Институт ядерной физики им. Г.И. Будкера
СО РАН

А.Д. Букин

КРИТЕРИЙ χ^2 ДЛЯ РАСПРЕДЕЛЕНИЯ ПУАССОНА

ИЯФ 97-104

Новосибирск

Критерий χ^2 для распределения Пуассона

А.Д. Букин

Институт ядерной физики им. Г.И.Будкера
Новосибирск, 630090

Аннотация

В работе рассматриваются несколько вариантов использования критерия χ^2 для статистики Пуассона и исследуются границы их применимости при уменьшении математического ожидания числа событий.

χ^2 Criteria for Poisson Statistics

A.D. Bukin

Budker Institute of Nuclear Physics
630090 Novosibirsk, Russia

Abstract

In the paper some variants of χ^2 criteria application for Poisson statistics are considered. Limits of their applicability are investigated for decreasing mean value of the number of events.

1. Введение

При математической обработке результатов экспериментов наиболее часто для проверки согласия теории и эксперимента применяется критерий χ^2 [1] в силу своей простоты и универсальности. Однако, довольно часто при обработке эксперимента статистика отсчётов мала, и критерий χ^2 становится неприменимым. Типичный пример — аппроксимация экспериментальных данных в диапазоне энергий вблизи резонанса при низком уровне фона. Тогда очевидный приём суммирования близких точек по энергии до достижения статистики в каждой точке, приемлемой для применения критерия χ^2, не всегда подходит, так как для этого может требоваться суммирование событий в точках, существенно удалённых по энергии. С другой стороны, жаль отказываться от применения критерия χ^2 по причине недостаточной статистики на "хвостах" резонанса, тогда как большая часть точек вблизи массы резонанса может вполне удовлетворять критерию применимости $n >> 1$.

В такой ситуации для проверки степени согласия теории и эксперимента можно действовать в двух направлениях: или воспользоваться другим критерием согласия, или ввести модифицированную функцию χ^2 такую, чтобы полученная величина соответствовала стандартному распределению вероятностей χ^2.

В данной работе исследуется второй путь: рассматривается несколько вариантов модификации χ^2 и исследуются границы применимости к ним стандартного распределения χ^2 при малых значениях математического ожидания.

2. Выбор вариантов модификации χ^2

Рассмотрим несколько вариантов модификации χ^2 при малой статистике отсчётов с целью изучить границы применимости стандартного распределения χ^2. Будем обозначать математическое ожидание числа событий как p, предполагая справедливость распределения Пуассона

$$W_n = \frac{p^n}{n!}e^{-p}, \quad p > 0, \quad n = 0, 1, \ldots \infty \tag{1}$$

Во всех вариантах модифицирования χ^2 будем использовать только те выражения, которые при большой статистике асимптотически стремятся к правильному виду χ^2 для распределения Пуассона для одной степени свободы:

$$\chi^2 = \frac{(n-p)^2}{p}, \text{ при } \ p >> 1 \tag{2}$$

Первые два варианта для вычисления χ^2 (по-видимому, совершенно неудовлетворительные при малой статистике) возьмём в том виде, как для гауссова распределения, с ошибками, оцениваемыми независимо от теоретической модели только по числу событий n, без привлечения математического ожидания p.

1. $\chi_1^2 = \dfrac{(n-p)^2}{n+1}$, ошибка измерения в каждой точке оценена, как $\sigma = \sqrt{n+1}$.

2. $\chi_2^2 = \dfrac{(n-p)^2}{n+0.5}$, ошибка измерения в каждой точке оценена, как $\sigma = \sqrt{n+0.5}$. Принять за ошибку измерения $\sigma = \sqrt{n}$ невозможно, так как в тех случаях, когда $n = 0$, придётся делить на ноль.

3. $\chi_3^2 = \dfrac{(n-p)^2}{p}$. Этом вариант хорош тем, что обеспечивает правильное значение среднего $< \chi_3^2 >= 1$ при всех значениях p, однако, дисперсия $D(\chi_3^2) = 2 + 1/p$, что при малых значениях p сильно отличается от правильного значения $D(\chi^2) = 2$.

4. $\chi_4^2 = 2 \cdot [p - n + n\log(n/p)]$. Это выражение является популярной записью логарифмической функции правдоподобия для пуассонова распределения [2–4]. Это выражение привлекает тем, что при своей простоте оно в пределе $p \to \infty$ переходит в асимптотическое выражение (2), к тому же оно используется в оптимизируемой логарифмической функции правдоподобия.

5. $\chi_5^2 = 2 \cdot \left[p - n\ln p + \begin{cases} 0, & \text{при } n = 0 \\ \ln(n!) - 0.5\ln(2\pi n), & \text{при } n > 0 \end{cases} \right]$

   Это выражение также может быть использовано для записи логарифмической функции правдоподобия и также переходит в правильное выражение для χ^2 при большой статистике.

6. Следующий вариант χ_6^2 получим из χ_3^2 введением таких поправок, которые обеспечивают правильное значение дисперсии 2 и среднего 1.

$$\chi_6^2 = \frac{(n-p)^2}{\sqrt{p(p+0.5)}} + \frac{1}{2p + 1 + \sqrt{2p(2p+1)}} \tag{3}$$

7. Наконец, последний вариант χ_7^2 получим из χ_4^2 введением аналогичных поправок с целью получения правильных дисперсии и среднего. Однако, ситуация здесь усложняется тем, что такая операция может быть проделана только численно в ограниченном интервале p. Поправку будем искать в виде

$$\chi_7^2 = F_1(p)\chi_4^2 + F_2(p) \tag{4}$$

в два приёма. Вначале получим множитель $F_1(p)$ в виде

$$F_1(p) = \frac{p^2 + ap + b + c\cdot\ln p + d\cdot\ln^2 p}{g + p^2 + h\cdot\ln p + r\cdot\ln^2 p} \tag{5}$$

методом наименьших квадратов (дисперсия $D(\chi_7^2)$ не зависит от $F_2(p)$). Численной минимизацией суммы квадратов отклонений функции $F_1(p)$ от требуемых значений в одиннадцати точках в интервале $p \in (0.01, 10)$ (логарифмически равномерно распределённых) получены значения коэффициентов

$$\begin{aligned} a &= -0.8991, \quad b = 1.318, \quad c = 0.548, \quad d = 0.1308, \\ g &= 0.1684, \quad h = -9.09\cdot 10^{-2}, \quad r = 1.0\cdot 10^{-2} \end{aligned} \tag{6}$$

Затем аналогичным приёмом получена аппроксимация функции $F_2(p)$ в виде

$$F_2(p) = \frac{s + t\cdot\ln p + u\cdot\ln^2 p}{p^2 + x\cdot\ln^2 p + y} \tag{7}$$

с коэффициентами

$$\begin{aligned} s &= -0.5543, \quad t = 9.338\cdot 10^{-2}, \quad u = 0.2174 \\ x &= 0.1878, \quad y = 0.4273 \end{aligned} \tag{8}$$

Очевидно, что при $p << 0.01$ значения дисперсии и среднего для χ_7^2 будут принимать значения, далёкие от правильных, однако, применение критерия χ^2 в этой области значений p выглядит очень искусственным.

