



Сибирское отделение Российской Академии наук
ИНСТИТУТ ЯДЕРНОЙ ФИЗИКИ им.Г.И. Будкера

*Б.90
1998*

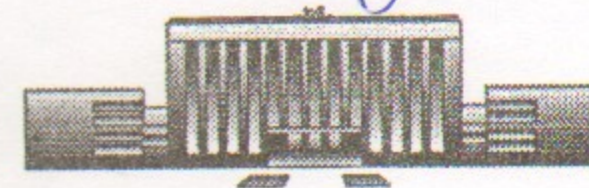
А.Д. Букин

КРИТЕРИЙ СОГЛАСИЯ КОЛМОГорова

ИЯФ 98-81

<http://www.inp.nsk.su/publications>

БИБЛИОТЕКА
Института ядерной
Физики СО АН СССР
ИНВ. № *2 мф*



НОВОСИБИРСК
1998

Критерий согласия Колмогорова

А.Д.Букин

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

Аннотация

В работе рассматривается известный критерий согласия Колмогорова, проблемы, возникающие при вычислении функции Колмогорова, поправки к асимптотической функции распределения для случая небольшого числа случайных величин. На ряде примеров проводится сравнение мощностей критериев Колмогорова и $P(\chi^2)$ Пирсона

Kolmogorov confidence level

A.Bukin

Budker Institute of Nuclear Physics,
630090, Novosibirsk, Russia

Abstract

In the paper the known Kolmogorov confidence level is considered, including the problems of Kolmogorov function evaluation, corrections for non-asymptotic case of low-number of random values. With a set of examples the comparison of power of the Kolmogorov criterion and $P(\chi^2)$ of Pearson is carried out.

©The Budker Institute of Nuclear Physics

В настоящее время становится популярным использование критерия Колмогорова в экспериментальной физике (возможно, благодаря включению этого критерия в набор команд системы PAW [1]). Для того, чтобы лучше представлять себе возможности этого критерия и область его применимости, а также сравнить его с широко используемым при анализе данных критерием χ^2 Пирсона [2], проведена настоящая работа.

1. Определение критерия Колмогорова

При обсуждении критерия Колмогорова для удобства будем проводить параллели с широко используемым критерием χ^2 Пирсона [2]. По определению, критерий Колмогорова применим к выборке n измерений x_i одной и той же случайной величины, распределённой в соответствии с непрерывным распределением вероятностей $f(x)$. Уже здесь можно увидеть общие черты с χ^2 и различие: критерий χ^2 тоже применим к n измерениям одной и той же случайной величины с непрерывным распределением вероятностей (число степеней свободы n), правда, распределение это должно быть нормальным (Гауссовым), а не произвольным. Для использования критерия Колмогорова для сравнения экспериментального распределения величины x_i и теоретического $f(x)$ вычисляются две интегральные функции распределения: теоретическая

$$F(x) = \int_{-\infty}^x f(\xi) d\xi \quad (1)$$

и выборочная

$$F_n(x) = k/n, \quad x_k < x < x_{k+1}, \quad (2)$$

где массив зарегистрированных значений x_i упорядочен по возрастанию, и принято, что $x_0 = -\infty$, $x_{n+1} = +\infty$. Максимальное отклонение тестируемой функции распределения от выборочной и используется в критерии Колмогорова [3,4]:

$$\lambda = \max_x |F_n(x) - F(x)| \quad (3)$$

Для фиксированного n область определения величины λ есть интервал $(\frac{1}{2n}, 1)$. В тех случаях, когда выборка x_i получена для событий, соответствующих правильной функции распределения $f(x)$, вероятность

$$W \left(\max_x |F_n(x) - F(x)| < \lambda \right) = K_n(\lambda) \quad (4)$$

является универсальной функцией, не зависящей от конкретной функции распределения $f(x)$. И в этом есть определённое сходство с критерием χ^2 . Правда, эту функцию $K_n(\lambda)$ крайне трудно вычислить, в отличие от $P(\chi^2)$.

Нетривиальный результат, полученный Колмогоровым, заключается в выводе асимптотического поведения функции распределения при больших n :

$$\lim_{n \rightarrow \infty} W\left(\sqrt{n} \times \max_x |F_n(x) - F(x)| < \lambda\right) = \lim_{n \rightarrow \infty} K_n(\lambda/\sqrt{n}) =$$

$$= K(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k \exp(-2k^2 \lambda^2) \quad (5)$$

Для конечного числа n , конечно, правильнее было бы пользоваться точной вероятностью $K_n(\lambda)$.

2. Вычисление функции $K(\lambda)$

В формуле (5) приведено выражение для функции $K(\lambda)$ в виде ряда. При не слишком близких к нулю значениях λ этот ряд быстро сходится, но даже и в этом случае его лучше переписать в несколько другом виде:

$$K(\lambda) = 1 + 2 \sum_{k=1}^{\infty} (-1)^k \exp(-2k^2 \lambda^2) \quad (6)$$

Однако, при стремлении $\lambda \rightarrow 0$ значение суммы ряда по этой формуле трудно вычислять, а при $\lambda = 0$ сумма ряда вообще становится неопределённой, так как общий член бесконечного ряда стремится к ± 1 . Воспользовавшись тождеством

$$\sum_{m=0}^{\infty} \left[e^{-2(2m)^2 \lambda^2} - e^{-2(2m+2)^2 \lambda^2} \right] = 1 \quad (7)$$

можно преобразовать выражение (6) к бесконечному ряду, общий член которого стремится к нулю при $\lambda \rightarrow 0$:

$$K(\lambda) = \sum_{m=0}^{\infty} \left[e^{-2(2m)^2 \lambda^2} + e^{-2(2m+2)^2 \lambda^2} - 2e^{-2(2m+1)^2 \lambda^2} \right] \quad (8)$$

Всё же и такое выражение неудобно для суммирования, так как при $\lambda \sim 0$ общий член ряда медленно меняется в интервале суммирования m до $1/\lambda$, к тому же меняя знак при $m \sim \frac{1}{4\lambda}$.

