауссову поверхность следует представить в виде произведения независимых одномерных кривых Гаусса. Мы имеем

$$\overline{\gamma_a \gamma_b} = \delta_{ab} h_a^{-1},$$

тогда  $\overline{\beta_i \beta_j} = \sum_a \sum_b \overline{\gamma_a \gamma_b} U_{ai} U_{bj} = \sum_a U_{ia}^{-1} h_a^{-1} U_{aj} =$

$$= (\underline{U}^{-1} \cdot \underline{h} \cdot \underline{U})^{-1}_{ij}$$

Согласно уравнению (10),  $\underline{H} = \underline{U}^{-1} \cdot \underline{h} \cdot \underline{U}$   
 что окончательный результат будет

так,

$$\overline{(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)} = (\underline{H}^{-1})_{ij} \text{ где } H_{ij} = \frac{\partial^2 W}{\partial \alpha_i \partial \alpha_j}$$

Магическая формула III

(II)

Усредненная по повторным экспериментам  
 будет:

$$\overline{H_{ij}} = N \int \frac{1}{f} \left( \frac{\partial f}{\partial \alpha_i} \right) \left( \frac{\partial f}{\partial \alpha_j} \right) dx$$

Для вычисления обратной матрицы  $\underline{H}^{-1}$  имеется следующее правило:

$$H_{ij}^{-1} = (-1)^{i+j} \times \frac{i,j \text{ минор от } \underline{H}}{\text{детерминант от } \underline{H}}$$

Пример. Предположим, что пробеги моноэнергетических частиц распределены по Гауссу со средним пробегом  $\alpha_1$ , и коэффициент страглинга равен  $\alpha_2$  (стандартное отклонение). Наблюдалось  $N$  частиц, имеющих пробеги  $X_1, \dots, X_N$ . Найти  $\alpha_1^*, \alpha_2^*$  и их ошибки.

Тогда:  $\mathcal{L}(\alpha_1, \alpha_2) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\alpha_2}} \exp \left[ -\frac{(X_i - \alpha_1)^2}{2\alpha_2^2} \right]$

$$W = -\frac{1}{2} \sum_i \frac{(x_i - a_1)^2}{a_2^2} - N \ln a_2 - N \ln(2\pi),$$

$$\frac{\partial W}{\partial a_1} = \sum_i \frac{(x_i - a_1)}{a_2^2}, \quad \frac{\partial W}{\partial a_2} = \frac{1}{a_2^3} \sum_i (x_i - a_1)^2 - \frac{N}{a_2}$$

Максимально правдоподобное решение получается путем приравнивания двух вышеприведенных уравнений нулю.

$$\left\{ \begin{array}{l} a_1^* = \frac{1}{N} \sum_i x_i \\ a_2^* = \sqrt{\frac{\sum (x_i - a_1^*)^2}{N}} \end{array} \right\} .$$

Читателю известна формула стандартного отклонения, в которой  $N$  заменено на  $(N-1)$ :

$$\bar{\sigma}_2 = \sqrt{\frac{\sum (x_i - a_1^*)^2}{N-1}} .$$

Это имеет место, поскольку в этом случае наиболее вероятная величина,  $a_2^*$ , и средняя величина,  $\bar{\sigma}_2$ , не совпадают.

Средние значения таких величин изучаются в разделе I7. Матрица получается путем вычисления следующих величин при  $a_1^*$  и  $a_2^*$ :

$$\frac{\partial^2 W}{\partial a_1^2} = -\frac{N}{a_2^2}, \quad \frac{\partial^2 W}{\partial a_2^2} = -\frac{3}{a_2^4} \sum (x_i - a_1)^2 + \frac{N}{a_2^2} .$$

$$\frac{\partial^2 W}{\partial \alpha_1 \partial \alpha_2} = -\frac{2}{\alpha_2^2} \sum (x_i - \alpha_1),$$

$$\underline{H} = \begin{pmatrix} \frac{N}{\alpha_2^{*2}} & 0 \\ 0 & \frac{2N}{\alpha_2^{*2}} \end{pmatrix} \quad \text{и} \quad \underline{H}^{-1} = \begin{pmatrix} \frac{\alpha_2^{*2}}{N} & 0 \\ 0 & \frac{\alpha_2^{*2}}{2N} \end{pmatrix}.$$

Согласно уравнению (II), ошибки  $\alpha_1$  и  $\alpha_2$  являются квадратными корнями из диагональных элементов матрицы ошибки  $\underline{H}^{-1}$

$$\Delta \alpha_1 = \frac{\alpha_2^*}{\sqrt{N}} \quad \text{и} \quad \Delta \alpha_2 = \frac{\alpha_2^*}{\sqrt{2N}}$$

(Это иногда называется ошибкой ошибки)

### 8. Перенос ошибок (propagation of errors), матрица ошибки

Рассмотрим случай, в котором некоторая физическая величина  $y$  является функцией  $\alpha$ .  $y = y(\alpha_1, \dots, \alpha_m)$ . Тогда "наилучшая" величина для  $y$  будет  $y^* = y(\alpha_1^*, \dots, \alpha_m^*)$ . С точностью до первого порядка в  $(\alpha_i - \alpha_i^*)$  мы имеем

$$y - y^* = \sum \frac{\partial y}{\partial \alpha_a} (\alpha_a - \alpha_a^*),$$

$$\overline{(y - y^*)^2} = \sum_a \sum_b \frac{\partial y}{\partial \alpha_a} \cdot \frac{\partial y}{\partial \alpha_b} \overline{(\alpha_a - \alpha_a^*)(\alpha_b - \alpha_b^*)},$$

$$(\Delta y)_{\text{ср. кв.}} = \sqrt{\sum_a \sum_b \frac{\partial y}{\partial \alpha_a} \cdot \frac{\partial y}{\partial \alpha_b} H_{ab}^{-1}}$$

(I2)

Хорошо известный частный случай уравнения (12), который справедлив, когда переменные являются совершенно не коррелированными, выглядит следующим образом:

$$(\Delta y)_{\text{ср. кв}} = \sqrt{\sum_a \left(\frac{\partial y}{\partial \alpha_a}\right)^2 (\Delta \alpha_a)^2}.$$

Обычная задача состоит в том, что интересуются  $M$  параметрами  $y_1, \dots, y_M$ , которые являются известными функциями  $\alpha$ . Если матрица ошибок  $\underline{H}^{-1}$  для  $\alpha_i$  известна, то мы имеем:

$$\overline{(y_i - y_i^*)(y_j - y_j^*)} = \sum_a \sum_b \frac{\partial y_i}{\partial \alpha_a} \cdot \frac{\partial y_j}{\partial \alpha_b} H_{ab}^{-1}.$$

(13)

В некоторых подобных случаях  $\frac{\partial y_i}{\partial \alpha_a}$  не может быть получена непосредственно, но  $\frac{\partial \alpha_i}{\partial y_a}$  легко находится. Тогда

$$\frac{\partial y_i}{\partial \alpha_a} = (\underline{J}^{-1})_{ia}, \text{ где } J_{ij} = \frac{\partial \alpha_i}{\partial y_j}$$

### 9. Систематические ошибки

Систематические ошибки являются общей категорией, которая включает в себя такие эффекты как фон, порог селекции, эффективность просмотра, энергетическое разрешение, угловое распределение, изменение эффективности счетчика в зависимости от положения пучка и его энергии, мертвое время и т.д. Неточность в оценке такого систематического эффекта называется "систематической ошибкой". Часто такие систематические ошибки оцениваются путем отдельных экспериментов, предназначенных специально для этой цели. Вообще же, метод максимального правдоподобия может быть использован в такого рода эксперименте для определения систематического эффекта и его ошибки. В этом случае систематический эффект и его ошибка вкладываются в функцию распределения главного эксперимента. В идеальном случае оба эксперимента могут быть рассмотрены как единый эксперимент с добавочным параметром  $\alpha_{n+1}$ , учитывающим систематическую ошибку.

