ОБЪЕДИНЕННЫЙ ИНСТИТУТ ЯДЕРНЫХ ИССЛЕДОВАНИЙ

10-92-475

ЗРЕЛОВ
Петр Валентинович

УДК 519.234.3:539.1

ИССЛЕДОВАНИЕ ИНТЕГРАЛЬНЫХ НЕПАРАМЕТРИЧЕСКИХ КРИТЕРИЕВ СОГЛАСИЯ И ИХ ПРИМЕНЕНИЯ ДЛЯ ИДЕНТИФИКАЦИИ ЧАСТИЦ В ФИЗИКЕ ВЫСОКИХ ЭНЕРГИЙ

Специальность: 05.13.16 - применение вычислительной техники, математического моделирования и математических методов в научных исследованиях

Автореферат диссертации на соискание ученой степени кандидата физико-математических наук

Дубна 1992

Работа выполнена в Лаборатории вычислительной техники и автоматизации Объединенного института ядерных исследований.

Научный руководитель :

кандидат физико - математических наук
старший научный сотрудник ИВАНОВ
Виктор Владимирович

Официальные оппоненты :

доктор физико - математических наук
профессор ТЯПКИН
Алексей Алексеевич

кандидат физико - математических наук
старший научный сотрудник УФИМЦЕВ
Михаил Валентинович

Ведущее научно - исследовательское учреждение : Институт физики высоких энергий, г.Протвино.

Защита диссертации состоится "____"_____ 1993 г.
в _____ часов на заседании специализированного совета Д047.01.04
при Лаборатории вычислительной техники и автоматизации ОИЯИ,
г.Дубна, Московской области, конференц - зал ЛВТА.

С диссертацией можно ознакомится в библиотеке ОИЯИ.
Автореферат разослан "____"_____ 1992 г.

Ученый секретарь
специализированного совета
кандидат физико - математических наук З.М.Иванченко

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность исследований. Одной из специфических черт экспериментальной физики высоких энергий является последовательная регистрация отдельных событий, каждое из которых имеет свою внутреннюю структуру. С другой стороны, стандартный набор статистических методов в значительной степени ориентирован на анализ представляемых в гистограммированном виде экспериментальных распределений, вызванных регистрацией большого количества событий. При анализе же отдельных событий, обычно представляемых малыми выборками, используются не всегда эффективные полуэмпирические методы. Поэтому очень важным направлением является внедрение в эту область строгих методов математической статистики. Наиболее подходящими здесь выглядят методы, разработанные для анализа малых выборок и предназначенные для обработки негистограммированных данных . С этой точки зрения большие перспективы открываются при использовании непараметрических критериев согласия, построенных на выборочной функции распределения. Эти критериии основаны на сравнении функции распределения F_0, предсказываемой гипотезой H_0, с эмпирической функцией распределения F_n при помощи различных мер "расстояния" между этими функциями. Свойство непараметричности для этих критериев обеспечивается выбором таких мер отклонения F_0 от F_n, распределения которых от F_0 не зависят. Такие критерии универсальны и в большинстве случаев просты и удобны в использовании.

Высокая мощность критериев этого типа, а также возможность анализа непосредственно наблюдаемых в эксперименте событий, обусловили ряд попыток применения одного из таких критериев, критерия омега-квадрат, для некоторых задач идентификации частиц в физике высоких энергий.

В настоящей работе развиваются идеи последовательного использования непараметрических критериев интегрального типа для построения статистических методов идентификации частиц на основании результатов измерений их времен пролета или ионизационных потерь энергии одновременно несколькими детекторами физической установки, а также использования построенных методов для решения ряда задач конкретных физических экспериментов.

Цель работы. Построение и исследование новых интегральных непараметрических критериев согласия и сравнение этих критериев по

мощности для различных типов альтернативных гипотез. Разработка на основе этих критериев статистических методов идентификации частиц, регистрируемых физическими установками в экспериментах физики высоких энергий.

Научная новизна работы. Разработан метод исследования интегральных непараметрических статистик, включающий изучение моментных и корреляционных свойств этих статистик, а также численный метод вычисления их функций распределения для малых объемов эмпирической выборки. Рассмотрен класс интегральных статистик $\omega(n,k)$, представимых в виде интеграла от k - й степени эмпирического процесса. С высокой точностью вычислены функции распределения ряда статистик этого класса (при $n = 1(1)10$ и $k = 1(1)5$). Исследованы корреляционные связи некоторых статистик $\omega(n,k)$ с другими известными статистиками интегрального типа. Построены критерии согласия $\omega(n,k)$ и в общем виде изучены их некоторые свойства.

Обоснована целесообразность и возможность использования критериев согласия, основанных на выборочной функции распределения, в задачах статистической идентификации вторичных частиц, регистрируемых экспериментальными установками. На конкретных примерах критериев ω_n^2 и ω_n^3 построены методы, позволяющие практически однозначно идентифицировать частицы по их временам пролета или ионизационным потерям энергии, измеряемым одновременно несколькими детекторами установки. Показано преимущество новых методов перед традиционными методами, использовавшимися для решения таких задач.

Практическая ценность. Построенные методы статистической идентификации частиц были применены для анализа данных, накопленных в результате моделирования проектируемой установки с предварительным названием "O^0Facility", предназначенной для проведения эксперимента по изучению подпорогового рождения K^+- мезонов в столкновениях протонов с легкими ядрами (^2H, ^{12}C) на ускорителе COSY(Julich, Germany). В сочетании с другими методами это помогло обосновать принципиальную возможность выделения изучаемых частиц на подавляющем фоне π^+-мезонов ($N_{K^+}/N_{\pi^+} \approx 10^{-5}$).

Кроме того методы были опробованы на некоторых данных физических экспериментов, проводимых на синхрофазотроне ОИЯИ с помощью установки МАСПИК (магнитный спектрометр с проволочными камерами). Показана высокая эффективность разделения вторичных частиц с разными электрическими зарядами и частиц с разными массами в задачах подобного типа.