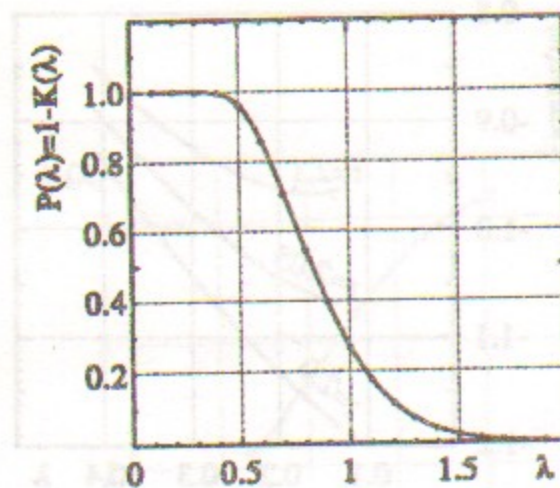


Рис. 1: График функции $P(\lambda) = 1 - K(\lambda)$

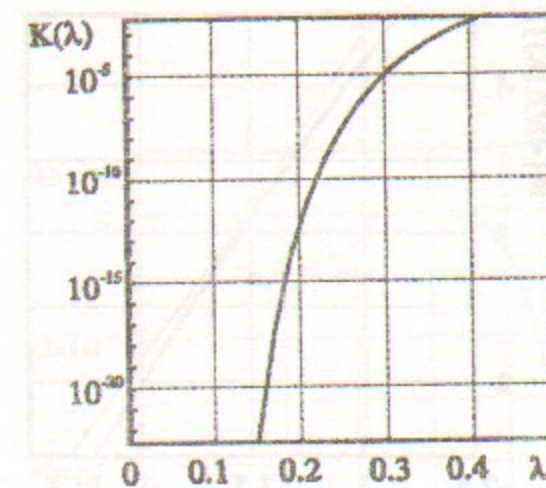


Рис. 2: График функции $K(\lambda) = 1 - P(\lambda)$ при малых λ

На рис.1 изображён график функции $P(\lambda) = 1 - K(\lambda)$ (вероятность получить отклонение интегральной функции больше λ , как раз то, что обычно приводится в таблицах). При сравнении значений функции, рассчитанных по формуле (6), с пятизначными таблицами в [4], не было замечено отличий, начиная с $\lambda = 0.3$. При больших значениях λ функция $P(\lambda)$ хорошо аппроксимируется первым членом ряда (6):

$$P(\lambda) = 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp(-2k^2 \lambda^2) \approx 2 \exp(-2\lambda^2), \quad \lambda \rightarrow \infty \quad (9)$$

Для меньших значений λ в указанных таблицах в [4] нет значений функции, так как $K(\lambda < 0.3) < 10^{-6}$. График функции $K(\lambda)$ при малых λ приведён на рис.2. Видно, что поведение функции не похоже на поведение аналитической функции (очень быстрое падение достаточно далеко от $\lambda = 0$). Хотя в литературе по статистике функция $K(\lambda)$ называется функцией Колмогорова, но специальная функция такого вида была введена задолго до Колмогорова (эллиптические θ -функции). Однако, удобного представления $K(\lambda)$ вблизи нуля не удалось найти в справочниках по специальным функциям. Попробуем получить достаточно точную аппроксимацию $K(\lambda)$ при $\lambda \leq 0.1$ с тем, чтобы не суммировать ряд даже при $\lambda \sim 0.1$, не говоря о меньших значениях.

Для начала посмотрим, не поможет ли сгладить функцию в районе нуля множитель вида $\exp(-A/\lambda^r)$. Для этого построим график зависимости $\ln\{-\ln[K(\lambda)]\}$ от $\ln \lambda$ (рис.3). Пунктирными линиями обозначено примерное направление касательных к графику функции в крайних точ-

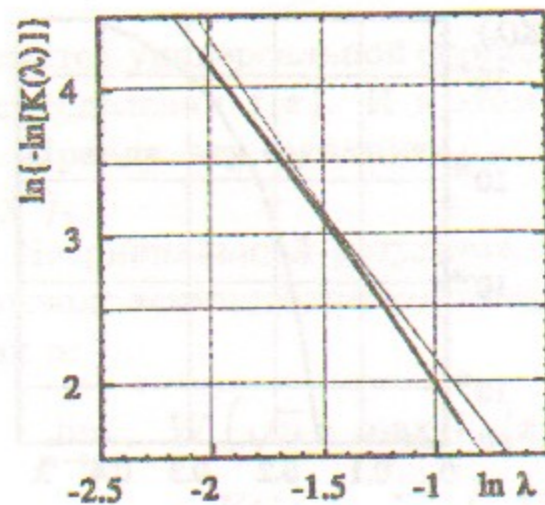


Рис. 3: Зависимость $\ln\{-\ln[K(\lambda)]\}$ от $\ln \lambda$

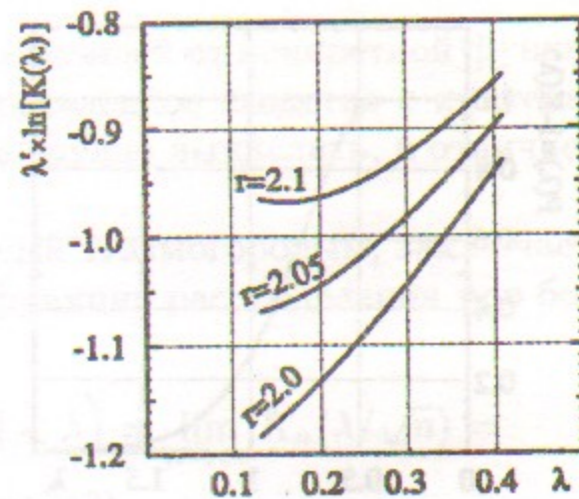


Рис. 4: График функции $\lambda^2 \times \ln[K(\lambda)]$ при малых λ

ках. При этом тангенс угла наклона касательной меняется от $k \approx -2.4$ до $k \approx -2.0$ при уменьшении λ . Это соответствует $r \approx 2$, т.е.

$$K(\lambda) \sim \exp(-A/\lambda^2) \quad (10)$$

Учтём эту зависимость и рассмотрим график функции $\lambda^r \times \ln[K(\lambda)]$ (рис.4) при нескольких значениях r около 2. При $r \approx 2.05$ график функции при $\lambda \rightarrow 0$ приближается к горизонтальной линии на уровне ~ -1.08 . Окончательно, попробуем представить функцию $K(\lambda)$ в виде:

$$K(\lambda) = \exp\left(-\frac{1.08}{\lambda^{2.05}}\right) \times \left[1 + e^{-2\lambda^2} R(\lambda)\right] \quad (11)$$