На рис. 1 приведена зависимость среднего значения для всех перечисленных здесь модификаций χ^2 для переменной, распределённой по закону Пуассона со средним значением p, в интервале $p \in (0.01, 1)$.

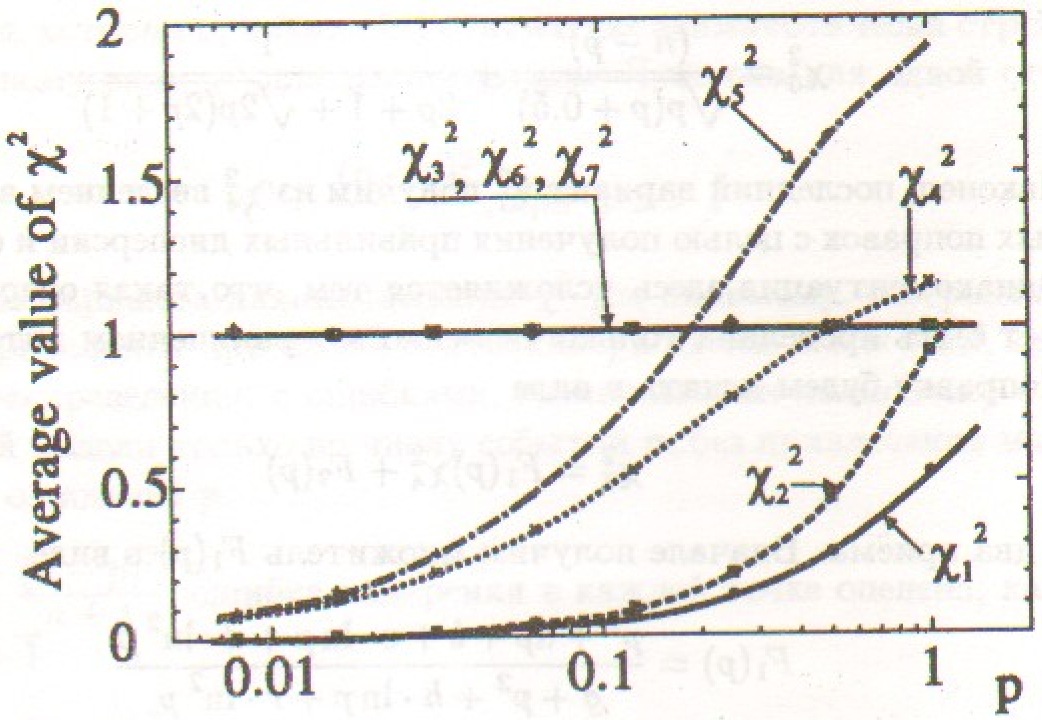

Рис. 1: Среднее значение χ^2 при одной степени свободы для нескольких вариантов модификации χ^2 при малом параметре p пуассонова распределения

Аналогично, рис. 2 демонстрирует зависимость дисперсии этих величин от p. Напомним, что здесь весь анализ χ^2 проводился с одной степенью свободы — в случае нескольких степеней свободы значения среднего и дисперсии складываются. Несмотря на то, что для некоторых вариантов вычисления χ^2 дисперсия и среднее могут быть вычислены аналитически, на рисунках представлены единообразные результаты численных расчётов методом Монте-Карло с генерацией псевдослучайных чисел, распределённых по закону Пуассона (статистика равна 10^6 событий).

Последний вариант χ_7^2 выглядит несколько громоздким. Очевидно, это обусловлено стремлением обеспечить асимптотический переход к правильным значениям дисперсии и среднего при $p \to \infty$, а также конечное значение поправочных функций F_1 и F_2 при $p \to 0$.

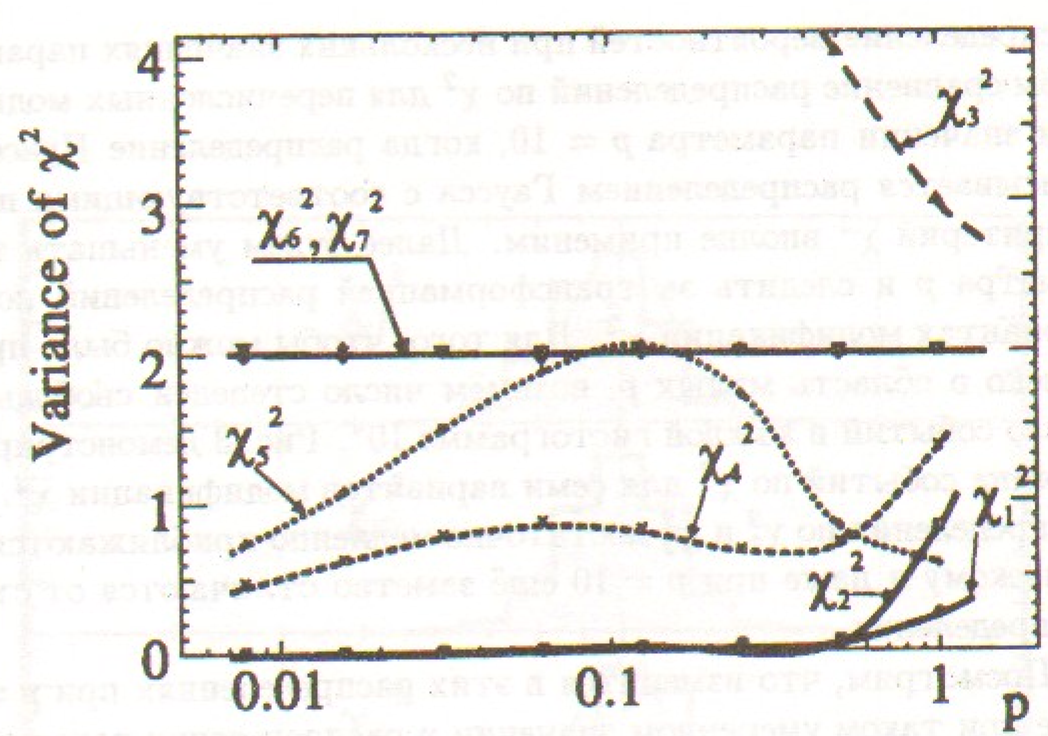

Рис. 2: Значение дисперсии χ^2 при одной степени свободы для нескольких вариантов модификации χ^2 при малом параметре p пуассонова распределения

3. Применимость критерия χ^2

Применимость критерия χ^2 для экспериментальных данных, подчиняющихся закону распределения вероятностей Пуассона, хорошо видна из рис. 1 и 2. Там же можно увидеть, какие варианты модификации χ^2 более предпочтительны при математическом ожидании числа событий p порядка или меньше единицы. Например, при $p = 0.1$ средний вклад каждой степени свободы для варианта вычисления χ_4^2 уже в два раза меньше стандартного, а вклад в дисперсию этого варианта в два раза меньше стандартного уже при $p = 0.7$.