В некоторых случаях систематический эффект не может быть найден из основного эксперимента. Пример, приведенный в разделе 7, может рассматриваться как такой случай. Допустим, что среди пучка моноэнергетических частиц имеется известный фон частиц, равномерно распределенный по пробегам. В этом случае функция распределения будет

$$f(\alpha_1, \alpha_2, \alpha_3, x) = \frac{1}{C} \left\{ \frac{1}{\sqrt{2\pi} \alpha_2} \exp[-(x-\alpha_1)^2/2\alpha_2^2] + \alpha_3 \right\},$$

$$\text{где } C(\alpha_1, \alpha_2, \alpha_3) = \int_{x_{\min}}^{x_{\max}} f \, dx.$$

Решение  $\alpha_3^*$  просто связано с процентным вкладом от фона.

### 10. Единственность максимально-правдоподобного решения

Обычно само по себе задачей является то, какую физическую величину следует выбрать за  $a$ . Например, в экспериментах по определению времени жизни, некоторые экспериментаторы выберут время жизни  $\tau^*$ , в то время как другие выберут  $\lambda^*$ , где  $\lambda = \frac{1}{\tau}$ . Некоторые экспериментаторы предпочитают использовать импульс, другие - энергию и т.д. Рассмотрим случай двух связанных физических параметров  $\lambda$  и  $a$ . Максимальное-правдоподобное решение для  $a$  получено из уравнения  $\partial w / \partial a = 0$ . Максимально-правдоподобное решение для  $\lambda$  получено из  $\partial w / \partial \lambda = 0$ . Но тогда мы имеем:

$$\frac{\partial w}{\partial a} \cdot \frac{\partial a}{\partial \lambda} = 0, \text{ и } \frac{\partial w}{\partial a} = 0.$$

Таким образом, условие для максимально-правдоподобного решения является единственным и не зависит от произвольности в выборе физического параметра.

Результат, полученный для времени жизни  $\tau^*$ , будет связан с решением  $\lambda^*$  соотношением  $\tau^* = \frac{1}{\lambda^*}$ .

Основным недостатком метода максимума правдоподобия является то, что нам неизвестно, как следует поступить тогда, когда имеется априорная вероятность  $Q$ . Если априорная вероятность  $a$  есть  $G(a)$ , а функция правдоподобия, полученная для самого эксперимента, есть  $H(a)$ , то суммарная функция правдоподобия будет:

$$\mathcal{L}(\alpha) = G(\alpha) H(\alpha);$$

$$W = \ln G + \ln H$$

$$\frac{\partial W}{\partial \alpha} = \frac{\partial}{\partial \alpha} \ln G + \frac{\partial}{\partial \alpha} \ln H$$

$$\frac{\partial}{\partial \alpha} \ln H(\alpha^*) = - \frac{\partial}{\partial \alpha} \ln G(\alpha^*),$$

что и дает максимально-правдоподобное решение. При отсутствии каких-либо априорных сведений член в правой части равен нулю. Другими словами стандартная процедура при отсутствии какой-либо априорной информации состоит в использовании априорного распределения, в котором все величины  $\alpha$  являются одинаково равновероятными. Точнее говоря, невозможно знать "истинное"  $G(\alpha)$ , поскольку оно, в свою очередь, должно зависеть от своей собственной априорной вероятности. Однако вышеприведенное уравнение применимо когда  $G(\alpha)$  является комбинированной вероятностной функцией всех предыдущих экспериментов и  $H(\alpha)$  является вероятностной функцией эксперимента, который анализируется.

Имеется целый класс проблем, в которых желательно найти неизвестное распределение по  $\alpha, G(\alpha)$ , а не одиночное значение  $\alpha_0$ .

Например, желают изучить импульсное распределение  $\mu$ -мезонов в космических лучах. Тогда наблюдается

$$\mathcal{L}(G) = \int G(\alpha) H(\alpha; x) d\alpha_1,$$

где  $H(\alpha; x)$  известна из природы эксперимента, а  $G(\alpha)$  является функцией, которую следует найти. Подобного рода проблема рассмотрена в работе 2.

## II. Доверительные интервалы и их зависимость от выбранного параметра

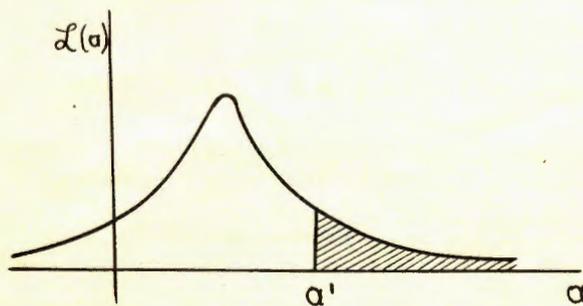
До сих пор мы работали только в терминах относительных вероятностей, и средне-квадратичные величины должны давать, по идее, точность определения  $\alpha = \alpha^*$ .

Какова вероятность того, что  $\alpha$  лежит между двумя определенными величинами такими как  $\alpha'$  и  $\alpha''$ ? Это называется доверительным интервалом

$$P(\alpha' < \alpha < \alpha'') = \int_{\alpha'}^{\alpha''} \mathcal{L} d\alpha / \int_{-\infty}^{\infty} \mathcal{L} d\alpha.$$

К сожалению, такая вероятность зависит от неопределенности в том, какая величина выбрана за  $\alpha$ . Для того, чтобы показать это, рассмотрим площадь, ограниченную кривой  $\mathcal{L}(\alpha)$  и заштрихованную на рисунке.

$$P(\alpha > \alpha') = \frac{\int_{\alpha'}^{\infty} \mathcal{L} d\alpha}{\int_{-\infty}^{\infty} \mathcal{L} d\alpha}$$



Если  $\lambda = \lambda(a)$  была выбрана в качестве физического параметра, то тот же самый доверительный интервал будет

$$P(\lambda > \lambda') = \frac{\int_{\lambda'}^{\infty} \mathcal{L} d\lambda}{\int_{-\infty}^{\infty} \mathcal{L} d\lambda} = \frac{\int_{a'}^{\infty} \mathcal{L} \frac{\partial \lambda}{\partial a} da}{\int_{-\infty}^{\infty} \mathcal{L} d\lambda} \neq P(a > a').$$

Таким образом, в общем случае численная величина доверительного интервала зависит от выбора физического параметра  $x'$ . Это справедливо до некоторой степени также в вычислении  $\Delta a$ . Только максимально-правдоподобное решение и относительные вероятности не зависят от выбора  $a$ . Для гауссовских распределений доверительные интервалы могут быть найдены путем использования таблиц интеграла вероятности. Имеются также таблицы совокупных биномиальных распределений и пуассоновских распределений.