где функция $R(\lambda)$ — полином. График функции $R(\lambda)$

$$R(\lambda) = e^{2\lambda^2} \times \left[K(\lambda) \cdot \exp\left(\frac{1.08}{\lambda^{2.05}}\right) - 1\right] \quad (12)$$

представлен на рис.5. Гладкой кривой изображён полином третьей степени

$$R(\lambda) \approx a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 \quad (13)$$

$$a_0 = 0.668838, a_1 = -2.97124, a_2 = 58.5242, a_3 = -89.3832$$

Полученную таким образом аппроксимацию функции $K(\lambda)$ можно анализировать на абсолютное (рис.6,7) и относительное (рис.8) отклонение от точных значений. Эти графики показывают, что в интервале зна-

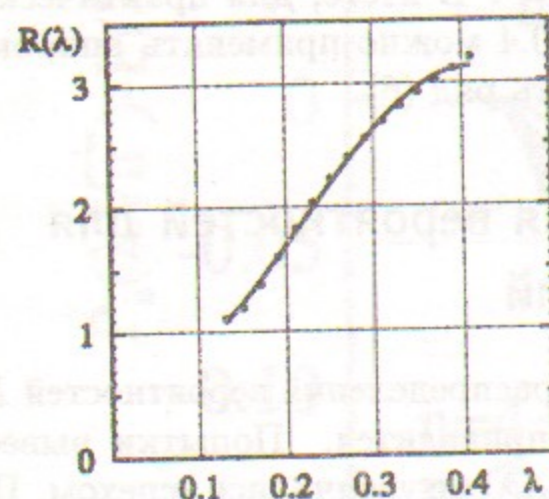


Рис. 5: Аппроксимация функции $R(\lambda)$ из уравнения (12) полиномом третьей степени.

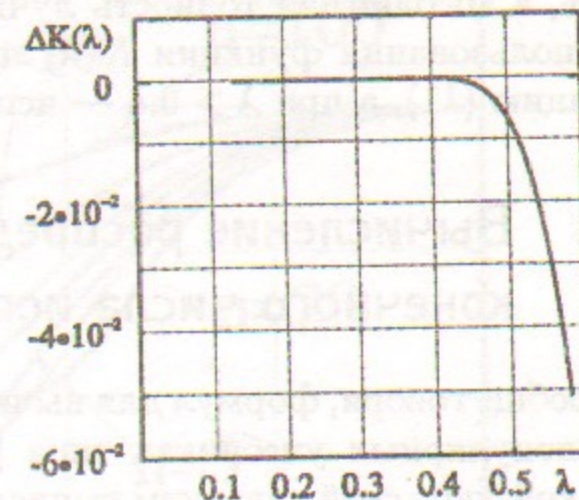


Рис. 6: Отклонение аппроксимирующей функции (11) от точного значения функции $K(\lambda)$

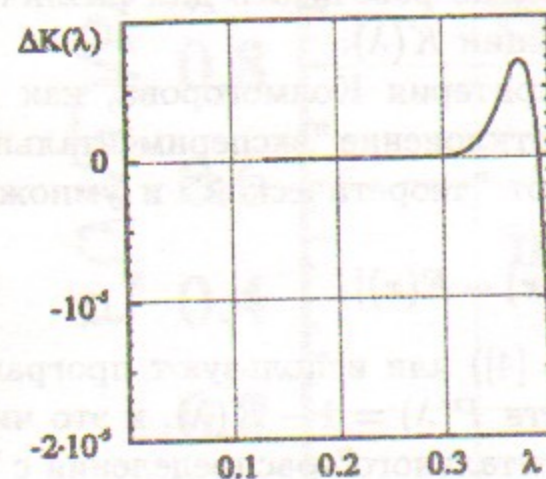


Рис. 7: Отклонение аппроксимирующей функции (11) от точного значения функции $K(\lambda)$

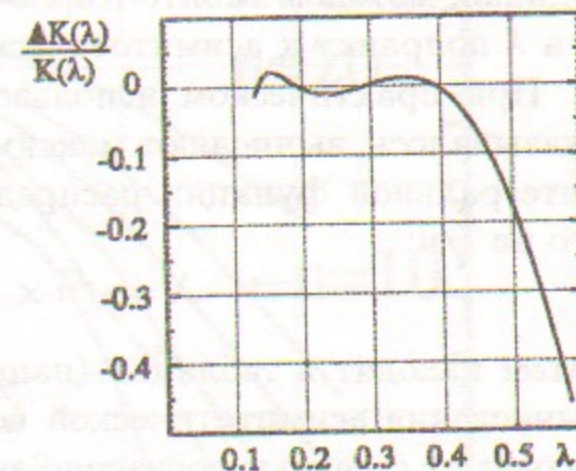


Рис. 8: Отклонение аппроксимирующей функции (11) от точного значения $K(\lambda)$, нормированное на $K(\lambda)$

чений λ от 0.13 до 0.4 относительная точность аппроксимации не хуже 2%, а абсолютная точность лучше $2 \cdot 10^{-5}$. В итоге, для практического использования функции $K(\lambda)$ при $\lambda < 0.4$ можно применять аппроксимацию (11), а при $\lambda > 0.4$ — использовать ряд (6).

3. Вычисление распределения вероятностей для конечного числа испытаний

Вообще говоря, формул для вычисления распределения вероятностей $K_n(\lambda)$ в популярных учебниках типа [3,4] не приводится. Попытки вывести какой-либо свой алгоритм вычисления $K_n(\lambda)$ не увенчались успехом. Правда, легко получаются формулы на краях области определения λ , однако, практического применения для них найти трудно. Имеются в виду такие две формулы. Пусть переменная λ определена как

$$\lambda = \max_x |F_n(x) - F(x)| \quad (14)$$

Тогда

$$K_n(\lambda) = n! \cdot (2\lambda - \frac{1}{n})^n, \quad \text{при } \frac{1}{2n} \leq \lambda \leq \frac{1}{n} \quad (15)$$

$$K_n(\lambda) = 1 - 2(1 - \lambda)^n, \quad \text{при } 1 - \frac{1}{2n} \leq \lambda \leq 1$$

Однако, методом Монте-Карло можно легко рассчитать для различных n и λ поправку к асимптотической функции $K(\lambda)$.