Аппробация работы и публикации. Результаты исследований, составившие диссертацию, обсуждались на научных семинарах ЛВТА и ЛЯП ОИЯИ. Часть результатов была доложена автором на "Одинадцатой Пражской конференции по теории информации, статистическим решающим функциям и случайным процессам" (Прага, 1990 г.) и на "Совещании по аналитическим вычислениям на ЭВМ в физических исследованиях" (Дубна, 1991).

Основные результаты диссертации изложены в 12 печатных работах, опубликованных в виде сообщений ОИЯИ, журнальных статей, а также в трудах конференций. Список работ приведен в конце автореферата.

Диссертация состоит из введения, двух частей, каждая из которых содержит две главы, а также заключения и приложения.

СОДЕРЖАНИЕ РАБОТЫ

Во введении дано обоснование актуальности темы диссертации, определены предмет исследования и цель работы. Кратко представлены типы непараметрических критериев согласия, основанных на статистиках, представляемых некоторыми функционалами эмпирического процесса. Определен класс задач для их эффективного использования. Обоснован выбор интегральных критериев, показана возможность расширения и обобщения этого класса.

Представлены примеры использования одного из критериев интегрального типа, критерия омега-квадрат, для решения ряда задач анализа экспериментальных данных в физике высоких энергий.

Дается краткое изложение содержания диссертации по главам.

Часть I

В первой главе вычисляются моментные функции различных порядков эмпирического процесса

$$y_n(t) = \sqrt{n}\ (F_n(t) - t), \qquad 0 \leqslant t \leqslant 1, \tag{1}$$

где $F_n(t)$ -- эмпирическая функция распределения по выборке $t_1, t_2, \ldots, t_n$ объема n из равномерного на $[0,1]$ распределения.

Приводится вид моментной функции эмпирического процесса произвольного порядка $p = k_1 + k_2 + \ldots + k_q$.

Для дальнейшего оказывается удобным ввести понятие обобщенной моментной функции процесса $y_n(t)$:

$$J_q(t^{(I)},t^{(II)},\ldots,t^{(q)};n,\varphi_1,\varphi_2,\ldots,\varphi_q) =$$

$$= E\left\{\varphi_1[t^I,y_n(t^I)]\varphi_2[t^{II},y_n(t^{II})],\ldots\varphi_q[t^{(q)},y_n(t^{(q)})]\right\}, \qquad (2)$$

где $\varphi_1, \varphi_2,\ldots, \varphi_q$ – некоторые функции (см. ниже). При дополнительных условиях интегрируемости функций φ_j по всем $t_i, i = 1, 2,\ldots, n$ оказывается возможным записать выражение для J_q:

$$J_q(t^I,t^{II},\ldots,t^{(q)};n,\varphi_1,\varphi_2,\ldots,\varphi_q) \underline{\underline{\quad t^I<t^{II}<\ldots<t^{(q)} \quad}}$$

$$= \sum_{s_1=0}^{n}\sum_{s_2=0}^{n-s_1}\ldots\sum_{s_q=0}^{n-(s_1+s_2+\ldots s_{q-1})} C_n^{s_1}C_{n-s_1}^{s_2}\ldots C_{n-(s_1+s_2+\ldots s_{q-1})}^{s_q} \times$$

$$\times (t^I)^{s_1}(t^{II}-t^I)^{s_2}\ldots(1-t^{(q)})^{n-(s_1+s_2+\ldots+s_q)} \times$$

$$\times \varphi_1[t^I,n^{-1/2}(s_1-nt^I)]\ldots\varphi_q[t^{(q)},n^{-1/2}(\sum_{i=1}^{n-1}s_i-nt^{(q)})], \qquad (3)$$

где C_n^s – биномиальные коэффициенты.

Для интегральных статистик определенного класса оказывается возможным построение единого метода исследования их моментных и корреляционных свойств. Указанный класс определяется статистиками, имеющими структуру следующего функционала эмпирического процесса $y_n(t)$:

$$\omega_n(\varphi) = \int_0^1 \varphi\,[t,\ y_n(t)]dt, \qquad (4)$$

где $\varphi\,[t,x]$ – обозначает измеримую относительно $\mathfrak{T}\times\mathfrak{X}$ и ограниченную на множестве $T\times X$ функцию ($\mathfrak{T}$ – σ-алгебра борелевских множеств на отрезке $T = [0,1]$, $\mathfrak{X}$ – σ-алгебра борелевских множеств в фазовом пространстве X процесса $y_n(t)$).

Подчеркивается, что поскольку наибольшую практическую ценность имеют статистики, имеющие невырожденное асимптотическое распределение, то на функцию $\varphi(t,y_n(t))$ может быть наложено дополнительное условие, обеспечивающее существование такого распределения. Этим условием является непрерывность φ, определенной на $D_{[0,1]}$[1] и рассматриваемой как функция $y_n(t)$, в метрике Скорохода d.[2]

[1] $D_{[0,1]}$ – пространство вещественных функций на $[0,1]$ непрерывных справа и имеющих левосторонние пределы.

Для измеримой относительно $\mathfrak{T}^q\times\mathfrak{X}^q$ и ограниченной на множестве $T^q\times X^q$ функции $g(t^I,\ldots, t^{(q)}; x^I,\ldots, x^{(q)}) = \varphi_1[t^I, x^I]\ldots\varphi_q\,[t^{(q)}, x^{(q)}]$ в силу теоремы Фубини моментные функции (2) существуют и имеет место равенство

$$E\left\{\int_0^1\ldots\int_0^1 dt^I\ldots dt^{(q)}\ \varphi_1[t^I,x^I]\varphi_2[t^{II},x^{II}]\ldots\varphi_q[t^{(q)},x^{(q)}]\right\} =$$

$$= \int_0^1\ldots\int_0^1 dt^I\ldots dt^{(q)}\ E\left\{\varphi_1[t^I,x^I]\varphi_2[t^{II},x^{II}]\ldots\varphi_q[t^{(q)},x^{(q)}]\right\}.$$