## 12. S - функция Бартлета

М.С.Бартлет (3) рассматривает такой вид доверительного интервала, который не содержит некоторых из вышеуказанных неприятностей. Он вводит функцию  $S(a)$ , которая всегда имеет среднее равное нулю и стандартное отклонение равное единице, независимо от выбора  $a$ .

$$S(a) \equiv \frac{1}{c} \frac{\partial W}{\partial a}, \text{ где } c^2 \equiv - \int_{a_{\min}}^{a_{\max}} \frac{\partial^2 W}{\partial a^2} \mathcal{L}(a) da.$$

x/ § 11 сформулирован Ориром так, что может создаться впечатление, что доверительный интервал не является объективной характеристикой данного экспериментального материала, а зависит от произвола в способе его обработки - от выбора параметра  $a$ . На самом деле, интервал не зависит от выбора параметра  $a$ , т.е. если  $\lambda = \lambda(a)$ ,  $\lambda' = \lambda(a')$ ,  $\lambda'' = \lambda(a'')$ , то всегда

$$P(a' > a > a'') = P(\lambda' > \lambda > \lambda'').$$

Важно только при вычислениях правильно обращаться с функцией  $\mathcal{L}(a)$ . Так как  $\mathcal{L}(a)$  есть плотность вероятности, то, по определению, должно быть  $\mathcal{L}(a) da = \mathcal{L}'(\lambda) d\lambda$ .

Отсюда вытекает следующее правило перехода

$$a \rightarrow \lambda$$

$$\mathcal{L}(a) \rightarrow \mathcal{L}'(\lambda) = \mathcal{L}[a(\lambda)] \frac{\partial a}{\partial \lambda}.$$

Этого правила надо придерживаться при манипуляциях со всеми функциями, имеющими смысл плотности вероятности. Действительно, в § 11 Д.Орир показывает, что формальные вычисления с  $\mathcal{L}$  по правилу "функция от функции" приводят к неверному результату. /Примеч.ред./.

Для  $\mathcal{L}(a)$ , которая представляет собой кривую Гаусса со стандартным отклонением  $\Delta a$ ,  $S(a)$  будет

$$S(a) = -\frac{a - a^*}{\Delta a}.$$

Бартлет предлагает, что поскольку распределение  $S$  больше похоже на распределение Гаусса, чем  $a$ , то 68,3%-ый доверительный интервал (одно стандартное отклонение) в  $a$  может быть получен путем решения для двух величин  $a$ , что дает  $S(a') = +1$  и  $S(a'') = -1$ .

Подобным же образом доверительный интервал для отклонения в пределах двух стандартов получается путем решения для  $S(a) = \pm 2$ . Работа Бартлета<sup>(3)</sup> содержит также дальнейшее уточнение поправок на несимметричность. Теперь мы покажем, что  $\bar{S} = 0$  и  $\bar{S}^2 = 1$ .

$$\bar{S} = \frac{1}{c} \int \frac{dw}{da} \mathcal{L} da = \int \frac{\partial \mathcal{L}}{\partial a} da = [\mathcal{L}(a_{\max}) - \mathcal{L}(a_{\min})] = 0$$

$$\bar{S}^2 = \frac{1}{c^2} \int \frac{1}{\mathcal{L}^2} \left( \frac{\partial \mathcal{L}}{\partial a} \right)^2 \mathcal{L} da = \frac{\int \frac{1}{\mathcal{L}} \left( \frac{\partial \mathcal{L}}{\partial a} \right)^2 da}{-\int \frac{\partial}{\partial a} \left( \frac{1}{\mathcal{L}} \frac{\partial \mathcal{L}}{\partial a} \right) \mathcal{L} da} =$$

$$= \frac{\int \frac{1}{\mathcal{L}} \left( \frac{\partial \mathcal{L}}{\partial a} \right)^2 da}{-\int \frac{\partial^2 \mathcal{L}}{\partial a^2} da + \int \frac{1}{\mathcal{L}} \left( \frac{\partial \mathcal{L}}{\partial a} \right)^2 da},$$

$\bar{S}^2 = 1$ , поскольку член

$$-\int \frac{\partial^2 \mathcal{L}}{\partial a^2} da = \frac{\partial \mathcal{L}}{\partial a} \Big|_{a_{\min}} - \frac{\partial \mathcal{L}}{\partial a} \Big|_{a_{\max}} = 0$$

### 13. Биномиальное распределение

Здесь мы рассмотрим случай, когда событие должно быть отнесено к одному из двух категорий, например, вверх или вниз, вперед или назад, положительно или отрицательно и т.д. Пусть  $p$  есть вероятность для события попасть в класс 1. Тогда  $(1-p)$  есть вероятность попасть в класс 2, и суммарная вероятность для наблюдения  $N_1$  событий в классе 1 из общего числа  $N$  событий будет:

$$P(N_1, N) = \frac{N!}{N_1!(N-N_1)!} \cdot p^{N_1} (1-p)^{N-N_1} \quad \text{биномиальное распределение} \quad (14)$$

Заметим, что  $\sum_{j=0}^N P(j, N) = [p + (1-p)]^N = 1$

Факториалы указывают на тот факт, что нам безразличен порядок, в котором происходят события.

Для данного экспериментального результата состоящего из  $N_1$  событий, относящихся к классу 1, функция правдоподобия  $\mathcal{L}(p)$  будет тогда

$$\mathcal{L}(p) = \frac{N!}{N_1!(N-N_1)!} p^{N_1} (1-p)^{N-N_1} \quad (15)$$

$$W = N_1 \ln p + (N-N_1) \ln(1-p) + \text{const} \quad (16)$$

$$\frac{\partial^2 W}{\partial p^2} = -\frac{N_1}{p^2} - \frac{N-N_1}{(1-p)^2} \quad (17)$$

Из уравнения (15) мы имеем

$$p^* = \frac{N_1}{N}$$

Из (I6) и (I7)

$$\overline{(p-p^*)^2} = \frac{1}{\frac{N_1}{p^{*2}} - \frac{N-N_1}{(1-p^*)^2}}$$

$$\Delta p = \sqrt{\frac{p^*(1-p^*)}{N}}$$

(I8)

Результаты, даваемые уравнениями (I7) и (I8) подобны тем, которые получаются при использовании прямой вероятности<sup>5)</sup>.