При практическом использовании критерия Колмогорова, как уже указывалось, вычисляют максимальное отклонение "экспериментальной" интегральной функции распределения от "теоретической" и умножают его на \sqrt{n} :

$$\lambda = \sqrt{n} \times \max_x |F_n(x) - F(x)|, \quad (16)$$

затем находят в таблицах (например, в [4]) или используют программу вычисления асимптотической вероятности $P(\lambda) = 1 - K(\lambda)$, и это число называют степенью согласия "экспериментального" распределения с "теоретическим". На рис.9 приведена поправка к асимптотической функции $P(\lambda)$, которая позволяет получить значения $P_n(\lambda/\sqrt{n})$ для некоторых n .

Можно заметить, что сходимость к асимптотической функции распределения наблюдается, хотя и довольно медленная. Наиболее интересная область $P(\lambda) \sim 0$ на этом графике неразличима. На другом рис.10 поправка изображена в виде множителя к функции $P(\lambda)$, что позволяет лучше представить область $\lambda > 1$.

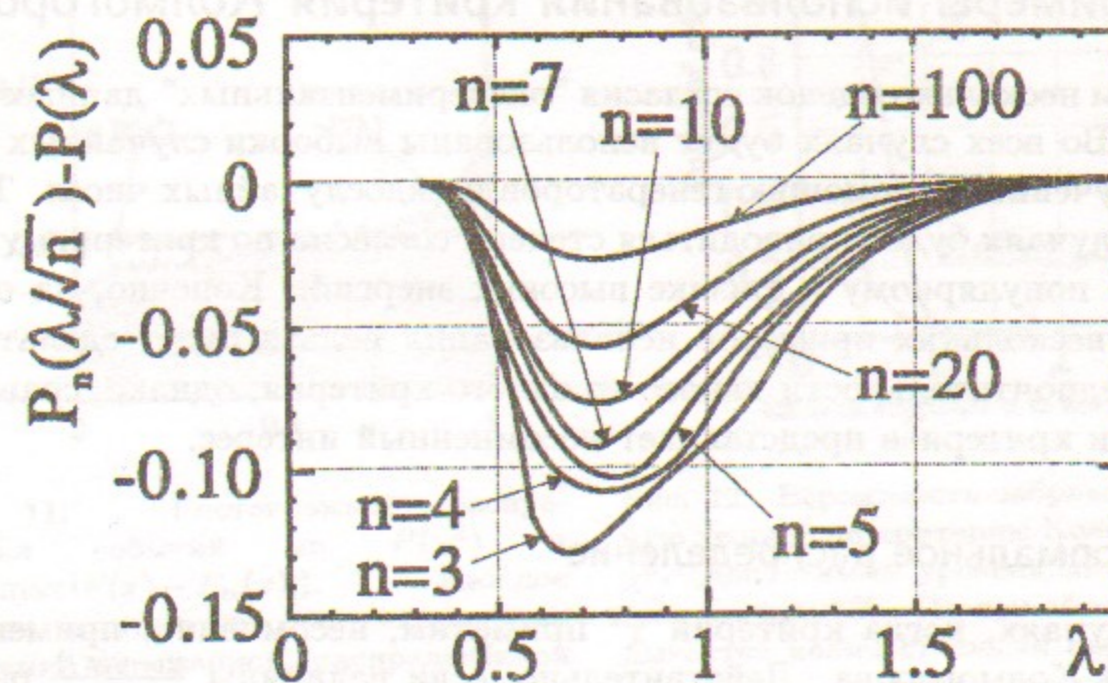


Рис. 9: Поправка $\Delta P_n = P_n(\lambda/\sqrt{n}) - P(\lambda)$ к асимптотической функции Колмогорова, рассчитанная методом Монте-Карло (точность расчёта лучше 10^{-3}).

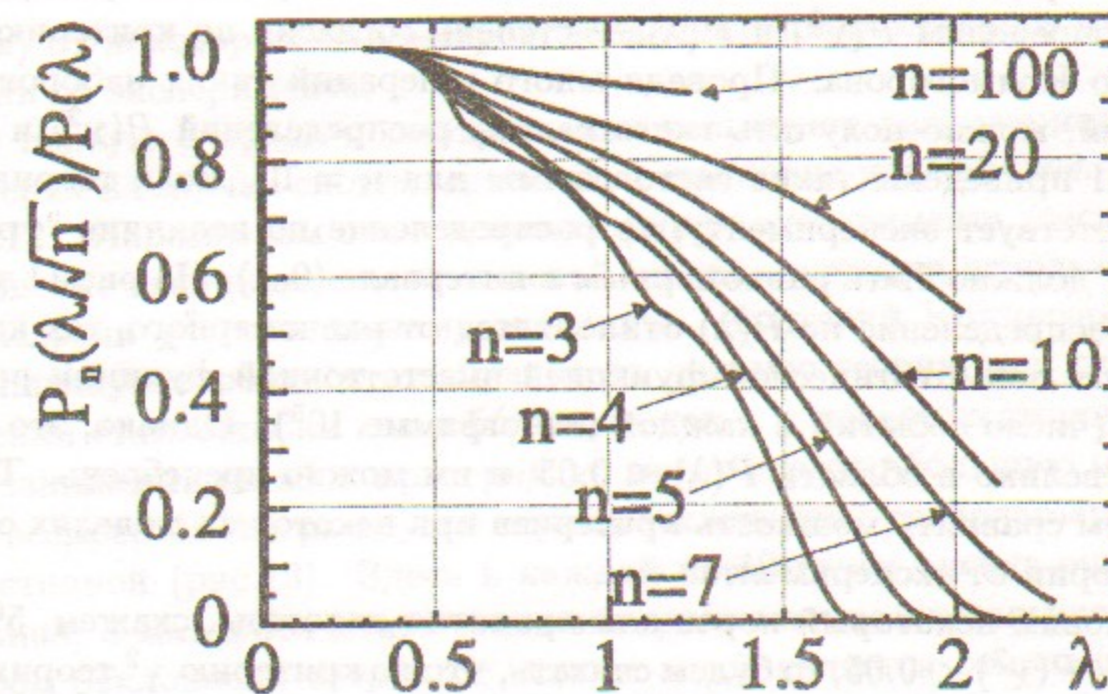


Рис. 10: Поправка $P_n(\lambda/\sqrt{n})/P(\lambda)$ к асимптотической функции Колмогорова, рассчитанная методом Монте-Карло (точность расчёта лучше 10^{-3}).