При проверке применимости критерия χ^2 в случае распределения Пуассона следует иметь в виду, что желательно выполнение соотношения $p \cdot n_D >> 1$ (произведение параметра распределения Пуассона на число степеней свободы), иначе в распределении по χ^2 только несколько дискретных значений χ^2 являются значимыми, и сравнивать полученное дискретное распределение по χ^2 со стандартным непрерывным распределением становится затруднительным.

Значения среднего и дисперсии, естественно, не полностью определяют распределение вероятностей по χ^2. К сожалению, провести какие-либо исследования общего характера представляется затруднительным, однако, можно привести несколько иллюстраций, как трансформируется распределение вероятностей при нескольких значениях параметра p. Начнём сравнение распределений по χ^2 для перечисленных модификаций χ^2 при значении параметра $p = 10$, когда распределение Пуассона хорошо описывается распределением Гаусса с соответствующими параметрами и критерий χ^2 вполне применим. Далее будем уменьшать значение параметра p и следить за трансформацией распределений по χ^2 во всех вариантах модификации χ^2. Для того, чтобы можно было продвинуться далеко в область малых p, возьмём число степеней свободы $n_D = 100$, число событий в каждой гистограмме 10^4. Рис. 3 демонстрирует распределение событий по χ^2 для семи вариантов модификации χ^2. Видно, что распределения по χ^2_1 и χ^2_2 достаточно медленно приближаются к асимптотическому и даже при $p = 10$ ещё заметно отличаются от стандартного распределения.

Посмотрим, что изменится в этих распределениях при $p = 1$ (рис. 4). Уже при таком умеренном значении p распределения разошлись и стали более различимы. Наилучшее совпадение остаётся у χ^2_6 и χ^2_7.

На следующем рисунке 5 значение p выбрано существенно меньше 1 ($p = 0.1$). Многие распределения здесь вообще не перекрываются со стандартным, поэтому ошибка в оценке $P(\chi^2)$ получается значительной. Так как для проверки совместимости теории и экспериментальных данных используется величина $P(\chi^2)$, то удобнее наблюдать отличия каждого из распределений от стандартного по графику следующего типа. Для каждого значения χ^2 точка графика для i-ой модели χ^2_i имеет абсциссу, равную $P(\chi^2)$, рассчитанную по стандартному распределению, а ординату, равную $P(\chi^2_i)$, расчитанную по распределению χ^2_i. Графики такого типа приведены на рис. 6, 7, 8, 9, 10 и 11 для значений $p = 4; 2; 1; 0.5; 0.2$ и 0.1, соответственно. Уже при умеренных $p \sim 1$ наблюдаются значительные отклонения интегральных распределений вероятности для вариантов вычисления $\chi^2_1, \chi^2_2, \chi^2_4$ и χ^2_5. Наилучшими приближениями остаются χ^2_3, χ^2_6 и χ^2_7. Неожиданным в этом выводе является то, что вариант χ^2_4 оказался заметно хуже, чем χ^2_3 в этой переходной области параметра p.

В области малых значений параметра $p << 1$ из этих рисунков однозначно следует, что наименьшие отклонения от стандартного распределения по χ^2 имеют варианты вычисления χ^2_6 и χ^2_7, предложенные в данной работе.

Посмотрим теперь, как сказывается дискретность распределения при $p = 0.1$. Для этого анализа оставим только варианты χ^2_6 и χ^2_7. Будем уменьшать постепенно количество степеней свободы. На рис. 12 количество степеней свободы $n_D = 50$, при этом уже $p \cdot n_d = 5$. Дискретность распределений по χ^2_6 и χ^2_7 отчётливо проявляется в виде точек излома