Тогда

$$\bar{N}_1 = pN$$

и  $\overline{(N_1 - \bar{N}_1)^2} = Np(1-p)$

---

Пример. В предыдущем примере (см. раздел 6) по угловому распределению  $\mu$ -e -распада мы нашли, что  $\Delta \alpha \approx \sqrt{\frac{3}{N}}$  является ошибкой параметра асимметрии  $\alpha$ . Предположим, что отдельное значение косинуса  $\chi_i$  для каждого события неизвестно.

Все, что мы знаем в этой проблеме, это только число случаев вверх и число вниз. Каково в этом случае будет  $\Delta \alpha$ ? Пусть  $p$  представляет собой вероятность распада в верхнюю полусферу, тогда мы имеем

$$p = \int_0^1 \frac{1+\alpha x}{2} dx = \frac{1+\frac{\alpha}{2}}$$

$dp = \frac{1}{4} d\alpha$  и, в пределе малых ошибок,  $\Delta \alpha = 4 \Delta p$ .

По уравнению (I8)

$$\Delta \alpha = 4 \sqrt{\frac{p^*(1-p^*)}{N}}$$

$$\Delta \alpha = \sqrt{\frac{4}{N} \left(1 - \frac{\alpha^{*2}}{4}\right)}$$

#### 14. Распределение Пуассона

Обычным типом задач, которые попадают в эту категорию, является определение поперечного сечения или определение среднего свободного пробега. Для среднего свободного пробега  $\lambda$  вероятность попадания события в интервал  $dx$  будет  $dx/\lambda$ . Пусть  $P(0, x)$  будет вероятностью непопадания события на отрезок  $x$ .

Тогда мы имеем

$$\begin{aligned}dP(0, x) &= -P(0, x) \times \frac{dx}{\lambda}, \\ \ln P(0, x) &= -\frac{x}{\lambda} + \text{const}, \\ P(0, x) &= e^{-x/\lambda} \quad (\text{при } x=0, P(0, x)=1). \end{aligned} \quad (19)$$

Пусть  $P(N, x)$  будет вероятностью нахождения  $N$  событий на отрезке  $x$ . Тогда элемент этой вероятности будет суммарной вероятностью для  $N$  событий на  $dx_1, \dots, dx_N$ , умноженной на вероятность непопадания событий на остаток прямой.

$$d^N P(N, x) = \prod_{i=1}^N \left( \frac{dx_i}{\lambda} \right) x e^{-\frac{x}{\lambda}}. \quad (20)$$

Полная вероятность получается путем интегрирования по  $N$ -мерному пространству. Следует заметить, что интеграл

$$\prod_{i=1}^N \int_0^x \frac{dx_i}{\lambda} = \left( \frac{x}{\lambda} \right)^N$$

проделывает всю эту операцию, но при этом каждый отдельный элемент вероятности из уравнения (20) он учитывает  $N!$  раз.

Деление на  $N!$  дает

$$P(N, x) = \frac{\left(\frac{x}{\lambda}\right)^N}{N!} e^{-\frac{x}{\lambda}} .$$

распределение Пуассона (21)

В качестве проверки, заметим, что

$$\sum_{j=1}^{\infty} P(j, x) = e^{-\frac{x}{\lambda}} \left( \sum \frac{\left(\frac{x}{\lambda}\right)^j}{j!} \right) = e^{-\frac{x}{\lambda}} \left( e^{\frac{x}{\lambda}} \right) = 1$$

$$\bar{N} = \sum_{N=1}^{\infty} N \frac{\left(\frac{x}{\lambda}\right)^N}{N!} e^{-\frac{x}{\lambda}} = \frac{x}{\lambda} .$$

Подобным же образом может быть показано, что  $(N - \bar{N})^2 = N$

Уравнение (21) часто дается через  $\bar{N}$  :

$$P(N, \bar{N}) = \frac{\bar{N}^N}{N!} e^{-\bar{N}}$$

распределение Пуассона (22)

Такая форма полезна при анализе экспериментов со счетом числа импульсов. В этом случае  $\bar{N}$  будет "истинной" скоростью счета.

Теперь мы рассмотрим случай, когда в некотором эксперименте наблюдалось  $N$  событий. Проблема состоит в нахождении максимально-правдоподобного решения для  $\alpha \equiv \bar{N}$  и его ошибки.

$$\mathcal{L}(\alpha) = \frac{\alpha^N}{N!} e^{-\alpha} ,$$

$$W = N \ln \alpha - \alpha - \ln N!$$

$$\frac{\partial W}{\partial a} = \frac{N}{a} - 1$$

$$\frac{\partial^2 W}{\partial a^2} = -\frac{N}{a^2} .$$

Тогда мы имеем  $a^* = N$  и из уравнения (7)

$$\Delta a = \frac{a}{\sqrt{N}} .$$

В экспериментах по определению поперечного сечения, мы имеем  $a = \rho x b$ , где  $\rho$  - число ядер мишени в  $\text{см}^3$  и  $x$  - полная длина пробега. Тогда

$$b^* = \frac{N}{\rho x} \quad \text{и} \quad \frac{\Delta b}{b^*} = \frac{1}{\sqrt{N}} .$$

В заключение мы заметим, что  $a^* \neq a$ :

$$\bar{a} = \frac{\int_0^{\infty} a \mathcal{L}(a) da}{\int_0^{\infty} \mathcal{L}(a) da} = \frac{\int_0^{\infty} a^{N+1} e^{-a} da}{\int_0^{\infty} a^N e^{-a} da} = \frac{(N+1)!}{N!} = N+1 .$$

### 15. Обобщенный метод максимума правдоподобия

Поскольку мы все время работали со стандартным формализмом максимального правдоподобия, то функции распределения всегда нормировались на единицу. Ферми показал, что требование нормировки, необязательно, поскольку рассматривается основной принцип, а именно: если аккуратно записать вероятность получения экспериментального материала, то соответствующая функция правдоподобия

дает относительные вероятности для нескольких параметров. Единственное накладываемое требование состоит в том, чтобы вероятность получения отдельного результата была записана точно. Мы теперь рассмотрим общий случай, в котором вероятность попадания события в  $dx$  будет  $F(x) dx$ , и

$$\int_{x_{\min}}^{x_{\max}} F dx \equiv \bar{N}(\alpha)$$

представляет собой среднее число событий, которое можно получить, если повторить эксперимент многократно. Согласно уравнению (19), вероятность непопадания события в малый конечный интервал будет

$$\Delta x \text{ есть } \exp\left(-\int_x^{x+\Delta x} F dx\right).$$

Вероятность непопадания в весь интервал  $x_{\min} < x < x_{\max}$  будет произведением таких экспонент, или же

$$\exp\left(-\int_{x_{\min}}^{x_{\max}} F dx\right) = e^{-\bar{N}}.$$

Элемент вероятности для отдельного экспериментального результата из  $N$  событий при  $x = x_1, \dots, x_N$ , тогда будет:

$$d^N p = e^{-\bar{N}} \prod_{i=1}^N F(x_i) dx_i.$$

Таким образом, мы имеем

$$\mathcal{L}(\alpha) = e^{-\bar{N}(\alpha)} \prod_{i=1}^N F(\alpha; x_i)$$

$$W(\alpha) = \sum_{i=1}^N \ln F(\alpha; x_i) - \int_{x_{\min}}^{x_{\max}} F(\alpha; x) dx$$

Решения  $\alpha_i = \alpha_i^*$  даются попарно системой  $M$  уравнений

$$\frac{\partial W}{\partial \alpha_i} = 0$$

Ошибки даются также выражением

$$\overline{(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)} = H_{ij}^{-1},$$

где

$$H_{ij} = - \frac{\partial^2 W}{\partial \alpha_i \partial \alpha_j}.$$

Единственная разница состоит в том, что  $N$  не присутствует в явной форме в формуле

$$- \frac{\partial^2 W}{\partial \alpha_i \partial \alpha_j} = \int \frac{1}{F} \left( \frac{\partial F}{\partial \alpha_i} \right) \left( \frac{\partial F}{\partial \alpha_j} \right) dx.$$

Вывод, подобный тому, которым выводилось равенство (8), показывает, что  $N$  уже учтено в интегрировании по  $F(x)$ .