4. Примеры использования критерия Колмогорова

Проведём несколько оценок согласия "экспериментальных" данных с "теорией". Во всех случаях будут использованы выборки случайных величин, полученные с помощью генераторов псевдослучайных чисел. Также во всех случаях будет приводиться степень согласия по критерию χ^2 , как наиболее популярному в физике высоких энергий. Конечно, на основе анализа нескольких примеров использования нельзя будет сделать вывод о предпочтительности какого-то одного критерия, однако, сравнение мощностей критериев представляет несомненный интерес.

4.1. Нормальное распределение

В тех случаях, когда критерий χ^2 применим, несомненно, применим и критерий Колмогорова. Действительно, если величины $\frac{y_i - y(x_i)}{\sigma}$ распределены по гауссову закону, то можно нормированные отклонения эксперимента от теории рассматривать, как независимые испытания одной и той же случайной величины, и применять критерий Колмогорова для степени согласия с нормальным распределением.

Возьмём n независимых измерений одной и той же случайной величины, распределённой в точном соответствии с гауссовым распределением, и определим $P(\chi^2)$ и $P(\lambda)$ — степень согласия по критерию χ^2 и критерию Колмогорова. Проведя много генераций таких наборов по n измерений, можно получить гистограммы распределений $P(\chi^2)$ и $P(\lambda)$. На рис.11 приведены такие гистограммы для $n = 10$. Если теория точно соответствует эксперименту, то распределение по величине "уровень доверия" должно быть равномерным в интервале $(0, 1)$. На рис.11 заметно, что распределение по $P(\lambda)$ отклоняется от равномерного, так как мы пользуемся асимптотической функцией вместо точной функции распределения (число событий в каждой гистограмме 10^5). Однако, это искажение невелико в области $P(\lambda) \leq 0.05$, и им можно пренебречь. Теперь попробуем сравнить мощность критериев при некоторых моделях отклонения теории от эксперимента.

Установим некоторый порог для принятия гипотезы, скажем, 5%. То есть, если $P(\chi^2) < 0.05$, то будем считать, что по критерию χ^2 теория принимается ложной. Конечно, в 5% случаев мы будем браковать истинную теорию. Однако, если критерий способен правильно отличать истинную теорию от ложной, то для ложной теории мы будем получать результат $P(\chi^2) < 0.05$ гораздо чаще, чем в 5% случаев. Построим теперь графики зависимости вероятности забраковать ложную теорию (мощность кри-

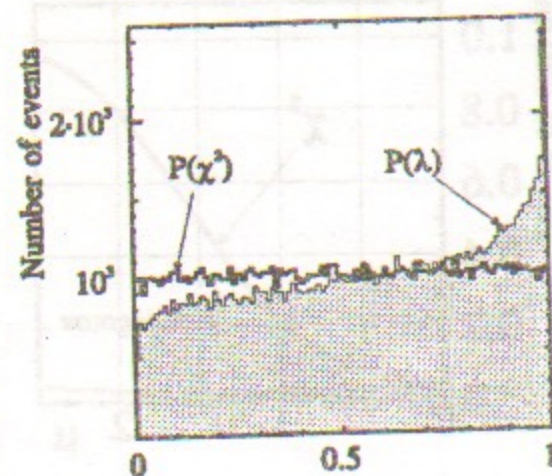


Рис. 11: Гистограммы распределений событий по $P(\chi^2)$ и $P(\sqrt{n} \max |F(x) - F_n(x)|)$. Каждое событие состоит из 10 независимых испытаний переменной, распределённой по нормальному закону. Полное число событий в каждой гистограмме 10^5

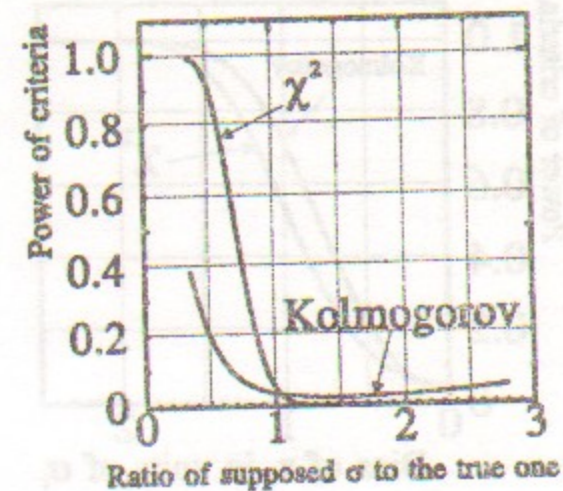


Рис. 12: Вероятность забраковать ложную теорию по критерию Колмогорова и χ^2 . Допустимый уровень согласия принят равным 5%. По оси абсцисс откладывается количественный параметр нарушения — отношение средне-кв. отклонения в каждой из 10 "экспериментальных" точек к его истинному значению (во всех 10 точках одновременно занижается или завышается ошибка)

терия) от некоторых параметров, задающих количественное отклонение теории от эксперимента.

В первую очередь посмотрим, как реагируют рассматриваемые два критерия на случай систематического завышения или занижения ошибок (рис.12). Параметром отклонения теории от эксперимента здесь является отношение величины статистической ошибки, которая использована при вычислении χ^2 , к истинному значению. Критерий Колмогорова имеет повышенную вероятность забраковать ложную теорию при сильном искажении стат. ошибки как в большую, так и в меньшую сторону, однако при занижении ошибок критерий χ^2 имеет гораздо большую мощность.

Рассмотрим теперь другой случай отклонения предполагаемой теории от истинной (рис.13). Здесь в каждой из 10 точек оценка ошибки правильная, а математическое ожидание систематически завышено. Параметром отклонения теории от истины здесь является отношение смещения среднего к средне-кв. отклонению. Для данного типа отклонения гипотезы от истины критерий Колмогорова оказывается более мощным.