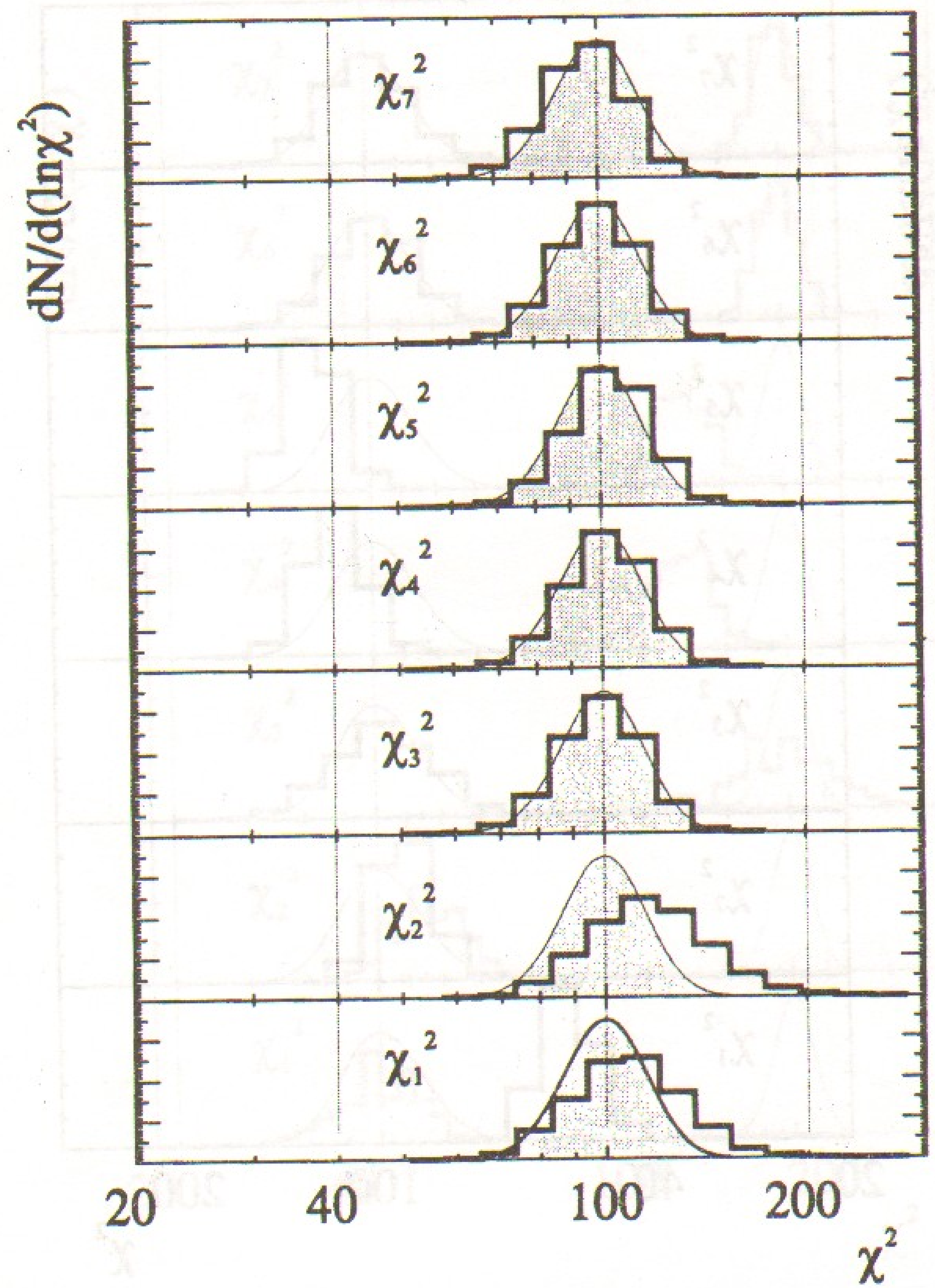

Рис. 3: Распределение по χ^2 событий с числом степеней свободы $n_D = 100$, $p = 10$. Сплошная кривая — стандартное распределение χ^2.

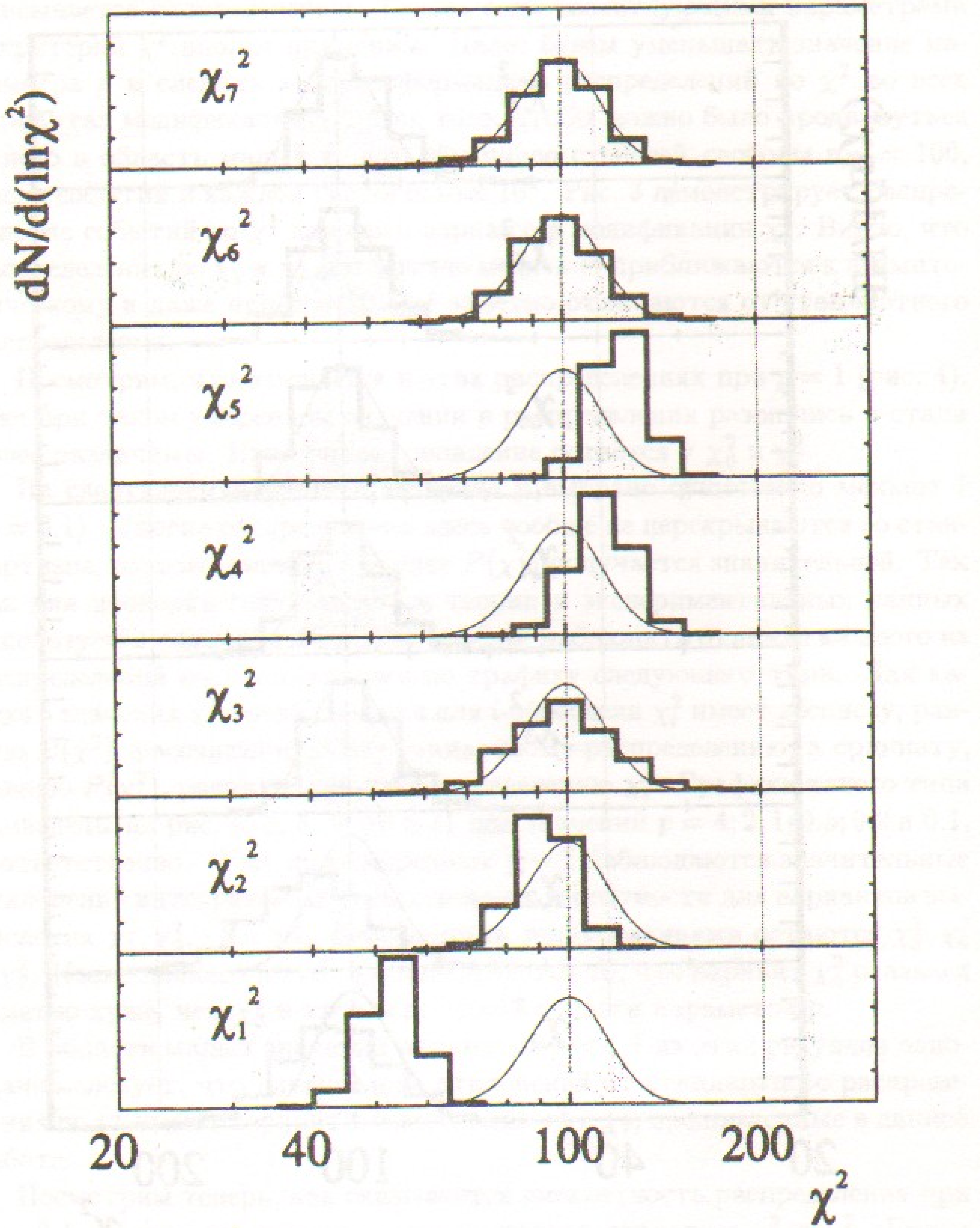

Рис. 4: Распределение по χ^2 событий с числом степеней свободы $n_D = 100$, $p = 1$. Сплошная кривая — стандартное распределение χ^2.

