Основные свойства распределения “гиперболический Гаусс”

Зайцев Александр Сергеевич
гр. **7351 ФФ НГУ**
лаб. **2 ИЯФ СО РАН**

8 окт. 2001

Введение

В данной работе приведены основные свойства и некоторые параметризации плотности статистического распределения “гиперболический Гаусс”.

“Экспериментальное происхождение” *распределения.*

Экспериментальные распределения событий по величинам импульса нейтрального каона, среднего импульса пионов, массы нейтрального каона для процесса $K_S^0 \to \pi^+\pi^-$ имеют, как показано в работах [1, 2], существенно негауссовый вид, что хорошо видно по форме распределений, приведенных на Рис. 1. Характерными особенностями всех упомянутых распределений являются экспоненциальный закон падения плотности распределения на больших расстояниях от пика и гауссово поведение плотности вблизи вершины. Распределения подобного рода уже встречались ранее [3] при построении методов определения импульсного разрешения дрейфовой камеры детектора **КМД-2** и абсолютной калибровки энергии пучков ускорительно-накопительного комплекса **ВЭПП-2М** с детектором **КМД-2** [4]. Ввиду такой широкой распространенности представляют интерес построение удобных параметризаций и исследование различных свойств подобных распределений, чему и посвящена данная работа.

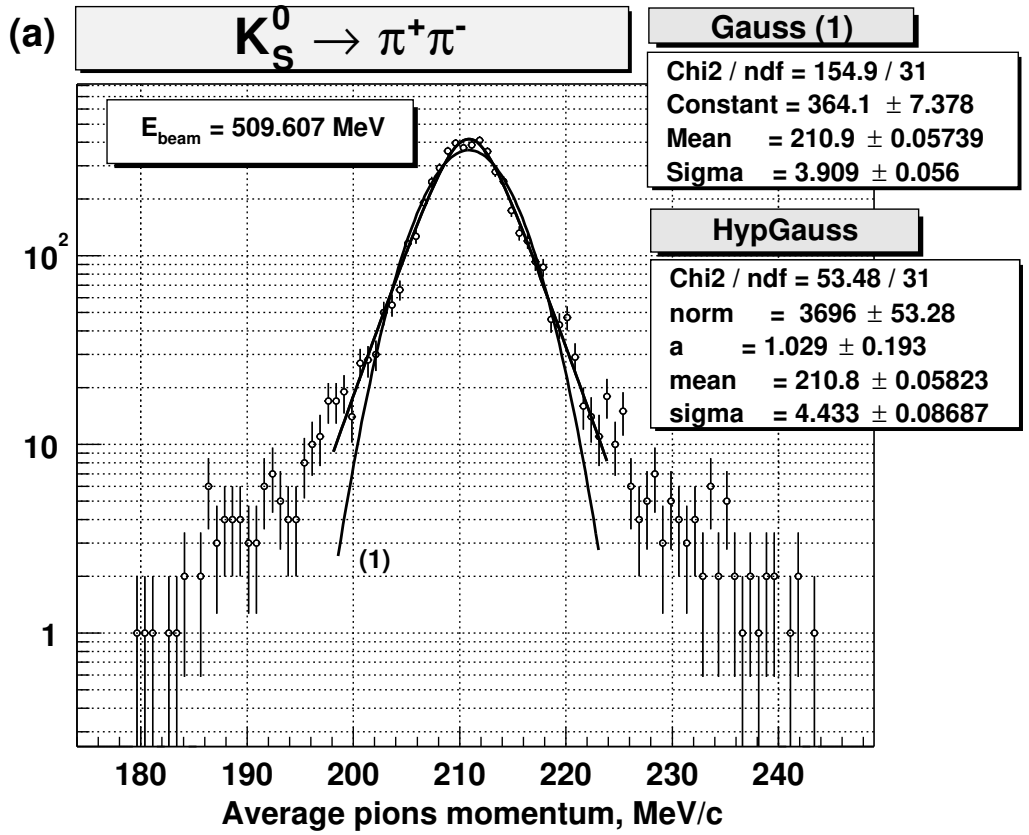
Общие соображения о выборе параметризации плотности *распределения.*
Плотность симметричного распределения, обладающего гауссовой вершиной и экспоненциальными хвостами, может быть записана в виде:

$$F\left(x,\widehat{\beta},\mu,\alpha_1,\alpha_2\right)=\frac{1}{N_x}\exp\left\{-\left[\left(1-\widehat{\beta}\left(x-\mu\right)\right)\frac{|x-\mu|}{\alpha_1}+\widehat{\beta}\left(x-\mu\right)\left(\frac{x-\mu}{\alpha_2}\right)^2\right]\right\}, \qquad (1)$$