На рис.14 средние значения в 10 точках измерения лежат на наклонной прямой, а мы будем проверять гипотезу, что все 10 точек имеют

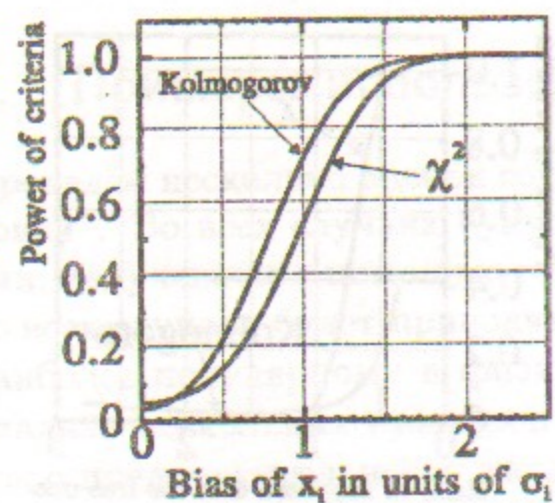


Рис. 13: Мощность критериев χ^2 и Колмогорова в случае, когда параметром отклонения теории от истины является смещение математического ожидания (в единицах ср.-кв. откл.)

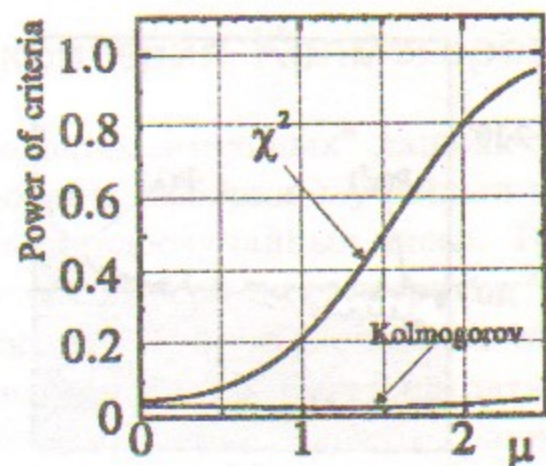


Рис. 14: Мощность критериев Колмогорова и χ^2 в случае, когда математическое ожидание для каждой из 10 точек измерения зависит от номера точки i : $\bar{x}_i = (i - 5.5)/4.5 \cdot \sigma \cdot \mu$, μ — максимальное отклонение в крайних точках в ед. σ . Проверяемая "ложная" теория предполагает, что математическое ожидание в каждой точке равно нулю.

среднее значение, равное нулю. В этом случае критерий χ^2 оказался намного предпочтительнее критерия Колмогорова.

Рис. 15 демонстрирует мощность критериев χ^2 и Колмогорова при проверке следующей "ложной" гипотезы: предполагается, что все 10 точек измерения имеют гауссово распределение со ср.-кв. отклонением σ относительно среднего значения, равного нулю, в то время как на самом деле среднее значение зависит от номера точки i следующим образом: $\bar{x}_i = \mu \cdot \sigma \sin(2i)$, где μ — безразмерный параметр "ложности" гипотезы.

Наконец, рассмотрим пример, очень невыгодный для критерия χ^2 , где модель отклонения гипотезы от истины выглядит несколько искусственно (рис. 16). Пусть наша гипотеза состоит в том, что все 10 измерений имеют среднее 0 и ср.-кв. отклонение σ , в то время как истинное распределение первой точки имеет среднее значение $\bar{x}_1 = -\sigma\mu/\sqrt{1+\mu^2}$, для оставшихся девяти точек $\bar{x}_i = \sigma\mu/\sqrt{1+\mu^2}$, а истинное ср.-кв. отклонение от среднего для всех точек $\sigma_i = \sigma/\sqrt{1+\mu^2}$. Здесь μ — некоторый параметр модели отклонения гипотезы от истины. Как и ожидалось, в этом случае критерий Колмогорова обладает большой мощностью, в то время как критерий χ^2 тем меньше бракует ложную гипотезу, чем она дальше от истины.

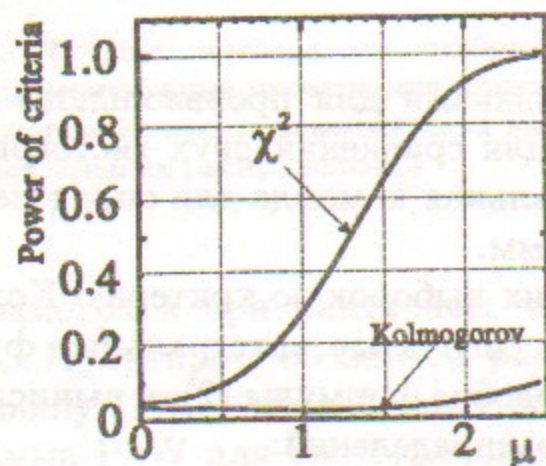


Рис. 15: Мощность критериев Колмогорова и χ^2 в случае, когда математическое ожидание для каждой из 10 точек измерения зависит от номера точки i : $\bar{x}_i = \sin(2i) \cdot \sigma \cdot \mu$, μ — параметр отклонения "ложной" гипотезы от истины. Проверяемая "ложная" теория предполагает, что математическое ожидание в каждой точке равно нулю.

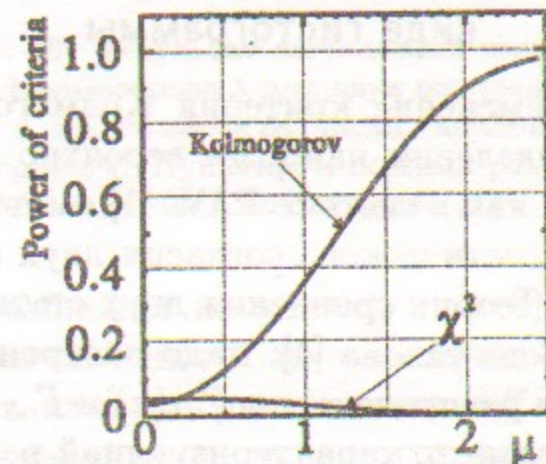


Рис. 16: Мощность критериев Колмогорова и χ^2 в случае, когда истинное математическое ожидание для первой точки измерения $\bar{x}_1 = -\sigma\mu/\sqrt{1+\mu^2}$, для оставшихся девяти точек $\bar{x}_i = \sigma\mu/\sqrt{1+\mu^2}$, истинное ср.-кв. отклонение $\sigma_i = \sigma/\sqrt{1+\mu^2}$. μ — некоторый параметр отклонения гипотезы от истины. Проверяемая "ложная" теория предполагает, что математическое ожидание в каждой точке равно нулю, а ср.-кв. откл. равно σ .