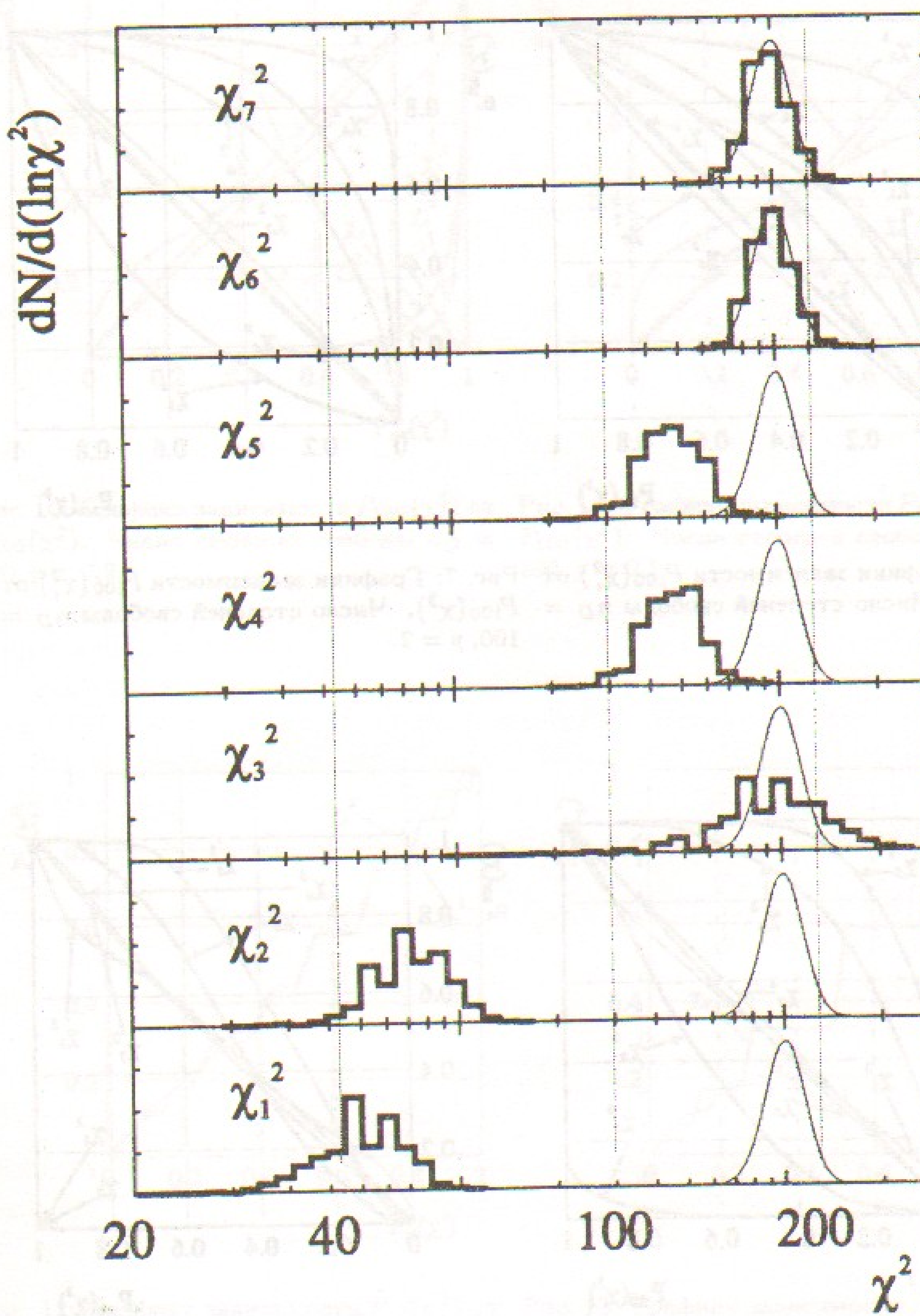

Рис. 5: Распределение по χ^2 событий с числом степеней свободы $n_D = 100$, $p = 0.1$. Сплошная кривая — стандартное распределение χ^2.

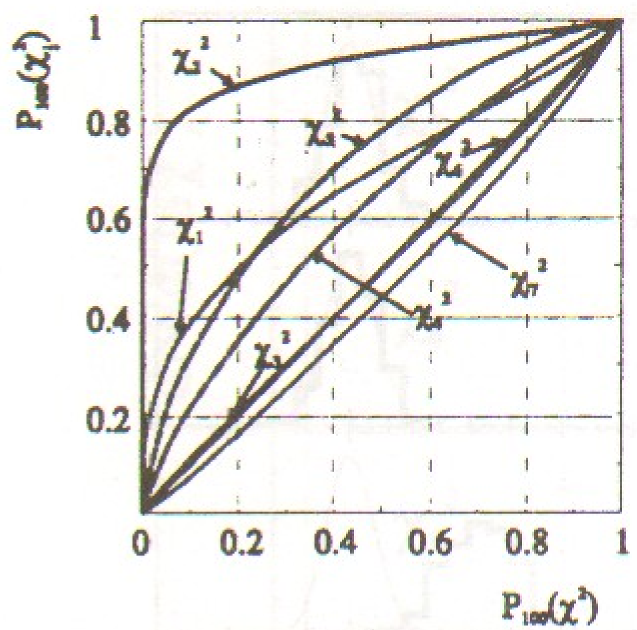

Рис. 6: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 4$.

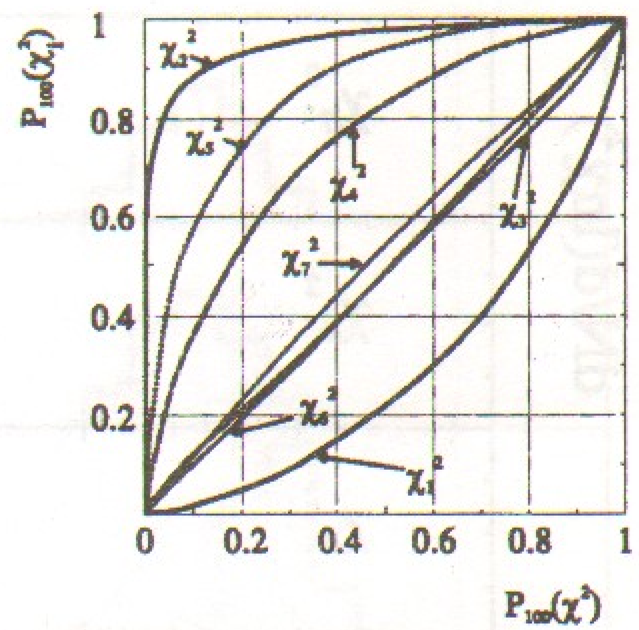

Рис. 7: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 2$.

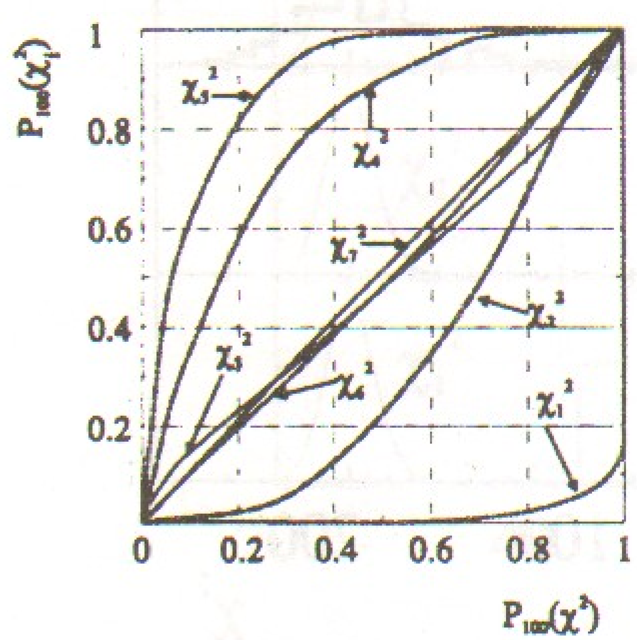

Рис. 8: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 1$.