где N_x — нормировочный множитель, $\alpha_1>0$, $\alpha_2>0$ — параметры распределения, μ — первый момент распределения, $\widehat{\beta}\left(q\right)$ — некоторая четная непрерывная при $-\infty<q<+\infty$ и монотонная при $0\le q<+\infty$ *переходная функция*, такая что

$$\widehat{\beta}\left(0\right)=1\,,\quad \lim_{q\to+\infty}\widehat{\beta}\left(q\right)=0\,. \qquad (2)$$

Для применения на практике параметризации (1) часто необходимо знать аналитические выражения для величины N_x и моментов распределения второго и более высокого порядков, так что представляется целесообразным рассматривать функции $\widehat{\beta}\left(q\right)$,



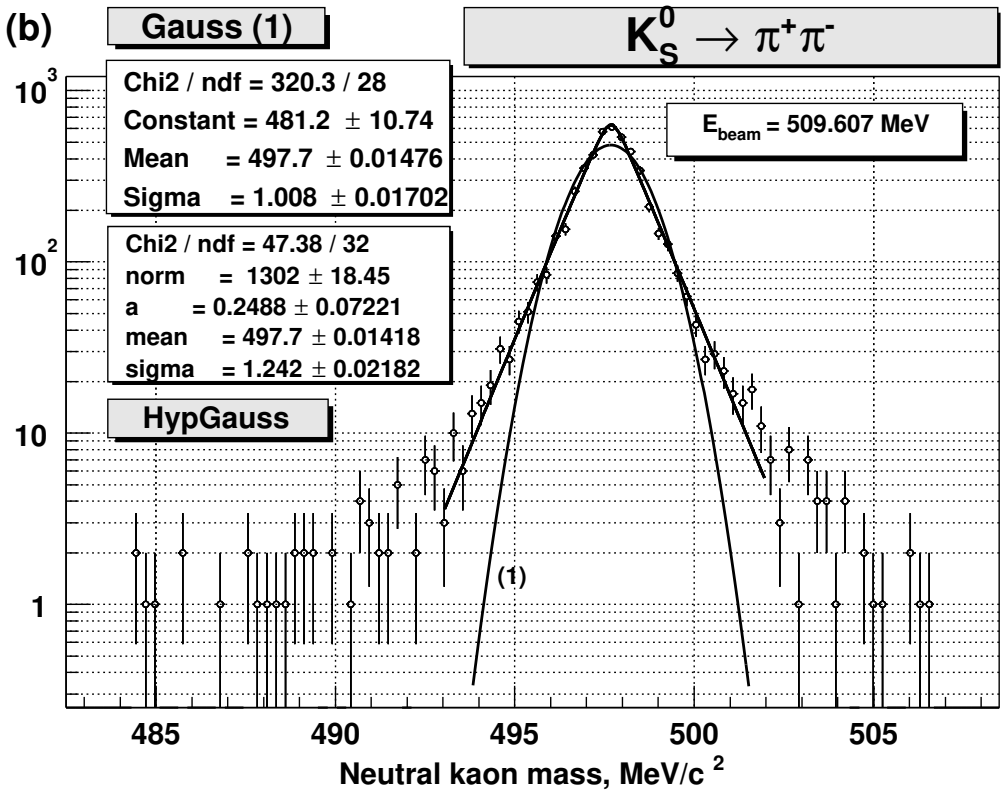


Рис. 1: Примеры аппроксимации некоторых распределений в рамках работы по определению массы нейтрального каона в распаде $K^0_S \to \pi^+\pi^-$ [2]: (a) — распределение по среднему импульсу пионов, (b) — распределения значения массы нейтрального каона в одной из точек по энергии пучков. Приведены аппроксимации с гауссовой ("Gauss") и *гипергауссовой* ("HypGauss") плотностями.

максимально упрощающие этот анализ. Однако, как показывает опыт, удобно вовсе отказаться от параметризации (1) и рассматривать немного менее общий случай:

$$F\left(x,\mu,\alpha_1,\alpha_2\right)=\frac{1}{N_x}\exp\left\{\alpha_1\ \left[1-\sqrt{1+\left(\frac{x-\mu}{\alpha_2}\right)^2}\right]\right\}\ . \tag{3}$$

Анализ плотности вида (3) оказывается менее сложной задачей, чем рассмотрение плотности общего вида (1), однако мы не исключаем возможности найти *переходную функцию* $\widehat{\beta}\left(q\right)$, приводящую к более простому для анализа выражению.

Все последующее изложение относится к параметризации вида (3) и ее дальнейшим модификациям. Распределение, плотность которого имеет вид (3), далее будем называть *гиперболическим Гауссом*, либо *гипергауссом*, либо просто *распределением*.