Этот факт используется для представления нецентральных моментов степени q для статистик, объединяемых классом $\omega_n(\varphi)$, в виде

$$M_q[\omega_n(\varphi)] = \int_0^1 dt^{(q)}\int_0^{t^{(q)}} dt^{(q-1)}\ldots\int_0^{t^{II}} dt^I\ J_q(t^I<t^{II}<\ldots<t^{(q)};n,\varphi,\ldots,\varphi) +$$

$$+\ldots+\int_0^1 dt^I\int_0^{t^I} dt^{II}\ldots\int_0^{t^{(q-1)}} dt^{(q)}\ J_q(t^{(q)}<t^{(q-1)}<\ldots<t^I;n,\varphi,\ldots,\varphi), \qquad (5)$$

и для представления, например, среднего значения произведения $\omega_n(\varphi_1)$ и $\omega_n(\varphi_2)$, необходимого для изучения корреляционных свойств $\omega_n(\varphi)$ в виде

$$E\,[\ \omega_n(\varphi_1)\omega_n(\varphi_2)\ ] = \int_0^1 dt^{II}\int_0^{t^{II}} dt^I\ J_2(\ t^I<t^{II};\ n,\varphi_1,\varphi_2) +$$

$$+\int_0^1 dt^I\int_0^{t^I} dt^{II}\ J_2(\ t^{II}<t^I;\ n,\varphi_1,\varphi_2). \qquad (6)$$

Формулы (3), (5) и (6) определяют метод исследования моментных и корреляционных свойств статистик класса $\omega_n(\varphi)$.

Указанный метод использовался для изучения свойств статистик

$$\omega(n,k) = \int_0^1 y_n^k(t)dt \qquad \text{и} \qquad \omega_A(n,k) = \int_0^1 |y_n(t)|^k dt,$$

[2] $d(x,y) = Inf_\lambda[Sup_t|x(t) - y(\lambda(t))| + Sup_t|t - \lambda(t)|]$, где $x(t)$, $y(t)\in D_{[0,1]}$, $\lambda\in\Lambda$, Λ – совокупность непрерывных монотонно возрастающих на $[0,1]$ числовых функций таких, что $\lambda(0) = 0$ и $\lambda(1) = 1$.

принадлежащих классу $\omega_n(\varphi)$ соответственно при $\varphi[t, y_n(t)] = y_n^k(t)$ и $\varphi[t, y_n(t)] = |y_n(t)|^k$. Моменты и ковариации статистик $\omega(n,k)$ и $\omega_A(n,k)$ исследовались двумя способами.

Первый способ заключался в численном определении этих характеристик для конкретных n и k с использованием алгебраических формул типа

$$E[\omega(n,k)] = \frac{1}{n^{k/2}} \sum_{s=0}^{n} \sum_{m_1=0}^{n-s} \sum_{m_2=0}^{k} C_n^s \, C_{n-s}^{m_1} \, C_k^{m_2} \, \frac{(-1)^{m_1+m_2} \, n^{m_2} \, s^{k-m_2}}{m_1 + m_2 + s + 1},$$

$$E[\omega(n,k_1)\omega(n,k_2)] = \frac{1}{n^{(k_1+k_2)/2}} \sum_{s_1=0}^{n} \sum_{s_2=0}^{n-s_1} \sum_{m_1=0}^{s_2} \sum_{m_2=0}^{n-s_1-s_2} \sum_{m_3=0}^{k_1} \sum_{m_4=0}^{k_2} C_n^{s_1} \times$$

$$\times C_{n-s_1}^{s_2} C_{s_2}^{m_1} C_{n-s_1-s_2}^{m_2} C_{k_1}^{m_3} C_{k_2}^{m_4} (-1)^{m_1+m_2+m_3+m_4} \, n^{m_3+m_4} \, \frac{1}{(s_1+s_2+m_2+m_3+m_4+2)} \times$$

$$\times \left[\frac{s_1^{k_1-m_3} \, (s_1+s_2)^{k_2-m_4}}{s_1+m_1+m_3+1} + \frac{s_1^{k_2-m_4} \, (s_1+s_2)^{k_1-m_3}}{s_1+m_1+m_4+1} \right],$$

где полагается $0^0 = 1$.

Второй способ состоял в реализации вычислений интегралов (5) и (6) для конкретных n и k при помощи системы аналитических вычислений *muMATH*. Получаемые выражения в этом случае имели точный вид, что позволило получить посредством составления и решения соответствующих систем линейных уравнений аналитические зависимости от n для ряда моментов распределений $\omega(n,k)$.

Показано, что статистики $\omega_A(n,k)$ имеют сильную корреляционную связь. То же относится и к статистикам $\omega(n,k_1)$ и $\omega(n,k_2)$ при четном $|k_1 - k_2|$.

Получен алгебраический вид статистик $\omega(n,k)$ и $\omega_A(n,k)$, удобный для практических применений :

$$\omega(n,k) = -n^{k/2}/(k+1) \sum_{i=1}^{n} \left\{ \left[(i-1)/n - t_i\right]^{k+1} - \left[i/n - t_i\right]^{k+1} \right\},$$

$$\omega_A(n,k) = n^{k/2}/(k+1) \sum_{i=1}^{n} \left\{ \left[i/n - t_i\right]^{k+1} sign^k(i/n - t_i) - \right.$$

$$\left. - \left[(i-1)/n - t_i\right]^{k+1} sign^k\left[(i-1)/n - t_i\right] \right\}.$$

Определены также пределы изменений этих статистик.

Вторая глава посвящена построению критериев согласия на основе статистик $\omega(n,k)$, изучению их свойств и сравнению по мощности.

Построенные критерии проверяют соответствие данных наблюдения нулевой гипотезе H_0: $F = F_0$.

В табл. 1 рассматривается весь спектр возможных критериев $\omega(n,k)$ в зависимости от типа альтернативной гипотезы H_1: $F \neq F_0$ (двусторонней) или H_1: $F < F_0$ ($F > F_0$)(односторонней).