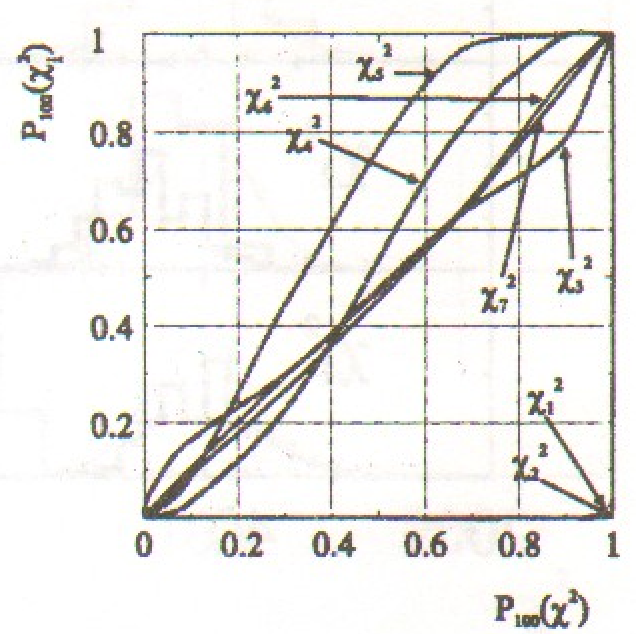

Рис. 9: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 0.5$.

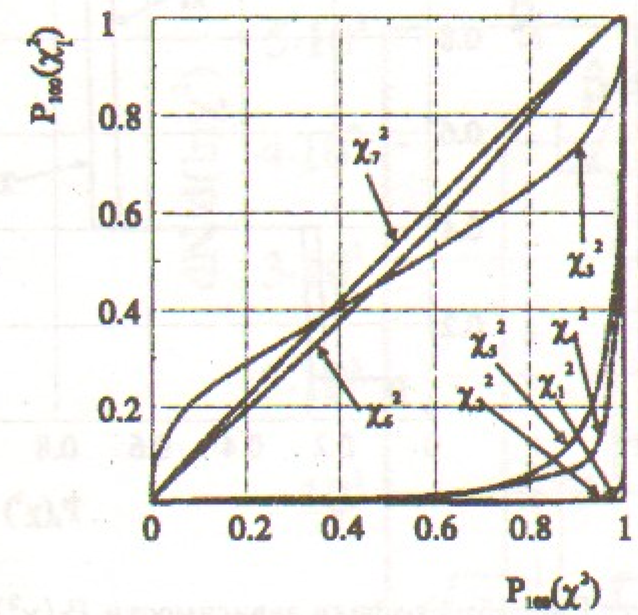

Рис. 10: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 0.2$.

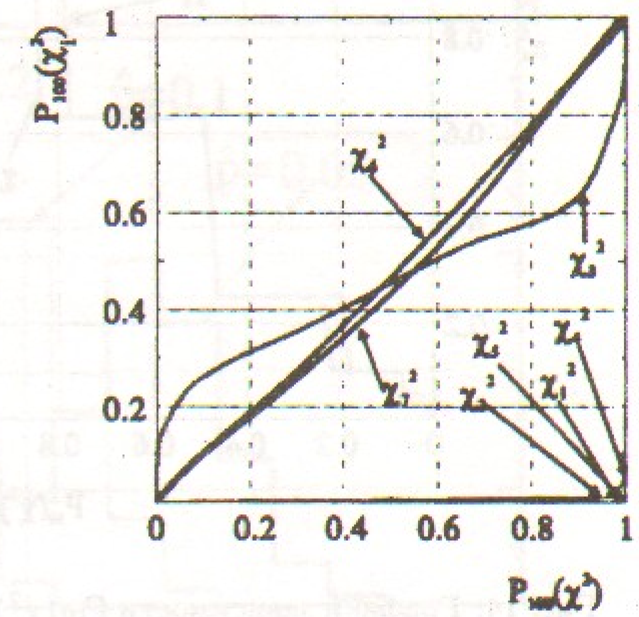

Рис. 11: Графики зависимости $P_{100}(\chi_i^2)$ от $P_{100}(\chi^2)$. Число степеней свободы $n_D = 100$, $p = 0.1$.

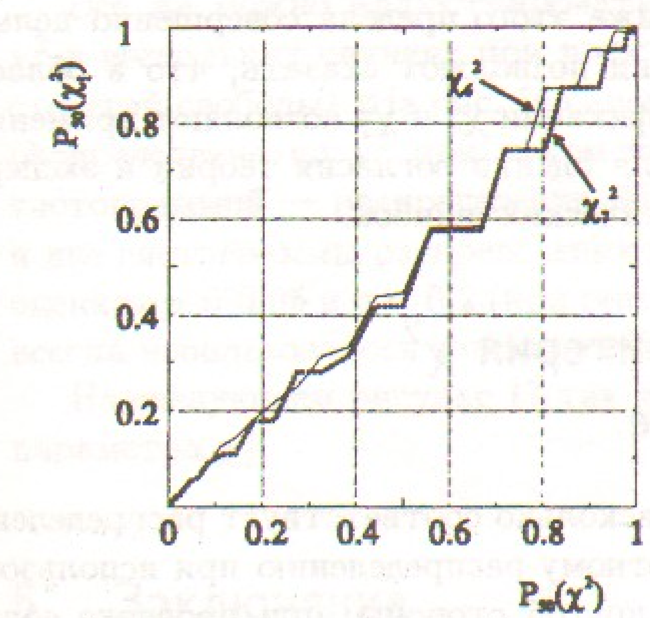

Рис. 12: Графики зависимости $P_{50}(\chi_i^2)$ от $P_{50}(\chi^2)$. Число степеней свободы $n_D = 50$, $p = 0.1$.

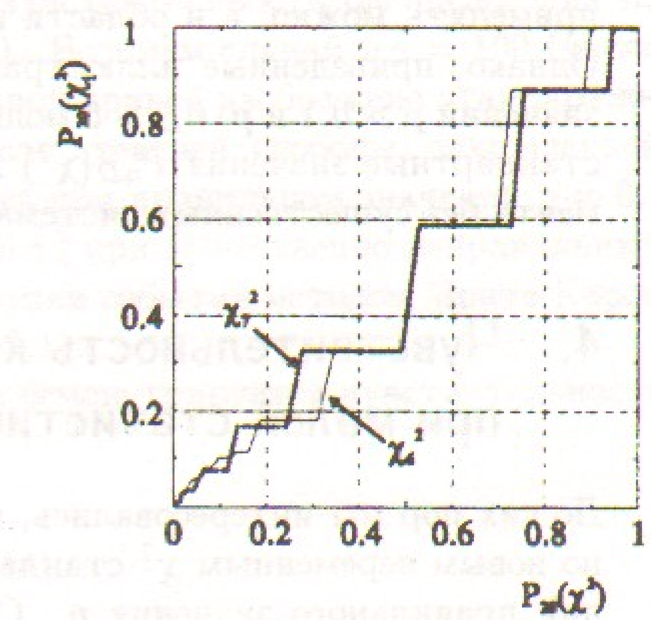

Рис. 13: Графики зависимости $P_{20}(\chi_i^2)$ от $P_{20}(\chi^2)$. Число степеней свободы $n_D = 20$, $p = 0.1$.