В результате анализа перечисленных примеров применения критерия Колмогорова к нормальному распределению можно сделать вывод, что мощность критерия Колмогорова вполне сравнима с мощностью критерия χ^2 , а в некоторых случаях заметно превышает её. Поэтому вполне оправдано во всех случаях, где это допустимо по формальным соображениям, приводить уровни согласия теории с экспериментом по обоим критериям.

4.2. Использование распределения вероятностей, полученного в виде гистограммы

Применение критерия Колмогорова физиками для произвольного распределения наиболее вероятно именно для сравнения двух гистограмм, так как в системе PAW [1] имеется специальная команда для определения статистического согласия двух гистограмм.

Теория сравнения двух статистических выборок по критерию Колмогорова такова [4]: надо построить две ступенчатые интегральные функции распределения $F_{n1}(x)$ и $F_{n2}(x)$, наподобие формулы (2), и вычислить параметр, характеризующий различие распределений:

$$\lambda = \sqrt{\frac{N_1 N_2}{N_1 + N_2}} \times \max_{-\infty < x < \infty} |F_{n1}(x) - F_{n2}(x)|, \quad (17)$$

где N_1, N_2 — полное число событий в первой и второй выборке, соответственно. При достаточно больших числах N_1, N_2 распределение вероятностей по этому параметру λ близко к асимптотическому распределению Колмогорова, и для степени согласия двух выборок можно пользоваться функцией $P(\lambda)$ в формуле (9).

Очевидно, что если мы пользуемся для определения параметра λ гистограммой, то мы вносим некоторую погрешность. Рассмотрим величину этого эффекта на каком-либо примере. Возьмём две выборки с одинаковым количеством событий, распределённых на интервале (0, 1) по немного отличающимся распределениям вероятностей:

$$\frac{dW_1}{dx} \sim \frac{1}{1+x^2}, \quad \frac{dW_2}{dx} \sim \frac{1}{1+x^2} + \frac{\beta}{1+(x-0.5)^2} \quad (18)$$

Выберем $\beta = 2$, количество каналов гистограммы $n = 50$, полную статистику событий $N_1 = N_2 = 1000$. Полученные гистограммы приведены на рис.17. Согласуются ли они статистически с одним и тем же распределением вероятностей? На рис.18 приведены результаты вычисления

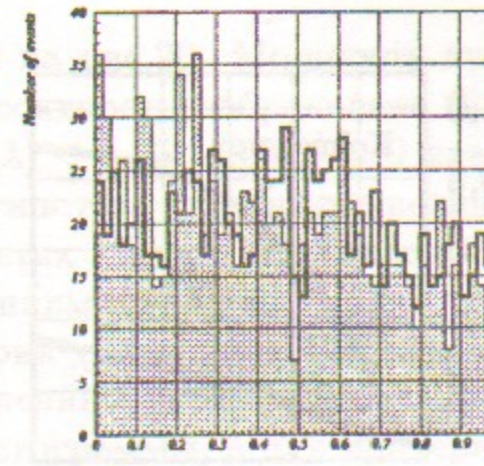


Рис. 17: Две гистограммы, выбранные для демонстрации применения критерия Колмогорова при сравнении двух экспериментальных распределений

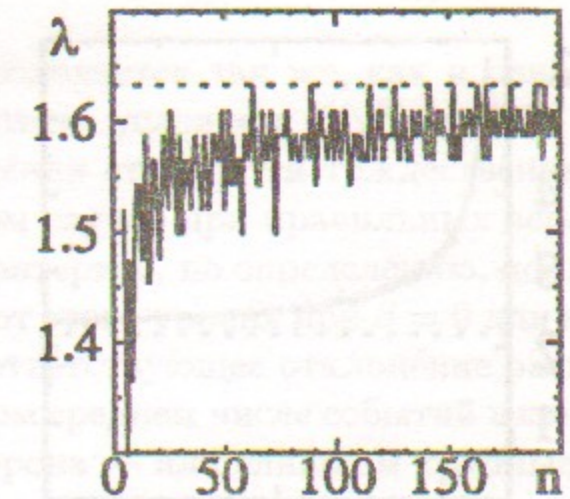


Рис. 18: Результат вычисления параметра Колмогорова λ для двух распределений в зависимости от числа каналов гистограммы. Пунктиром показан результат вычисления λ по формуле (17)

параметра λ при различном числе n разбиения на каналы гистограммы. Пунктиром обозначен результат вычисления λ по каноническому правилу (17) — для бесконечного числа каналов (по команде *diff* программа PAW для гистограмм на рис.17 вычислила степень согласия по критерию Колмогорова $P(\lambda) = 1.96 \cdot 10^{-2}$, $\max_x |F_{n1}(x) - F_{n2}(x)| = 0.068$, $\lambda = 0.068 \cdot \sqrt{1000/2} = 1.52$, что согласуется с нашими расчётами). Видно, что при вычитании двух ступенчатых функций распределения эффект "укрупнения ступенек" при объединении событий в более крупные каналы гистограммы не очень велик, но всё же надо иметь в виду, что правильнее сравнивать гистограммы с более мелким шагом.

При сравнении экспериментального распределения с некоторой функцией распределения $F(x)$ погрешность при увеличении шага гистограммы может быть больше, чем при сравнении двух гистограмм. Проверим это на примере сравнения гистограммы, полученной по распределению вероятностей dW_2/dx из (18), с распределением вероятностей dW_1/dx . Точное значение параметра Колмогорова в этом случае

$$\lambda = \sqrt{N} \max_x |F_n(x) - F(x)| = \sqrt{1000} \cdot 6.87 \cdot 10^{-2} = 2.17 \quad (19)$$

и уровень согласия получается $P(\lambda) = 1.6 \cdot 10^{-4}$. Результаты вычисления λ непосредственно по гистограмме для разного количества каналов приведены на рис. 19.