Плотность *распределения*. Простейшая параметризация плотности *распределения* имеет вид:

$$g\left(x,\mu,a,b\right)=\frac{1}{N_x}\exp\left\{a\ \left[1-\sqrt{1+\left(\frac{x-\mu}{b}\right)^2}\right]\right\}\ , \tag{4}$$

где $a\ ,\ b>0$, N_x — нормировочный множитель

$$N_x=b\,e^a\int_{-\infty}^{+\infty}\cosh x\,e^{-a\cosh x}dx=2b\,e^a K\left(1,a\right)\ , \tag{5}$$

$K\left(\alpha,a\right)$ — семейство функций Бесселя мнимого аргумента 2-го рода (функции Макдональда). Функции $y=K\left(\alpha,x\right)$ удовлетворяют уравнению

$$x^2y''+x\,y'-\left(x^2+\alpha^2\right)y=0 \tag{6}$$

и обладают следующими асимптотиками:

$$\sqrt{\tfrac{\pi}{2x}}e^{-x}+O\left(x^{-3/2}\right)\ ,\quad x\to+\infty\,, \tag{7}$$

$$\tfrac{\Gamma(\alpha)}{2}\left(\tfrac{x}{2}\right)^{-\alpha}+O\left(x^{2-\alpha}\right)\ ,\quad x\to+0\,,\quad \alpha\neq 0\,, \tag{8}$$

$$\ln\tfrac{2}{x}-\gamma+O\left(x^2\right)\ ,\quad x\to+0\,,\quad \alpha=0\,, \tag{9}$$

где γ — константа Эйлера:

$$\gamma=\lim_{n\to\infty}\left(\sum_{i=1}^{n}\frac{1}{i}-\ln n\right)=0.5572156649\ldots\,.$$

Полезное интегральное представление функции Макдональда имееет вид:

$$K\left(\alpha,a\right)=\int_0^{+\infty}\cosh\left(\alpha x\right)e^{-a\cosh x}dx\,. \tag{10}$$

Статистические характеристики плотности. Математическое ожидание E_x и дисперсия D_x случайной величины X, распределенной с плотностью вероятности (4), имеют вид:

$$E_x\ =\ \int_{-\infty}^{+\infty}x\,g\left(x,\mu,a,b\right)dx\,, \tag{11}$$

$$\mu_x^{[2]}\equiv D_x\ =\ \int_{-\infty}^{+\infty}\left(x-\mu\right)^2 g\left(x,\mu,a,b\right)dx=b^2\frac{K\left(2,a\right)}{a\,K\left(1,a\right)}\,. \tag{12}$$

Высшие центральные моменты $\mu_x^{[k]}$ могут быть найдены при помощи следующих соотношений:

$$\mu_x^{[2n+2]} = \left\{ b^2 \left(\frac{\partial^2}{\partial a^2} - 1 \right) \right\} \mu_x^{[2n]} , \quad \mu_x^{[2n+1]} = 0 , \quad n \in \aleph , n > 0 ; \tag{13}$$

$$\frac{\partial}{\partial x} K(\alpha, x) = -K(\alpha+1, x) + \frac{a}{x} K(\alpha, x) , \tag{14}$$

$$\frac{\partial^2}{\partial x^2} K(\alpha, x) = \frac{1}{x} K(\alpha+1, x) + \left[\frac{a(a-1)}{x^2} - 1 \right] K(\alpha, x) , \tag{15}$$

например

$$\mu_x^{[4]} = \left(\frac{b}{a} \right)^4 \left\{ \frac{K(0,a)}{K(1,a)} \left[2\frac{K(0,a)}{K(1,a)} - 3 \right] a^3 + \left[\left(\frac{K(0,a)}{K(1,a)} \right)^2 - 1 \right] a^2 + 12 \right\} . \tag{16}$$

Применение критерия *минимальной граничной вариации.* Для *распределения* возможно аналитическое построение величины $\sigma_0^{(\eta)}$:

$$\left\{ \sigma_0^{(\eta)} \left[f(N, \eta, \{\alpha_i\}) \right] \right\}^{-2} = N \int_{-\infty}^{+\infty} f(x, \mu, \{\alpha_i\}) \left[\frac{\partial \ln f(x, \eta, \{\alpha_i\})}{\partial \eta} \right]^2 dx , \tag{17}$$

знание которой необходимо для применения критерия *минимальной граничной вариации* (**MVB**) [**5, 6**] наличия систематического эффекта при наличии двух оценок $\widetilde{\eta}_1$ и $\widetilde{\eta}_2$ параметра η распределения $f(x, \eta, \{\alpha_i\})$ случайной величины x. Величина $\sigma_0^{(\eta)}$ дает максимально эффективную оценку статистической ошибки величины $\Delta = \widetilde{\eta}_1 - \widetilde{\eta}_2$ [**6**].