Таблица 1

Критерии согласия на основе статистик $\omega(n,k)$

Тип критерия	левосторонний	правосторонний	двусторонний
Альтернатива	$F > F_0$	$F < F_0$	$F \neq F_0$
Крит. область k нечетное	$\omega(n,k) > Z_\alpha$	$\omega(n,k) < Z_\alpha$	$\lvert\omega(n,k)\rvert > Z_\alpha$
k четное	———	———	$\omega(n,k) > Z_\alpha$
Ур.-ние для Z_α k нечетное	$\Phi_{k,n}(Z_\alpha) = 1-\alpha$	$\Phi_{k,n}(Z_\alpha) = \alpha$	$\Phi_{k,n}(Z_\alpha) = 1-\alpha/2$
k четное	———	———	$\Phi_{k,n}(Z_\alpha) = 1-\alpha$

Показано, что критерии $\omega(n,k)$ устойчивы, а в случае односторонних альтернативных гипотез также несмещены и состоятельны, т.е. уровень значимости критерия не превосходит соответствующего значения мощности $\alpha \leqslant P$, и надежность разделения гипотез возрастает с увеличением числа наблюдений

$$\lim_{n\to\infty} Pr\{\omega(n,k) \in B_\alpha | H_1\} = 1.$$

Определены функции распределения $\Phi_{k,n}(z)$ статистик $\omega(n,k)$ для ряда значений n и k.

Для статистики $\omega(n,1)$ это сделано аналитически, а для случая произвольного k был построен обобщенный численный метод определения характеристической функции распределения $\omega(n,k)$ и получения на ее

основе функций $\Phi_{k,n}(z)$. Этим методом с высокой точностью (от одной – двух единиц пятого знака до единицы четвертого знака после запятой) получены таблицы процентных точек распределений $\omega(n,k)$ при $k = 1,2,\ldots,5$ и $n = 1,2,\ldots,10$.

Мощность критериев $\omega(n,k)$ и их сравнение. В работе показано, что в случае принадлежности эмпирической выборки $t_1, t_2,\ldots, t_n$ ($t_i \in [0,1]$, $i = 1,2,\ldots n$) односторонней альтернативной гипотезе G (для определенности – правосторонней $G < F_0$), для больших значений n распределение статистик $\tilde{\omega}(n,k)$ приближенно описывается гауссовским распределением с параметрами

$$a = n^{k/2}(-1)^k \int_0^1 g_k(u)du \quad \text{и} \quad \sigma^2(k,n) = n^{k-1}k^2\, D(S), \text{ где}$$

$$D(S) = 2\int_0^1 G(t_2)g_{k-1}(t_2)dt_2 \int_{t_2}^1 g_{k-1}(t_1)dt_1 - \left[\int_0^1 G(t)g_{k-1}(t)dt\right]^2$$

$$\text{и } g_k(t) = (t - G(t))^k.$$

Это позволяет оценивать уровень мощности критерия $\omega(n,k)$ практически для любой функции $G < F_0$ альтернативной гипотезы H_1 по формулам

$$P \cong 1 - \Phi\left\{\frac{Z_\alpha(k,n) - n^{k/2}\int_0^1 g_k(u)du}{\sigma(k,n)}\right\} \quad \text{для четных } k,$$

$$P \cong \Phi\left\{\frac{Z_\alpha(k,n) + n^{k/2}\int_0^1 g_k(u)du}{\sigma(k,n)}\right\} \quad \text{для нечетных } k,$$

где $\Phi(x)$ – функция распределения плотности $N(0,1)$, $Z_\alpha(k,n)$ – критическое значение соответствующего критерия для выбранных α и n.

С целью получения некоторых общих результатов в исследованиях мощности Чэпмен[3] предложил рассматривать два типа альтернативных гипотез, G_{min_0} и G_M, одна из которых минимизирует, а другая – максимизирует мощность критерия для односторонних гипотез с фиксированным $\Delta = \sup_{-\infty<x<\infty} |G(x) - F_0(x)|$.

Для этих гипотез на примере выборок $n = 50$, 200 и 1000 с помощью моделирования и аналитических оценок было произведено сравнительное

[3] *Chapman D.G.// Ann. Math. Stat., 1958, v.29, p.655.*

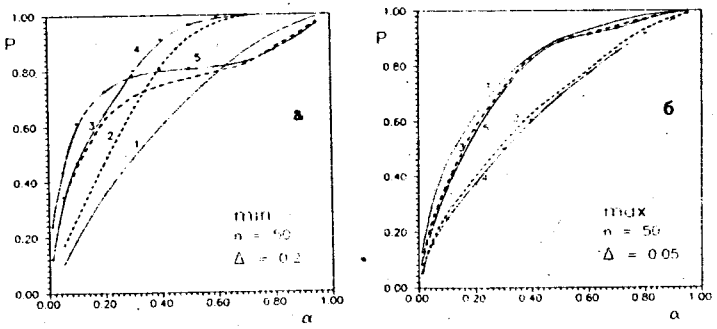

Рис.1. Зависимости мощностей критериев ω_n^1 (кривая 1), ω_n^2 (2), ω_n^3 (3), ω_n^4 (4) и ω_n^5 (5) от уровня значимости α при $n = 50$: а) для гипотезы G_{min_0} при $\Delta = 0.2$, б) для гипотезы G_M при $\Delta = 0.05$.

исследование мощности критериев $\omega(n,k)$ при $k = 1,2,\ldots,5$. Некоторые результаты представлены на рис. 1.

Для критериев ω_n^2 и ω_n^3 мощности сравнивались также и для малых $n < 10$. Нулевая и альтернативная гипотезы задавались гауссовскими распределениями с дисперсиями равными 1 и средними значениями,

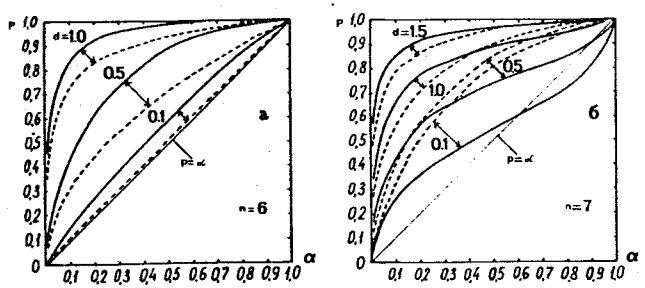

Рис.2. Сравнение мощностей критерия согласия ω_n^2 (пунктирные кривые) и одностороннего критерия ω_n^3 (сплошные кривые): а) в случае односторонней альтернативной гипотезы на примере выборки $n = 6$ и $d = 0.1$, 0.5 и 1.0; б) в случае двусторонней альтернативной гипотезы на примере выборки $n = 7$ и $d = 0.1$, 0.5, 1.0 и 1.5.