В частном сообщении Г.Бакус доказал, используя прямую вероятность, что теорема максимума правдоподобия также справедлива для обобщенного метода максимума правдоподобия, и что в пределе больших  $N$  не имеется метода, который бы давал более точные оценки.

При отсутствии обобщенного метода максимального правдоподобия наша процедура состояла бы в нормировке  $F(\alpha; x)$  на единицу путем использования соотношения

$$f(\alpha; x) = \frac{F(\alpha; x)}{\int F dx}$$

Например, рассмотрим образец, содержащий только два радиоактивных вещества со временами жизни  $\alpha_1$  и  $\alpha_2$ . Пусть  $\alpha_3$  и  $\alpha_4$  будут начальными концентрациями веществ.

Тогда мы имеем

$$F(\alpha_i; x) = \alpha_3 e^{-x/\alpha_1} + \alpha_4 e^{-x/\alpha_2},$$

где  $x$  есть время. Стандартный метод тогда будет состоять в использовании следующего выражения

$$f(\alpha; x) = \frac{e^{-x/\alpha_1} + \alpha_5 e^{-x/\alpha_2}}{\alpha_1 + \alpha_5 \alpha_2},$$

которое нормировано на единицу. Следует заметить, что четыре начальных параметра были сведены к трем путем использования

$\alpha_5 \equiv \alpha_4/\alpha_3$ . Далее  $\alpha_3$  и  $\alpha_4$  могут быть найдены путем использования вспомогательно уравнения

$$\int_0^{\infty} F dx = N$$

для общего числа отсчетов.

В этой стандартной процедуре уравнение

$$\bar{N}(\alpha_i) = N$$

должно быть всегда справедливо. Однако, в обобщенном методе максимального правдоподобия эти две величины необязательно равны.

Поэтому обобщенный метод максимума правдоподобия даст другое решение для  $\alpha_i$ , которое должно быть в принципе лучшим.

Другой пример состоит в том, что наилучшая величина для поперечного сечения  $\sigma$  не получается путем использования обычной процедуры принятия  $\rho \sigma L = N$  (число событий по длине пробега  $L$ ). Тот факт, что имеется дополнительная априорная информация, такая как форма углового распределения, делает возможным провести несколько более точное вычисление поперечного сечения. В частном сообщении Ф. Крауфорда показано, что два метода дают точно те же самые результаты в частном случае, когда  $F(\alpha_i; x)$  однородно по  $\alpha_i$ .

### 16. Метод наименьших квадратов

До сих пор мы рассматривали ситуацию, когда экспериментальные результаты представляли собой  $N$  событий, дающих точные величины  $x_1, \dots, x_N$ , где  $x_i$  могли быть, а могли и не быть, все различными. Случай, когда  $x_i$  имеют известные ошибки измерения рассмотрен в работе (I).

С данного момента мы обратим наше внимание к случаю  $p$  измерений (не  $p$  событий) в точках  $x_1, \dots, x_p$ . Экспериментальные результаты будут в виде  $(y_1 \pm \sigma_1), \dots, (y_p \pm \sigma_p)$ . Одним из экспериментов такого типа будет опыт, в котором каждое измерение состоит из  $N_i$  событий.

Тогда  $y_i = N_i$  и  $N_i$  распределены по Пуассону с  $\sigma_i = \sqrt{N_i}$ . В этом случае функция правдоподобия будет

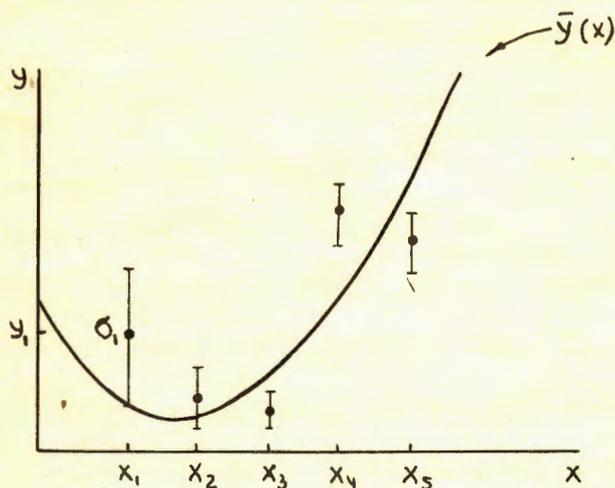
$$\mathcal{L} = \prod_{i=1}^p \frac{[\bar{y}(x_i)]^{N_i}}{N_i!} e^{-\bar{y}(x_i)}$$

и

$$W = \sum_{i=1}^p N_i \ln \bar{y}(x_i) - \sum_{i=1}^p \bar{y}(x_i) + \text{const.}$$

Мы используем обозначение  $\bar{y}(\alpha_i, x)$  для кривой, которая должна быть сопоставлена с экспериментальными точками. Кривая наилучшего соответствия будет определяться  $\alpha_i = \alpha_i^*$ . В этом случае, когда точки распределены по Пуассону, решения получаются из системы  $M$  уравнении

$$\sum_{a=1}^P \frac{\partial \bar{y}(x_a)}{\partial \alpha_i} = \sum_{a=1}^P \frac{N_a}{\bar{y}(x_a)} \frac{\partial \bar{y}(x_a)}{\partial \alpha_i} .$$



Остальная часть этого раздела посвящена случаю, в котором  $y_i$  распределены по Гауссу со стандартным отклонением  $\sigma_i$ . Здесь приложим знаменитый метод наименьших квадратов. Мы теперь увидим, что метод наименьших квадратов математически эквивалентен методу максимума правдоподобия. В этом случае с гауссовым распределением функция правдоподобия будет:

$$\mathcal{L} = \prod_{a=1}^p \frac{1}{\sqrt{2\pi} \sigma_a} \exp \left[ -\frac{(y_a - \bar{y}(x_a))^2}{2\sigma_a^2} \right] \quad (23)$$

$$W = \frac{1}{2} \mathcal{M} - \sum_{a=1}^p \ln \sqrt{2\pi} \sigma_a,$$

где

$$\mathcal{M} = \sum_{a=1}^p \frac{[y_a - \bar{y}(x_a)]^2}{\sigma_a^2} \quad (24)$$

Решения  $\alpha_i = \alpha_i^*$  получаются путем нахождения минимума для  $\mathcal{M}$  (максимума для  $W$ ).

$$\frac{\partial \mathcal{M}}{\partial \alpha_i} = 0 \quad (25)$$

Эта минимальная величина  $\mathcal{M}$  называется суммой наименьших квадратов  $\mathcal{M}^*$ .