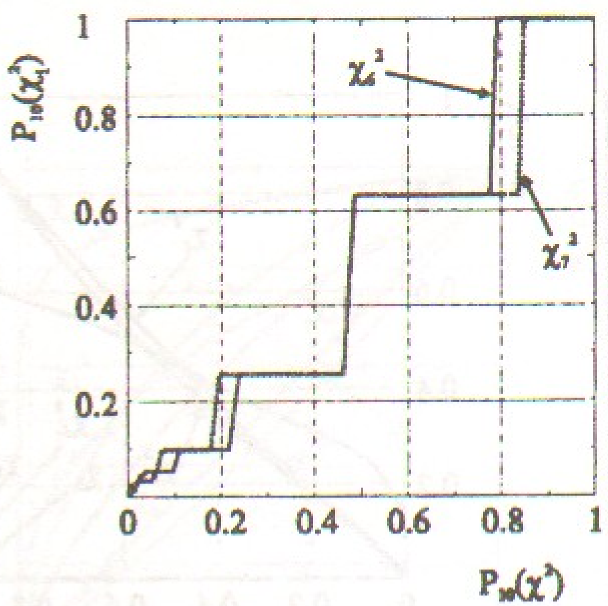

Рис. 14: Графики зависимости $P_{10}(\chi_i^2)$ от $P_{10}(\chi^2)$. Число степеней свободы n_D = 10, $p = 0.1$.

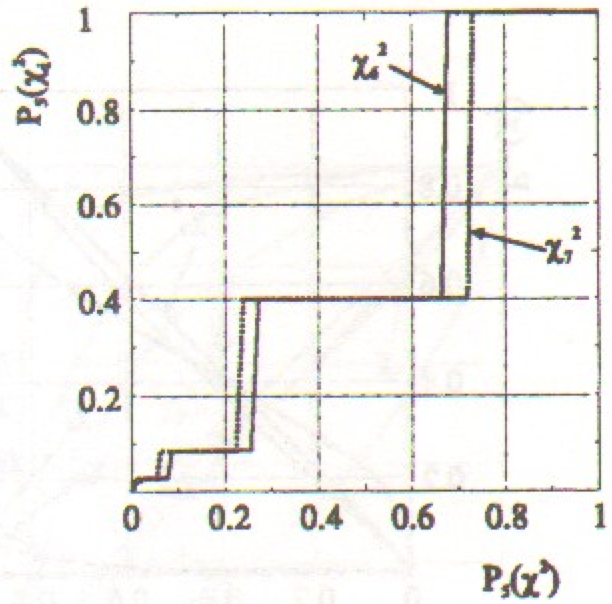

Рис. 15: Графики зависимости $P_5(\chi_i^2)$ от $P_5(\chi^2)$. Число степеней свободы $n_D = 5$, $p = 0.1$.

кривых. На последующих рисунках 13, 14 и 15, с количеством степеней свободы, соответственно, равным 20, 10 и 5, это явление проявляется всё более ярко. Очевидно, что нельзя поставить точного предела по параметру p или количеству степеней свободы, выше которого критерий χ^2 применять можно, а в области ниже этого предела совершенно нельзя. Однако, приведённые иллюстрации позволяют сказать, что в области значений $p > 0.1$ и $p \cdot n_D > 5$ модификации χ_6^2 и χ_7^2 позволяют применять стандартные значения $P_{nD}(\chi^2)$ для оценки согласия теории и эксперимента без существенных систематических ошибок.

4. Чувствительность критерия χ^2 при малой статистике

До сих пор мы интересовались, насколько соответствует распределение по новым переменным χ^2 стандартному распределению при использовании правильного значения p. С другой стороны, при проверке согласия экспериментальных данных теоретической модели не менее важен вопрос, насколько чувствителен данный критерий к неправильному значению p. К сожалению, ответ на такой вопрос неоднозначен и сильно зависит от того, сколько степеней свободы в распределении χ^2 и в какой области значений p проводятся исследования.

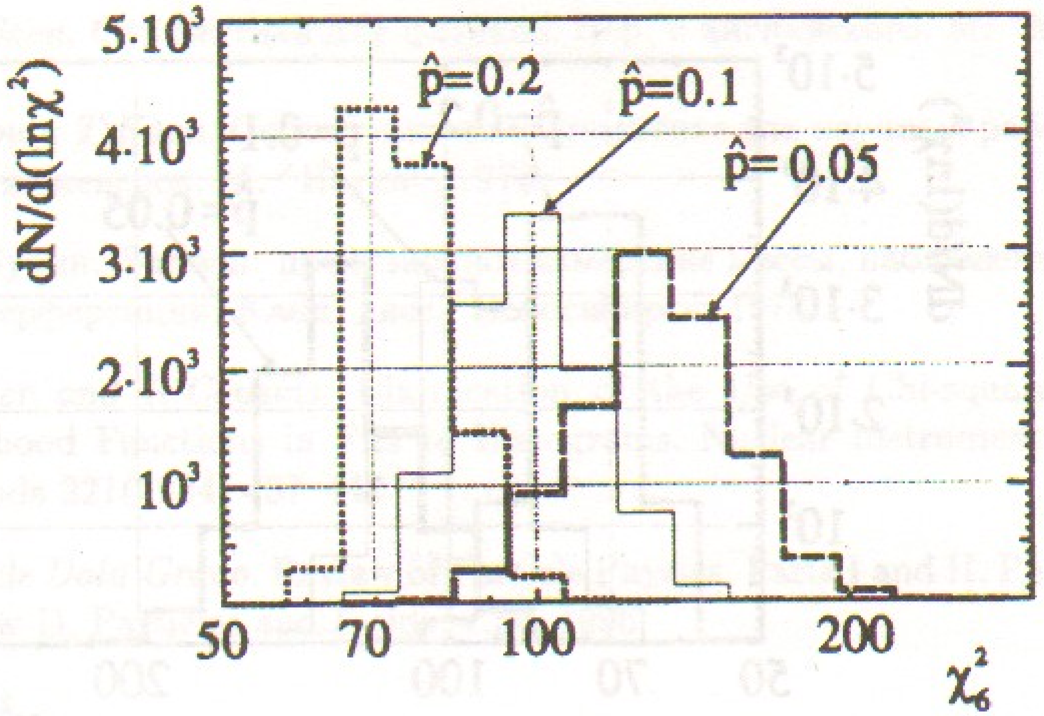

Рис. 16: Чувствительность критерия χ^2 при количестве степеней свободы $n_D = 100$ для варианта вычисления $\chi^2 = \chi_6^2$ при малом параметре $p = 0.1$ пуассонова распределения