различающимися на величину d.

Кроме того, с целью проверки эффективности односторонних критериев $\omega(n,k)$ для задач с двусторонними гипотезами было произведено вычисление функций мощности для гипотез H_0: $N(0,1)$, H_1: $N(d,4)$. На рис. 2 приведены типичные результаты сравнения критериев ω_n^2 и ω_n^3 в случае малых объемов эмпирической выборки n.

Отмечается, что существует тип двусторонней альтернативной гипотезы, для которой применение критериев $\omega(n,k)$ с нечетным k оправдано и эффективно. Это составные гипотезы двустороннего типа (рис. 3б). Мощность двустороннего критерия $\omega(n,k)$ для таких гипотез определяется равенством $P = C_1P_1 + C_2P_2$, где P_1, P_2 – мощности критериев по отношению к соответствующим односторонним гипотезам.

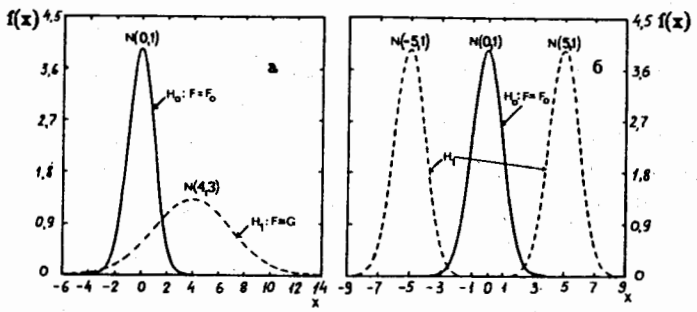

Рис.3. Два основных типа альтернативных двусторонних гипотез H_1: $F = G$ по отношению к нулевой гипотезе H_0 : $F = F_0$, заданной плотностью $N(0,1)$: а) гипотеза H_1 представлена распределением $N(4,3)$; б) составная гипотеза H_1 определена двумя гауссовскими распределениями $N(-5,1)$ и $N(5,1)$.

ЧАСТЬ II

Вторая часть диссертации посвящена применению интегральных критериев для построения статистических методов идентификации частиц и использованию этих методов в ряде задач конкретных физических экспериментов.

Рассматриваемые методы ориентированы на два класса физических задач – идентификацию частиц по суммарным спектрам масс и идентификацию частиц на основе спектров ионизационных потерь.

И в том и в другом случаях использование методов предполагает параллельные и независимые измерения таких физических величин, как времена пролета частиц или их ионизационные потери энергии, одновременно несколькими детекторами физической установки (в реальных экспериментах, как правило, предусмотрено дублирование измерений этих величин).

Измеряемые спектры являются суммами распределений от частиц разных сортов (в том числе и фона), и их идентификация означает разделение вкладов распределений от разных частиц в этих спектрах.

Физические и статистические основы метода.

Физическое содержание состоит в том, что измеряемое событие (набор измеряемых физических величин), принадлежащее частице определенного сорта имеет такую структуру, что если отсчет в первом детекторе (точка в спектре этого детектора) принадлежит какому-то из составляющих спектр распределений, то отсчеты во всех последующих детекторах также принадлежат распределениям этого типа.

Т.о. можно говорить о некоторой "коррелированности" измеряемых спектров по отношению к отдельному событию.

Статистическое содержание состоит в том, что величины отсчетов в детекторах, принадлежащих одному событию, преобразованные некоторым образом, можно рассматривать как одномерную выборку и подвергать ее проверке на соответствие какому-либо распределению.

Реализация методов (единая схема).

а) В измеренном спектре каждого детектора выделяется распределение, отвечающее частицам выбранного сорта (обычно это доминирующее распределение).

б) Все спектры преобразуются таким образом, чтобы распределения от выбранного сорта частиц можно было описывать одной и той же функцией плотности для всех детекторов.

в) Для каждого события, вызванного регистрацией отдельной частицы, из величин, одновременно принадлежащих преобразованным спектрам, конструируется выборка. Эта выборка при помощи критериев согласия $\omega(n,k)$ проверяется на соответствие гипотезе, определяемой распределением, которое было положено в основу преобразования (б).

При этом, события, отвечающие этой гипотезе, выделяются в допустимой области, а события, принадлежащие распределениям от частиц других сортов – в критической области критерия.

Реализация в конкретных задачах.

1) С помощью спектрометра МАСПИК[4] при бомбардировке дейтронами с импульсом *9* ГэВ/с мишеней из CD_2, CH_2 и *C* под углом *140* Мрад к оси первичного пучка измерялся импульсный спектр вторичных частиц в интервале от *3,5* до *5,5* ГэВ/с. Основной вклад в спектр в этой области дают протоны, примесь дейтронов составляет примерно *1 %*.

Идентификация заряженных частиц (протонов или дейтронов) осуществлялась путем измерения их импульсов и времен пролета. Импульсы частиц вычислялись по углам их отклонения в поле анализирующего магнита, а для определения времен пролета использовались сигналы от двух пар сцинтилляционных счетчиков, расстояния между которыми составляли соответственно *16,7* и *21,9* м.

Для разделения протонов и дейтронов применялся критерий ω_n^2. В соответствии с описанной выше схемой метода для преобразования спектров масс частиц использовались доминирующие в них распределения протонов, хорошо аппроксимирующиеся в данной задаче гауссовскими кривыми. На рис. *4* представлен спектр масс частиц, полученный с

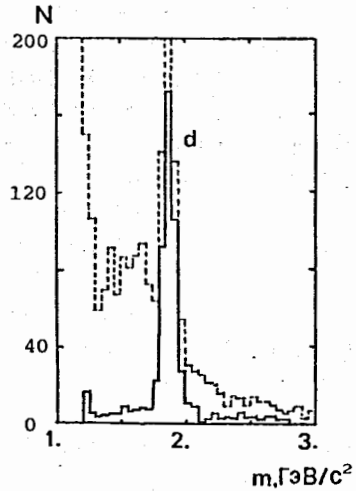

Рис.4. Спектр масс вторичных частиц, полученный с учетом данных от первой пары счетчиков: а) для масс, больших 1,25 ГэВ/с (пунктирная кривая); б) для событий с $\omega_2^2 >$ 0.666 (сплошная кривая).