Величины  $\alpha_i$ , по которым производится минимализация  $\mathcal{M}$ , называются решениями наименьших квадратов (коэффициентами регрессии). Таким образом, решения метода максимума правдоподобия и наименьших квадратов одинаковы. Согласно уравнению (II) ошибки наименьших квадратов будут:

$$\overline{(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)} = (\underline{H}^{-1})_{ij}, \text{ где } H_{ij} = \frac{1}{2} \frac{\partial^2 \mathcal{M}}{\partial \alpha_i \partial \alpha_j},$$

Наконец, мы рассмотрим частный случай, в котором  $\bar{y}(\alpha_i, x)$  линейна по  $\alpha_i$  :

$$\bar{y}(\alpha_i; x) \equiv \sum_{a=1}^M \alpha_a f_a(x) .$$

Тогда

$$\frac{\partial \eta}{\partial \alpha_i} = -2 \sum_{a=1}^P \left[ \frac{y_a - \sum_{b=1}^M \alpha_b f_b(x_a)}{\sigma_a^2} \right] f_i(x_a) , \quad (26)$$

и

$$H_{ij} = \sum_{a=1}^P \frac{f_i(x_a) f_j(x_a)}{\sigma_a^2} . \quad (27)$$

Определим

$$v_i \equiv \sum_{a=1}^P \frac{y_a f_i(x_a)}{\sigma_a^2} . \quad (28)$$

Тогда

$$\frac{\partial \eta}{\partial \alpha_i} = -2 \left[ v_i - \sum_{b=1}^M \alpha_b H_{bi} \right] .$$

В матричной форме система  $M$  уравнений, дающих решение наименьших квадратов будет:

$$\begin{aligned} 0 &= \underline{u} - \underline{\alpha}^* \cdot \underline{H} && \vee \text{ это } \underline{u} \\ \underline{\alpha}^* &= \underline{u} \cdot \underline{H}^{-1} \end{aligned} \quad (29)$$

$$\alpha_i^* = \sum_{a=1}^M \left[ \sum_{e=1}^P \frac{y_e f_a(x_e)}{\sigma_e^2} \right] (\underline{H}^{-1})_{ai}$$

$u_a$

$$(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) = \underline{H}_{ij}^{-1} \quad \text{где } H_{ij} = \sum_{a=1}^P \frac{f_i(x_a) f_j(x_a)}{\sigma_a^2}$$

(30)

Уравнение (30) является полной процедурой для вычисления решений метода наименьших квадратов и их ошибок. Следует заметить, что хотя эта процедура называется проведением кривой по методу наименьших квадратов, однако, нет никогда никакой необходимости проводить какую-либо кривую. Очень часто эксперимент может представлять собой комбинацию нескольких экспериментов, в котором должны быть совмещены вместе несколько различных кривых (все функции  $\alpha_i$ ). Тогда величина  $\eta$  будет суммой по всем точкам на всех кривых.

Пример. Известно, что кривая должна быть параболой. Имеется четыре экспериментальные точки  $x = -0,6; -0,2; 0,2$  и  $0,6$ . Экспериментальные результаты следующие  $5 \pm 2, 3 \pm 1, 5 \pm 1$  и  $8 \pm 2$ .

Найдем наилучшим образом соответствующую кривую.

$$\bar{y}(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2$$

$$f_1 = 1, f_2 = x, f_3 = x^2$$

$$H_{11} = \sum_a \frac{1}{\sigma_a^2}; \quad H_{22} = \sum_a \frac{x_a^2}{\sigma_a^2}; \quad H_{33} = \sum_a \frac{x_a^4}{\sigma_a^2}.$$

$$H_{12} = \sum_a \frac{x_a}{\sigma_a^2}; \quad H_{13} = \sum_a \frac{x_a^2}{\sigma_a^2} = H_{22}; \quad H_{23} = \sum_a \frac{x_a^3}{\sigma_a^2}.$$

$$\underline{H} = \begin{pmatrix} 2,5 & 0 & 0,26 \\ 0 & 0,26 & 0 \\ 0,26 & 0 & 0,068 \end{pmatrix}$$

$$\underline{H}^{-1} = \begin{pmatrix} 0,664 & 0 & -2,54 \\ 0 & \mathbf{3,847} & 0 \\ -2,54 & 0 & 24,418 \end{pmatrix}$$

$$\underline{u} = (11,25 \quad 0,85 \quad 1,49). \quad (\text{до } \checkmark)$$

$$\alpha_1^* = 3,685$$

$$\Delta \alpha_1 = 0,815$$

$$\overline{\Delta \alpha_1, \Delta \alpha_2} = 0$$

$$\alpha_2^* = 3,27$$

$$\Delta \alpha_2 = 1,96^?$$

$$\alpha_3^* = 7,808$$

$$\Delta \alpha_3 = 4,94$$

$$\overline{\Delta \alpha_1, \Delta \alpha_3} = -2,54$$

$$\bar{y}(x) = (3,685 + 0,815) + (3,27 + 1,96)x + (7,808 + 4,94)x^2$$

будет кривой наилучшего соответствия.

### 17. Точность соответствия, $\chi^2$ -распределение

Численные значения функции правдоподобия при  $\alpha = \alpha^*$ ,  $\mathcal{L}(\alpha^*)$  могут быть, в принципе, использованы в качестве проверки того, правильная ли функция используется для  $f(\alpha; x)$ . Если используется неверная  $f$ , то функция правдоподобия будет меньшей по высоте и большей по ширине. В принципе, используя понятие прямой вероятности, можно вычислить распределение  $\mathcal{L}(\alpha^*)$ , принимая отдельное значение  $f(\alpha_0, x)$  за истинное. Тогда вероятность получения  $\mathcal{L}(\alpha^*)$  меньшей, чем наблюдавшаяся величина, будет указывать, что для  $f$  взята не совсем правильная функция. Если для конкретного эксперимента получили ответ, что имеется один шанс против  $10^4$  на получение такой малой величины  $\mathcal{L}(\alpha^*)$ , тогда может возникнуть серьезный вопрос в отношении эксперимента или функции  $f(\alpha; x)$ , которая была использована.

$$\underline{H} = \begin{pmatrix} 2,5 & 0 & 0,26 \\ 0 & 0,26 & 0 \\ 0,26 & 0 & 0,068 \end{pmatrix}$$

$$\underline{H}^{-1} = \begin{pmatrix} 0,664 & 0 & -2,54 \\ 0 & \mathbf{3,847} & 0 \\ -2,54 & 0 & 24,418 \end{pmatrix}$$

$$\underline{u} = (11,25 \quad 0,85 \quad 1,49). \quad (\text{до } \checkmark)$$

$$\alpha_1^* = 3,685$$

$$\Delta \alpha_1 = 0,815$$

$$\overline{\Delta \alpha_1, \Delta \alpha_2} = 0$$

$$\alpha_2^* = 3,27$$

$$\Delta \alpha_2 = 1,96^{\text{?}}$$

$$\alpha_3^* = 7,808$$

$$\Delta \alpha_3 = 4,94$$

$$\overline{\Delta \alpha_1, \Delta \alpha_3} = -2,54$$

$$\bar{y}(x) = (3,685 + 0,815) + (3,27 + 1,96)x + (7,808 + 4,94)x^2$$

будет кривой наилучшего соответствия.