Для плотности вида (4) и выражение для $\sigma_0^{(\eta)}$ имеет вид:

$$\left\{ \sigma_0^{(\mu)} \left[g(N, \mu, a, b) \right] \right\}^{-2} = \frac{aN\mu}{2b^2 K(1,a)} \int_{-\infty}^{+\infty} e^{-a \cosh x} dx ,$$

$$\left\{ \sigma_0^{(\mu)} \left[g(N, \mu, a, b) \right] \right\}^{2} = \frac{b^2}{aN\mu} \frac{K(1,a)}{K(0,a)} . \tag{18}$$

Альтернативная параметризация плотности *распределения.* Для нахождения значения дисперсии по результатам аппроксимации $\left(\widetilde{a} \pm \delta a, \widetilde{b} \pm \delta b \right)$ в параметризации вида (4) необходимо осуществить дополнительное вычисление; погрешность результата можно оценить по следующей формуле:

$$\sigma_{\sigma_x}^2 = \frac{K(2,\widetilde{a})}{\widetilde{a} K(1,\widetilde{a})} \left\{ \delta_b^2 + \left(\frac{\widetilde{b}\,\delta_a}{2} \right)^2 \left[\frac{K(2,\widetilde{a})}{K(1,\widetilde{a})} - \frac{K(3,\widetilde{a})}{K(2,\widetilde{a})} \right]^2 \right\} . \tag{19}$$

На практике удобнее параметризовать плотность распределения в терминах σ и свободного параметра a путем замены

$$b \to \sigma \sqrt{\frac{a K(1,a)}{K(2,a)}} \tag{20}$$

следующим образом:

$$G(x, \mu, a, \sigma) = \frac{1}{2\sigma} \sqrt{\frac{K(2,a)}{a K^3(1,a)}} \exp \left[-a \sqrt{1 + \left(\frac{x-\mu}{\sigma} \right)^2 \frac{K(2,a)}{a K(1,a)}} \right] . \tag{21}$$

Оценка параметров *распределения* по выборке. *Метод моментов* позволяет легко получить аналитические выражения[1] для оценок $\widetilde{\mu}\,, \widetilde{\sigma}$ параметров μ и σ соответственно по экспериментальной выборке $\{x_i\}$ объема N:

$$\widetilde{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i\,, \quad \widetilde{\sigma} = \sqrt{\frac{\sum_{i=1}^{i=N}\left(x_i - \widetilde{\mu}\right)^2}{N-1}}\,. \tag{22}$$

Метод максимального правдоподобия [7, 8] в нашем случае не столь удобен для нахождения оценок, так как система уравнений правдоподобия может быть решена лишь численно, к тому же в рамках этого метода невозможно обойти построение оценки параметра a. Как показывает практика, при аппроксимации разумным начальным значением параметра a является любое число из интервала $0 < a < 2$.

Литература

[1] Л.М. Барков, И.Б. Вассерман, П.В. Воробьев *и др.*, Прецизионное измерение массы нейтрального каона. ЯФ **46**, 1088—1094, 1987.

[2] А.С. Зайцев, Квалификационная работа на соискание степени бакалавра "Прецизионное измерение массы нейтрального каона". НГУ, Новосибирск, 2001.

[3] К.Ю. Михайлов (Лаб. **2**, ИЯФ), частное сообщение.

[4] R.R. Akhmetshin, G.A. Aksenov, E.V. Anashkin *et al.*, Measurement of ϕ-meson parameters with **CMD-2** detector at **VEPP-2M** collider. Preprint INP **95-35**, Novosibirsk, 1995.

[5] M. Kendall, A. Stuart and T.K. Ord, Advanced Theory of Statistics (1991), 623. New York: Oxford Univ. Press.

[6] Rodger Barlow, **SLUO** Lectures on Statistics and Numerical Methods in HEP, Lecture **5**: Systematic Errors. **PS files (all lectures):** `http://www.slac.stanford.edu/grp/rd/sluo/Stat_Lectures.html`

[7] В.Е. Гмурман, Теория вероятностей и математическая статистика. М.:"Высшая школа", 2001.

[8] Timothy E. Holy, The Analysis of Data from Continuous Probability Distributions. **physics/9706015**.

[1] Задача существенно упрощена тем обстоятельством, что нет необходимости строить оценку параметра a, если мы не интересуемся высшими моментами распределения.