учетом данных от первой пары счетчиков. Пунктирной кривой отмечен результат без применения какой-либо статистической процедуры, сплошной кривой показано распределение масс для событий, выделенных в критической области $\omega_2^2 >$ *0.666* критерия омега-квадрат.

2) В другом эксперименте на спектрометре МАСПИК, также под углом *140* Мрад при бомбардировке ядрами ^4He с импульсом *4,5* ГэВ/с/нуклон тех же мишеней измерялся импульсный спектр вторичных частиц в интервале от *5,0* до *15,0* ГэВ/с. Основной вклад в спектр в этой области дают "однозарядные" частицы (протоны – *p*, дейтроны – *d* и тритий), примесь "двузарядных" частиц (ядер ^3He и ^4He) не превышает *0,1 %*.

Кроме времен пролета идентификация частиц проводилась по ионизационным потерям в пяти сцинцилляционных счетчиках.

На рис. *5* представлено распределение амплитуд выходных импульсов для одного из сцинцилляционных счетчиков основного плеча спектрометра. Доминирующее распределение в этом спектре вызвано регистрацией однозарядных частиц. На рисунке также отмечена область

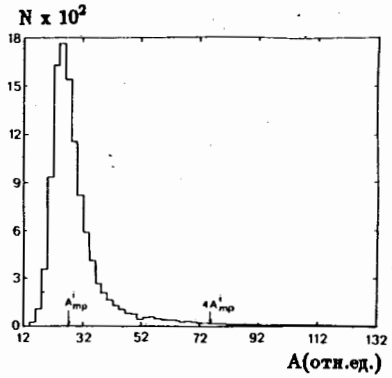

Рис.5. Распределение амплитуд выходных импульсов для одного из сцинцилляционных счетчиков основного плеча спектрометра МАСПИК.

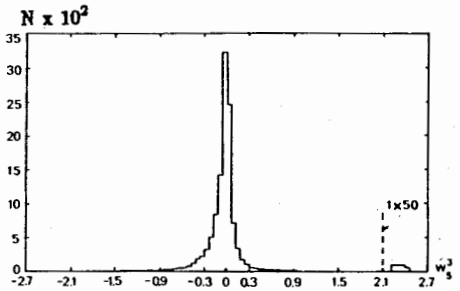

Рис.6. Распределение случайных величин ω_n^3, полученное в результате обработки амплитуд выходных импульсов от пяти сцинцилляционных счетчиков основного плеча спектрометра МАСПИК.

ожидаемого проявления событий, вызванных двузарядными частицами. Для проверки этого предположения использовался односторонний критерий ω_n^3. В качестве нулевой гипотезы бралось распределение Ландау. События , вызванные двузарядными частицами, выделялись в критической области критерия ω_5^3 (рис. *6*). Всего было выделено *7* событий (^3He и

[4]Описание установки и основных задач эксперимента дано в работах *L.S.Azhgirey et all.//Nuclear Physics A528 (1991) 621-646*, Ажгирей Л.С. и др.//Ядерная физика, т.*46*, вып.*4(10)*, *1987*.

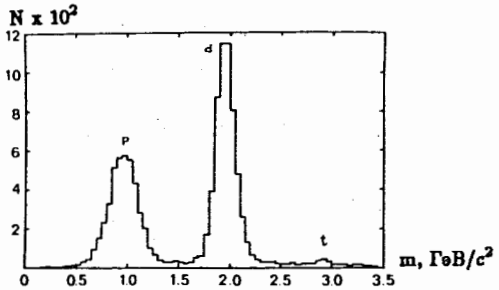

Рис.7. Спектр масс однозарядных частиц, зарегистрированных основным плечом спектрометра МАСПИК.

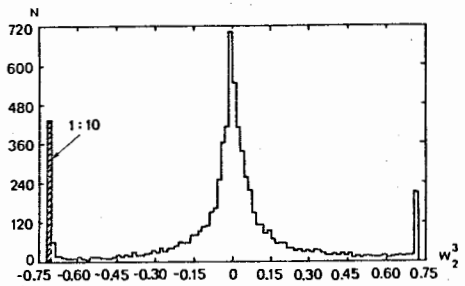

Рис.8. Распределение случайных величин ω_n^3, полученное в результате обработки спектров масс однозарядных частиц, измеряемых с помощью двух пар времяпролетных сцинцилляционных счетчиков в основном плече спектрометра МАСПИК.

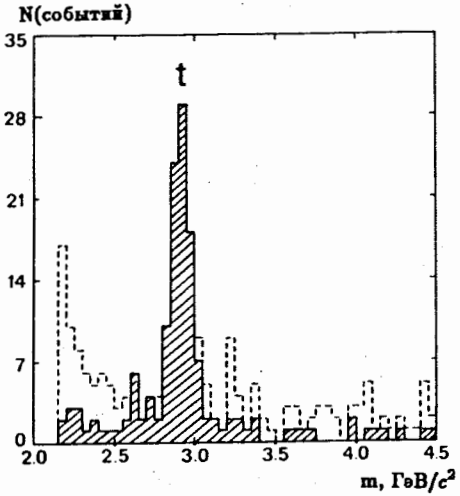

Рис.9. Спектр масс однозарядных частиц, отвечающий событиям с $\omega_n^3 >$ 0.66, - пунктирная гистограмма; сплошной гистограммой представлен результат повторной обработки указанных событий с помощью критерия ω_n^3.

^4He). Для разделения вкладов однозарядных частиц (p,d и t) спектры масс вторичных частиц (рис. 7) подвергались обработке при помощи двустороннего критерия ω_n^3. Распределение величин ω_2^3 показано на рис. 8. В критических областях $\omega_2^3 < -0.66$ и $\omega_2^3 > 0.66$ выделены события, отвечающие протонам и тритию соответственно. Окончательный спектр масс частиц, представленный на рис.9, демонстрирует хорошее выделение частиц трития с помощью применяемого метода.