### 17. Точность соответствия, $\chi^2$ -распределение

Численные значения функции правдоподобия при  $\alpha = \alpha^*$ ,  $\mathcal{L}(\alpha^*)$  могут быть, в принципе, использованы в качестве проверки того, правильная ли функция используется для  $f(\alpha; x)$ . Если используется неверная  $f$ , то функция правдоподобия будет меньшей по высоте и большей по ширине. В принципе, используя понятие прямой вероятности, можно вычислить распределение  $\mathcal{L}(\alpha^*)$ , принимая отдельное значение  $f(\alpha_0, x)$  за истинное. Тогда вероятность получения  $\mathcal{L}(\alpha^*)$  меньшей, чем наблюдавшаяся величина, будет указывать, что для  $f$  взята не совсем правильная функция. Если для конкретного эксперимента получили ответ, что имеется один шанс против  $10^4$  на получение такой малой величины  $\mathcal{L}(\alpha^*)$ , тогда может возникнуть серьезный вопрос в отношении эксперимента или функции  $f(\alpha; x)$ , которая была использована.

На практике, определение распределения  $\chi^2(\alpha^*)$  обычно представляет собой невероятно-трудное численное интегрирование в

*Заметка!*  
N-мерном пространстве. Однако, в частном случае метода наименьших квадратов, пределы интегрирования превращаются в радиус-вектор в p-мерном пространстве. В этом случае мы используем распределение  $\eta(\alpha^*)$ , а не  $\chi^2(\alpha^*)$ .

Мы сначала рассмотрим распределение  $\eta(\alpha_0)$ . Согласно уравнениям (23) и (24) элемент вероятности будет

$$d^p \rho \propto \exp[-\eta/2] d^p y_i.$$

Заметьте, что  $\eta = \rho^2$ , где  $\rho$  - есть модуль вектора в p-мерном пространстве. Объем p-мерной сферы  $V \propto \rho^p$ . Тогда элемент объема в этом пространстве будет:

$$d^p y_i \propto \rho^{p-1} d\rho \propto \eta^{(p-1)/2} \eta^{-1/2} d\eta.$$

Таким образом,

$$dP(\eta) \propto \eta^{p/2-1} e^{-\eta/2} d\eta.$$

Нормировка получается путем интегрирования от  $\eta=0$  до  $\eta=\infty$

$$dP(\eta_a) = \frac{1}{2^{p/2} \Gamma(p/2)} \eta_a^{(p/2)-1} e^{-\eta_a/2} d\eta_a.$$

$$\text{где } \eta_a \equiv \eta(\alpha_0). \quad (29)$$

Это распределение представляет собой хорошо известное  $\chi^2$  распределение с p степенями свободы.

Таблицы  $\chi^2$  для  $\int_{\eta_0}^{\infty} dP(\eta)$  для

некоторого числа степеней свободы приведены в справочнике

"Handbook of Chemistry and Physics", а также в других обычных математических таблицах.

Из определения  $\bar{m}$  (уравнение (24)) очевидно, что  $\bar{m}_0 = p$ . Используя уравнение (29), можно показать, что  $(\bar{m}_0 - \bar{m}_0)^2 = 2p$ . Следовательно, будет подозрительно, если экспериментальные результаты дадут величину  $\bar{m}$ , много большую чем  $(p + \sqrt{2p})$ . Обычно  $\alpha_0$  является неизвестным. В таком случае представляет интерес распределение

$$m^* \equiv m(\alpha^*)$$

К счастью, это распределение также совершенно простое. Это будет чистое  $\chi^2$  распределение с  $(p - m)$  степенями свободы, где  $p$  - число экспериментальных точек, а  $m$  - число искомых параметров. Таким образом, мы имеем

$$dP(m^*) = \chi^2 \text{ - распределение для } (p-m) \text{ степеней свободы.}$$

$$\bar{m}^* = (p-m) \text{ и } \Delta m^* = \sqrt{2(p-m)}$$

(31)

Поскольку вывод уравнения (31) несколько длинен, то мы приводим его в приложении П.

Пример I: Определим  $\chi^2$  вероятность для решения задачи, приведенной в конце раздела I6.

$$m = \left( \frac{5 - \bar{y}(-0.6)}{2} \right)^2 + \left( \frac{3 - \bar{y}(-0.2)}{1} \right)^2 + \left( \frac{5 - \bar{y}(0.2)}{1} \right)^2 + \left( \frac{8 - \bar{y}(0.6)}{2} \right)^2$$

$\chi^2 = 0,674$  по сравнению с  $\bar{\chi}^2 = 4 - 3 = 1$ .

Согласно таблице для  $\chi^2$  для одной степени свободы вероятность получения  $\chi^2 > 0,674$  составляет 0,41. Таким образом, экспериментальные данные вполне совместны с принятой теоретической формой

$$\bar{y} = a_1 + a_2 x + a_3 x^2.$$

Пример 2. Две различные лаборатории измерили время жизни  $\theta_1$ -частицы и дали  $(1,00 \pm 0,01) \cdot 10^{-10}$  сек и  $(1,04 \pm 0,02) \cdot 10^{-10}$  сек, соответственно. Действительно ли эти результаты не согласуются?

Согласно уравнению (6), взвешенное среднее будет

$$\alpha^* = 1,008 \cdot 10^{-10} \text{ сек.}$$

Таким образом,

$$\chi^2 = \left( \frac{1,00 - 1,008}{0,01} \right)^2 + \left( \frac{1,04 - 1,008}{0,02} \right)^2 = 3,2$$

$\bar{\chi}^2 = 2 - 1 = 1$ . Согласно таблице  $\chi^2$  для одной степени свободы вероятность получения  $\chi^2 > 3,2$  будет 0,074.

Поэтому согласно статистике, два измерения одной и той же величины будут принимать вышеуказанные различные значения с вероятностью в 7,4%.

Приложение I

Предсказание из соотношений правдоподобия

Для физика, планирующего новый эксперимент, важная работа состоит в предварительной оценке того, как много событий необходимо для доказательства некоторой гипотезы. Обычная процедура состоит в вычислении среднего логарифма соотношения правдоподобия. Средний логарифм ведет себя математически лучше, чем само среднее от соотношения.