Всё же можно иллюстрировать чувствительность распределения по χ^2 в нескольких случаях при $p << 1$. Возьмём случай $n_D = 100$ (число степеней свободы). На рис. 16 сплошной кривой изображено стандартное распределение по χ^2 при таком числе степеней свободы, закрашенной гистограммой — распределение по χ_6^2 при правильном значении $p = 0.1$ и две гистограммы распределения по χ_6^2 при существенно неправильных оценках $p = 0.05$ и $p = 0.2$ (при генерации событий методом Монте-Карло всегда использовалось значение $p = 0.1$).

На следующем рисунке 17 так же демонстрируется чувствительность параметра χ_7^2.

5. Заключение

В работе проверены пределы применимости критерия χ^2 согласия экспериментальных данных с теорией для нескольких вариантов вычисления величины χ^2 при математическом ожидании p числа экспериментальных событий порядка или меньше единицы, где распределение вероятностей Пуассона существенно отличается от распределения Гаусса, кото-

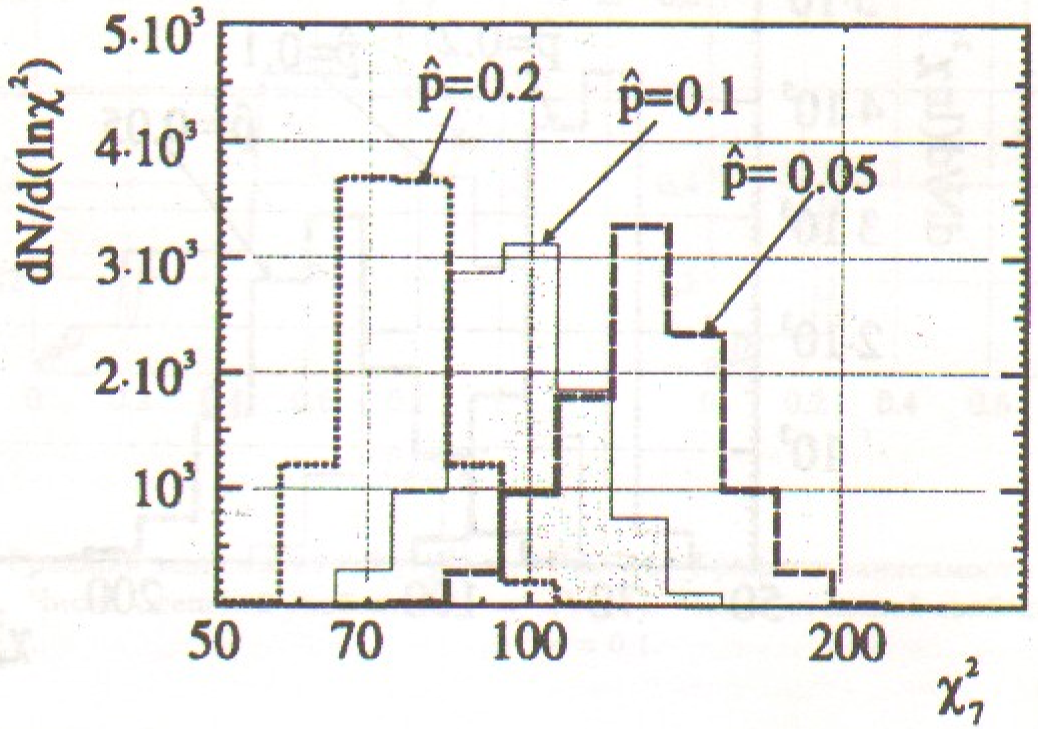

Рис. 17: Чувствительность критерия χ^2 при количестве степеней свободы $n_D = 100$ для варианта вычисления $\chi^2 = \chi^2_7$ при малом параметре $p = 0.1$ пуассонова распределения

рое, строго говоря, необходимо для правильного использования критерия χ^2.

В результате анализа получено, что уже при $p \sim 1$ предпочтительнее использовать предложенные в данной работе варианты модификации χ^2_6 и χ^2_7, которые обеспечивают правильное значение среднего и дисперсии χ^2. Вычисленные таким образом значения χ^2 дают больше оснований использовать стандартные значения вероятностей $P(\chi^2)$ для заданного числа степеней свободы. Особенно просто выглядит выражение для χ^2_6, которое, кстати, обеспечивает правильные значения среднего и дисперсии при всех возможных значениях $0 < p < \infty$:

$$\chi^2_6 = \frac{(n-p)^2}{\sqrt{p(p+0.5)}} + \frac{1}{2p+1+\sqrt{2p(2p+1)}}$$

где n — количество экспериментальных событий, p — математическое ожидание числа событий, определяемое теорией.

Литература

[1] *Д.Худсон.* Статистика для физиков. Пер. с английского. М: "Мир", 1970.
Г.Корн и Т.Корн. Справочник по математике для научных работников и инженеров. М: "Наука", 1973.

[2] А.Д.Букин. ϕ-мезон: прецизионное измерение массы, наблюдение $\omega - \phi$ интерференции. Канд. дисс., Новосибирск, 1978

[3] *S.Baker and R.Cousins.* Clarification of the Use of Chi-square and Likelihood Functions in Fits to Histograms. Nuclear Instruments and Methods 221(1984) 437–442

[4] *Particle Data Group.* Review of Particle Physics. Parts I and II. Physical Review D, Particles and Fields, v.54, 1996.

А.Д. Букин

Критерий χ^2

для распределения Пуассона

A.D. Bukin

χ^2 criteria
for Poisson statistics

ИЯФ 97-104

Ответственный за выпуск А.М. Кудрявцев
Работа поступила 24.12. 1997 г.

Сдано в набор 25.12.1997 г.
Подписано в печать 25.12.1997 г.
Формат бумаги 60×90 1/16 Объем 1.1 печ.л., 0.9 уч.-изд.л.
Тираж 80 экз. Бесплатно. Заказ N° 104

Обработано на IBM PC и отпечатано на
ротапринте ИЯФ им. Г.И. Будкера СО РАН,
Новосибирск, 630090, пр. академика Лаврентьева, 11.