Сравнение методик ω_n^1, ω_n^2 и ω_n^3 с традиционными методами и методом отношения функций правдоподобия было проведено на основе моделирования распределений, приведенных в работе по разделению протонов и заряженных пионов с энергией 100 ГэВ в потоках космических лучей[5]. Измерения ионизационных потерь проводились

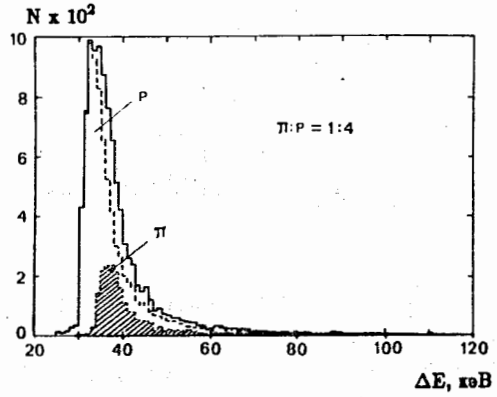

Рис.10. Суммарное распределение ионизационных потерь для 100 ГэВ-ных протонов и пионов в отдельном детекторе (сплошная гистограмма): вклад от протонов – пунктирная гистограмма, вклад от пионов – заштрихованная гистограмма.

одновременно несколькими детекторами экспериментальной установки (n = 6). Суммарное распределение ионизационных потерь для одного детектора показано на рис. 10.

Использовались методы ω_n^1, ω_n^2 и ω_n^3, метод отношения функций правдоподобия, а также традиционные методы статистического анализа измерений с многослойных детекторов – метод отбора по минимальной амплитуде, метод усреднения импульсов со всех детекторов с предварительным отбросом больших амплитуд и метод подсчета числа детекторов, импульсы с которых не превышают заранее выбранной величины. На рис.11 а,б представлены распределения случайных величин ω_6^2 и ω_6^3.

---

[5] *P.V.Ramaha Murty and G.D.Demeester//Nucl. Instr. and Meth., 1967,56, p.93.*

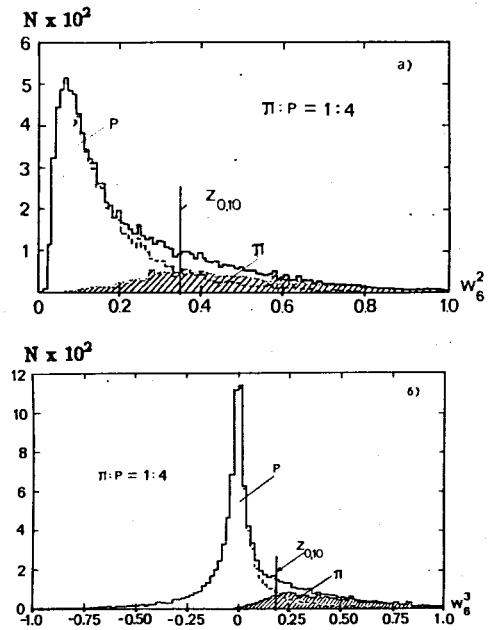

Рис.11. Распределения случайных величин ω_n^2 (а) и ω_n^3 (б), полученные в результате обработки смоделированных спектров ионизационных потерь (n = 6): суммарное распределение – сплошная гистограмма; вклад протонных событий – пунктирная гистограмма; вклад пионных событий – заштрихованная гистограмма.

Проведенное сравнение показало, что в данной задаче метод ω_n^3 превосходит как вышеперечисленные традиционные методы, так и методы ω_n^1 и ω_n^2 и уступает в мощности, хотя и очень незначительно, методу отношения функций правдоподобия, обладая по сравнению с ним рядом важных преимуществ.

Использование метода ω_n^3 для выделения подпороговых K^+ мезонов. Установка с предварительным названием *"O⁰Facility"*[6] предназначена для проведения эксперимента по изучению подпорогового рождения K^+-мезонов в столкновениях протонов с легкими ядрами (^2H, ^{12}C) путем проведения инклюзивных и эксклюзивных измерений дифференциальных сечений на ускорителе *COSY* (*Julich, Germany*), позволяющем ускорять ядра (от водорода до серы) до импульсов от *270* до *3300* Мэв/с.

Экспериментально необходимо решить задачу практически однозначной

[6] *W.Borgs et al. Proposal for K^+ - Meson Spectroscopy with "O⁰Facility" at TP2 in COSY.*

регистрации K^+ – мезонов на подавляющем фоне π^+ – мезонов. Соотношение вкладов этих частиц ожидается равным $N_{K^+}/N_{\pi^+} \approx 10^{-5}$. Решение этой задачи было проведено в два этапа. На первом этапе происходил отбор событий по измерениям ионизационных потерь в двух сцинцилляционных счетчиках установки с использованием метода отбора по минимальной амплитуде, что позволило подавить фон в $2\cdot10^3$ раз. На втором этапе к спектрам масс оставшихся частиц был применен критерий ω_n^3. Окончательный спектр масс K^+-мезонов представлен на рис. *12*. Примесь пионов среди выделенных частиц составляет менее *15 %*.

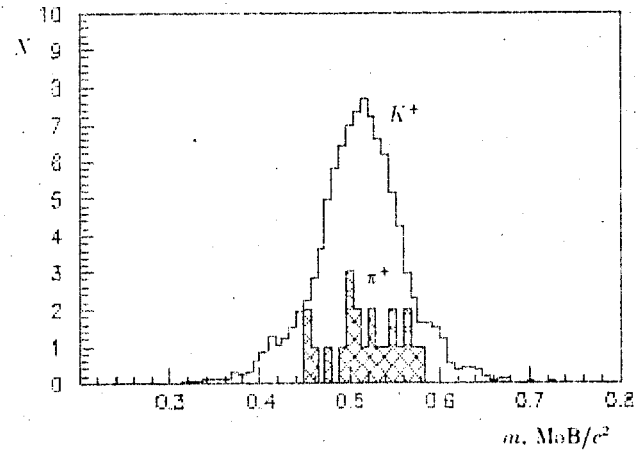

Рис.12. Спектр масс вторичных частиц, попавших в допустимую область критерия ω_2^3 с уровнем значимости α = 0.04.