Мы имеем:

$$\overline{\log R} = N \int \log \frac{f_A}{f_B} f_A(x) dx, \quad \begin{array}{l} \text{принимая, что} \\ \text{A верно.} \end{array} \quad (32)$$

или 
$$\overline{\log R} = N \int \log \frac{f_A}{f_B} f_B(x) dx, \quad \begin{array}{l} \text{принимая, что} \\ \text{B верно} \end{array}$$

Рассмотрим пример (данный в разделе 3 ) для  $\tau$  мезона. Мы полагаем, что спин 0 на самом деле имеет место и мы желаем установить, что имеется  $10^4$  шансов против 1, что нет спина 1. Сколько событий необходимо для этого? В этом случае уравнение (32) дает

$$\log 10^4 = 4 = \int_0^1 \log\left(\frac{1}{2x}\right) dx = -N \int_0^1 \log(2x) dx.$$

$$N = 30$$

Таким образом, необходимо в среднем 30 взаимодействий. Однако, если повезет, то может быть и не нужно стольких событий. Рассмотрим крайний случай с одним событием и с  $X = 0$ : тогда  $R$  будет бесконечной и это одно-единственное событие будет само по себе явным доказательством того, что спин  $\tau$ -мезона равен нулю. Флуктуация (средне-квадратичное распределение)  $\log R$  для данного  $N$  будет

$$\overline{(\log R - \overline{\log R})^2} = N \left[ \int \left( \log \frac{f_A}{f_B} \right)^2 f_A dx - \left( \int \log \frac{f_A}{f_B} f_A dx \right)^2 \right]$$

Приложение II

Распределение суммы наименьших квадратов

Мы определим:

$$\underline{z}_i \equiv \frac{y_i}{\sigma_i} \quad \text{и} \quad F_{ij} \equiv \frac{f_i(x_j)}{\sigma_i}.$$

Заметим, что  $\underline{H} = \underline{\tilde{F}} \cdot \underline{F}$  из уравнения (27),

$$\underline{z} \cdot \underline{F} = \underline{a}^* \cdot \underline{H} \quad \text{из уравнений (28) и (29).}$$

Тогда 
$$\underline{a}^* = \underline{z} \cdot \underline{F} \cdot \underline{H}^{-1} \quad (33)$$

$$\mathfrak{M}_0 = \sum_{a=1}^p \sum_{\epsilon=1}^m \left[ (\underline{z}_a - \underline{\tilde{a}}_\epsilon^* \cdot \underline{F}_{a\epsilon}) + (\underline{a}_\epsilon^* - \underline{\tilde{a}}_\epsilon) F_{a\epsilon} \right]^2 \quad (34)$$

где  $\underline{a}$  без звездочки используется для обозначения  $\underline{a}_0$ .

$$\mathfrak{M}_0 = \sum_a^p \sum_\epsilon^m \left( \frac{y_a}{\sigma_a} - \frac{a_\epsilon^* f_\epsilon(x_a)}{\sigma_a} \right)^2 + 2(\underline{z} - \underline{a} \cdot \underline{F}) \underline{F} (\underline{a}^* - \underline{a}) + (\underline{a}^* - \underline{a}) \underline{F} \cdot \underline{F} (\underline{a}^* - \underline{a}),$$

$$\mathfrak{M}_0 = \mathfrak{M}^* + 2(\underline{z} \cdot \underline{F} - \underline{a}^* \cdot \underline{F} \cdot \underline{F}) (\underline{a}^* - \underline{a}) + (\underline{z} \cdot \underline{F} \cdot \underline{H}^{-1} - \underline{a} \underline{H} \cdot \underline{H}^{-1}) \underline{H} (\underline{H}^{-1} \underline{F} \underline{z} - \underline{H}^{-1} \underline{H} \underline{a}),$$

(используется уравнение(34)). Второй член справа равен нулю по уравнения (33)

$$\mathfrak{M}^* = \mathfrak{M}_0 - (\underline{z} \cdot \underline{F} - \underline{a} \underline{F} \underline{F}) \underline{H}^{-1} \underline{H} \underline{H}^{-1} (\underline{\tilde{F}} \underline{z} - \underline{F} - \underline{\tilde{F}} \cdot \underline{\tilde{F}} \underline{a})$$

$$\mathfrak{M}^* = (\underline{z} - \underline{\tilde{z}}) (\underline{1} - \underline{S}) (\underline{z} - \underline{\tilde{z}}), \quad \text{где} \quad \underline{a} \cdot \underline{F} = \underline{\tilde{z}} \quad \text{и}$$

$$\underline{S} = \underline{F} \cdot \underline{H}^{-1} \cdot \underline{\tilde{F}}. \quad (35)$$

Заметим, что

$$\underline{S}^2 = (\underline{\tilde{F}} \cdot \underline{H} \underline{\tilde{F}}) (\underline{\tilde{F}} \cdot \underline{H}^{-1} \underline{\tilde{F}}) = \underline{F} \cdot \underline{H}^{-1} \underline{\tilde{F}} = \underline{S}.$$

Если  $s_i$ , есть собственное значение  $\underline{S}$ , то оно должно быть равно  $s_i^2$ , являющемуся собственным значением для  $\underline{S}^2$ . Таким образом,  $s_i = 0$  или 1. След  $\underline{S}$  будет

$$\text{Tr } \underline{S} = \sum_{a,b,c} F_{ab} H_{bc}^{-1} \tilde{F}_{ca} = \sum_{bc} H_{cb} H_{bc}^{-1} = \text{Tr } \underline{I} = M.$$

Поскольку след матрицы инвариантен при унитарных преобразованиях, то он всегда равен сумме собственных значений матрицы. Поэтому  $M$  собственных значений  $\underline{S}$  равны единице и  $(p-M)$  равно нулю. Пусть  $\underline{U}$  будет унитарной матрицей, которая диагонализует  $\underline{S}$  (и также  $(\underline{1} - \underline{S})$ ). Согласно уравнению (35),

$$\begin{aligned} m^* &= \eta \cdot \underline{U} (\underline{1} - \underline{S}) \underline{U}^{-1} \cdot \underline{\tilde{\eta}}, \\ &= \sum_{a=1}^p m_a \eta_a^2, \\ &= \sum_{a=1}^{p-M} \eta_a^2, \end{aligned}$$

где  $\eta = (\underline{z} - \underline{\bar{z}}) \cdot \underline{U}$ ;  $m$  - есть собственные значения матрицы  $(\underline{1} - \underline{S})$ ;

$m^* = \sum_{a=1}^{p-M} \eta_a^2$  имеет место поскольку  $M$  неравных нулю собственных значений  $\underline{S}$  вычеркивают  $M$  собственных значений  $\underline{1}$ . Таким образом,

$$dp(m^*) \propto e^{-m^*/2} d^{(p-M)} \eta_a,$$

где  $m^*$  - квадрат радиуса вектора в  $(p-M)$ -мерном пространстве. По определению (см. раздел I7) это как раз  $\chi^2$  распределение с  $(p-M)$  степенями свободы.

86-24

Издательский отдел Объединенного института ядерных исследований  
Заказ 734 Тираж 200 Техн. редактор В. Р. Саранцева V 1958 г.