В приложении приводятся таблицы процентных точек распределений статистик $\omega(n,k)$ при $k = 1,2,\ldots,5$ и $n = 1,2,\ldots,10$.

В заключении сформулированы основные выводы диссертации:

1) Предложен метод вычисления моментов и ковариаций для некоторого класса интегральных непараметрических статистик. Метод основан на свойствах моментных функций эмпирического процесса.

2) С помощью этого метода исследованы моментные и корреляционные свойства статистик $\omega(n,k)$ и $\omega_A(n,k)$, представимых соответственно в виде интеграла от k-й степени эмпирического процесса и интеграла от

модуля k-й степени процесса $y_n(t)$.

Исследованы корреляционные связи статистик Мозеса (U), Ватсона (U_n^2), омега-квадрат (ω_n^2) и Андерсона-Дарлинга (A_n^2) со статистиками $\omega(n,k)$.

3) Предложен обобщенный метод вычисления функций распределения статистик $\omega(n,k)$. Этим методом с высокой точностью вычислены таблицы процентных точек распределений $\omega(n,k)$ для ряда значений n и k.

4) На основе статистик $\omega(n,k)$ построены одно- и двусторонние критерии согласия и изучены их общие свойства.

С использованием численного моделирования, а также некоторых аналитических оценок исследованы мощности критериев $\omega(n,k)$ по отношению к ряду односторонних альтернативных гипотез.

5) Критерии ω_n^2 и ω_n^3 использовались для построения методов идентификации частиц по измерениям их времен пролета или ионизационных потерь энергии одновременно несколькими детекторами экспериментальной установки. Методы были опробованы на ряде данных реальных физических экспериментов, проводимых на установке МАСПИК.

6) На основе анализа данных, полученных моделированием физических распределений, опубликованных в других работах, произведено сравнение эффективности применения построенных методов и традиционных методов, предназначенных для решения некоторых типов подобных задач.

7) На основании результатов моделирования проектируемой установки для изучения подпорогового рождения K^+ - мезонов, с помощью новой методики была предложена процедура надежного выделения изучаемых частиц на фоне доминирующих процессов.

Работы положенные в основу диссертации:

1. Зрелов П.В. Обобщенные моментные функции эмпирического процесса и интегральные непараметрические статистики. Сообщение ОИЯИ, *Р11-92-398*, Дубна, *1992*.
2. Зрелов П.В., Иванов В.В. Проверочная статистика $\omega_n^3 = n^{3/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$ в теории непараметрических критериев согласия. Сообщение ОИЯИ , *Р10-88-321*, Дубна, *1988*.
P.V.Zrelov, V.V.Ivanov. Test Statistics $\omega_n^3 = n^{3/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$ *and its main properties. Transactions of the Eleventh Prague Conference on Information theory, Statistical decision functions, random processes, Academia, Prague 1992.*
3. П.В.Зрелов, В.В.Иванов. Функции распределения статистики Смирнова - Крамера - Мизеса для малых выборок. Препринт ОИЯИ, *Р10-86-547*, Дубна, *1986*.
4. Зрелов П.В., Иванов В.В. Критерии согласия, основанные на проверочной статистике $\omega_n^3 = n^{3/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$. Определения и свойства, мощность для малых n. Сообщение ОИЯИ *Р10-89-577*, Дубна, *1989*.
5. Зрелов П.В., Иванов В.В. Критерии согласия, основанные на проверочной статистике $\omega_n^3 = n^{3/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$. Исследование мощности одностороннего критерия для больших значений n. Сообщение ОИЯИ, *Р11-92-409*, Дубна, *1992*.
6. Зрелов П.В., Иванов В.В. Функции распределения статистики $\omega_n^3 = n^{3/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^3 dP(x)$ для малых n. Препринт ОИЯИ, *Д11-92-139*, Дубна, *1992*.
7. Зрелов П.В., Иванов В.В. Интегральные непараметрические статистики $\omega_n^k = n^{k/2}\int_{-\infty}^{\infty} [S_n(x) - P(x)]^k dP(x)$ и их основные свойства. Алгебраический вид, функции распределения и критерии согласия. Сообщение ОИЯИ, *Р10-92-461*, Дубна, *1992*.
8. *Zrelov P.V., Ivanov V.V. Determination of Distribution Parameters for* ω_n^2 *and* ω_n^3 *Test Statistics by the Computer Algebra Application. IV International Conference on Computer Algebra in Phisical Research, World Scientific. Singapore. New Jersey. London. Hong Kong. 1991, p.329-334.*
9. Зрелов П.В., Иванов В.В. Метод выделения маловероятных событий с помощью критерия согласия Смирнова- Крамера- Мизеса. Препринт ОИЯИ, Р10-86-812, Дубна, *1986*.
P.V.Zrelov, V.V.Ivanov. The small probability events separation method based on Smirnov - Cramer - Von Mises test. Совместный научный сборник Объединенного института ядерных исследований (Дубна, ОИЯИ) и центрального института физических исследований (Будапешт, ВНР). Выпуск шестой, *KFKI - 1989 - 62/M*. Будапешт, *1989*.
10. П.В.Зрелов, В.В.Иванов. Метод идентификации релятивистских частиц на основе критерия согласия ω_n^3 . Препринт ОИЯИ, *Р10-89-739*, Дубна, *1989*.
P.V.Zrelov, V.V.Ivanov. The Relativistic Charged Particles Identification Method Based on the Goodness-of-Fit ω_n^3*- Criterion. Nucl. Instr. and Meth. in Phys. Res., A310, 1991, 623.*

11. *P.V.Zrelov, V.V.Ivanov. Experimental Data Analysing Methods Based on Nonparametric Goodness-of-Fit Criterion* ω_n^3 *. Transactions of the Computing in High Energy Physics'91. Universal Academy Press, Inc., Tokyo, Japan, 1991.*
12. П.В.Зрелов, В.В.Иванов, В.И.Комаров, А.И.Пузынин, А.С.Хрыкин. Моделирование эксперимента по изучению процессов подпорогового рождения K^+- мезонов. Препринт ОИЯИ, *Р10-92-369*, Дубна, *1992*.

Рукопись поступила в издательский отдел
18 ноября 1992 года.