Интересно отметить, что в случае недостаточно мелкого шага гистограммы при сравнении двух гистограмм по критерию Колмогорова мож-

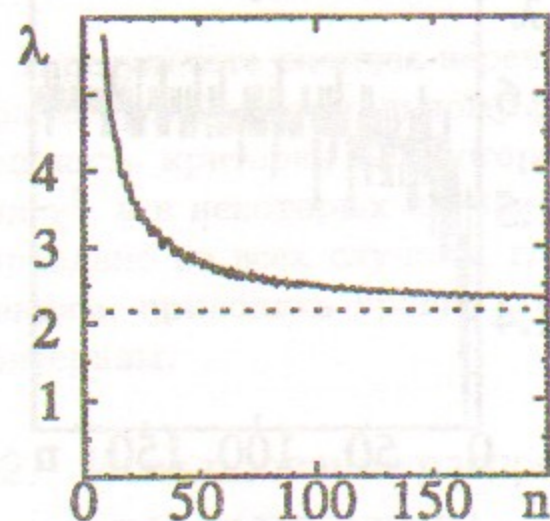


Рис. 19: Результат вычисления параметра Колмогорова λ для сравнения гистограммы, полученной в соответствии с dW_2/dx из (18), с функцией распределения dW_1/dx

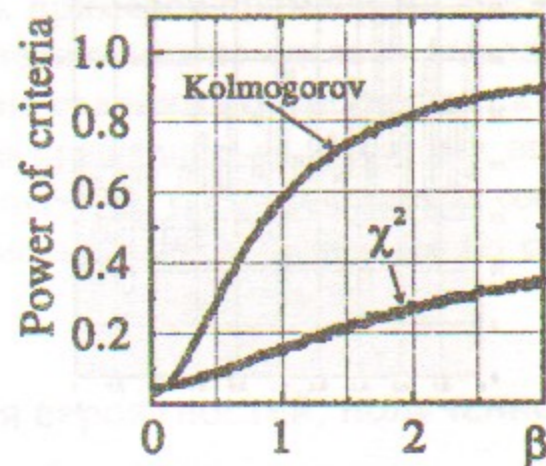


Рис. 20: Мощность критерия при сравнении двух гистограмм распределений (18) при $n = 50$ по критериям Колмогорова и χ^2 Пирсона в зависимости от параметра β

но завязать степень согласия, а при сравнении гистограммы с теоретической функцией распределения вероятностей — сильно занизить.

Как мы видели, даже при достаточно большом числе событий в каждом канале гистограммы можно применять критерий Колмогорова для проверки статистического согласия двух экспериментальных распределений. Но если типичное число событий в канале гистограммы порядка или больше 10, то для сравнения двух распределений можно использовать и критерий χ^2 , так как при большом числе каналов гистограммы биномиальное распределение по числу событий в канале можно заменить на Пуассоновое распределение, а оно, в свою очередь, при числе событий порядка или больше 10 переходит в гауссово, что является необходимым условием для применимости критерия χ^2 . Если обозначить числа событий в первой гистограмме через k_i , во второй гистограмме — m_i , то величина χ^2 может быть рассчитана по формуле:

$$\chi^2 = \sum_{i=1}^n \frac{\left(\frac{k_i}{N_1} - \frac{m_i}{N_2}\right)^2}{\frac{k_i}{N_1} + \frac{m_i}{N_2}} \quad (20)$$

Число степеней свободы следует уменьшить на единицу $n_D = n - 1$ для учёта нормировки обоих распределений. Графики мощности критерия Колмогорова $P(\lambda)$ и критерия χ^2 $P(\chi^2)$ при сравнении двух гистограмм при $n = 50$, $N_1 = N_2 = 1000$ в зависимости от параметра β , количественно определяющего различие двух распределений в (18), приведе-

ны на рис.20. Мощность критерия определяется так же, как и ранее — вероятностью забраковать ложную гипотезу условием $P(\chi^2) < 0.05$ или $P(\lambda) < 0.05$. При $\beta = 0$ два распределения становятся тождественными и гипотеза становится истинной. В этом случае при правильных вероятностях $P(\chi^2)$ и $P(\lambda)$ мощность обоих критериев, по определению, должна принимать значение 0.05. Отклонение от этого уровня при $\beta = 0$ для критерия χ^2 означало бы, скорее всего, соответствующее отклонение распределения Пуассона от гауссова при данном среднем числе событий в канале гистограммы, а для критерия Колмогорова — или слишком крупные каналы гистограммы, или недостаточное приближение к асимптотике при полном числе событий $N = 1000$. Как показывает рисунок, критерий Колмогорова имеет большую мощность при сравнении двух гистограмм, а если учесть, что при малой статистике критерий χ^2 неприменим и по формальным соображениям, то можно ожидать, что для сравнения двух распределений всегда предпочтительнее критерий Колмогорова.

5. Описание подпрограмм, использованных для получения результатов

При изучении свойств критерия Колмогорова были использованы некоторые подпрограммы, написанные на языке Фортран-77, которые, вообще говоря, могут оказаться полезными в реальной работе при анализе статистического согласия выборки с гипотезой. Поэтому здесь приводится краткое описание вызова этих подпрограмм.

call FKOLM(Fk,Fp,x)

real *8 Fk, Fp, x

Вычисление функций Колмогорова $Fk = K(x)$ и $Fp = P(x) = 1 - K(x)$. При $x < 0.4$ используется функция FKOLM3, при $x \geq 0.4$ — функция FKOLM1 (см. ниже).

Fp = FKOLM1(x)

real *8 FKOLM1, Fp, x

Вычисление функции Колмогорова $P(\lambda)$ по формуле (9).

Fk = FKOLM2(x)

real *8 FKOLM2, Fk, x

Вычисление функции Колмогорова $K(\lambda)$ по формуле (8).

$F_k = FKOLM3(x)$
real *8 FKOLM3, Fk, x
Аппроксимация функции Колмогорова $K(\lambda)$ вблизи $\lambda \sim 0$ в соответствии с формулой (11).

$F_p = FK_GAUSS(n, g_arr)$
integer *4 n
real *8 FK_GAUSS, Fp, g_arr(n)
Оценка степени согласия по критерию Колмогорова n измерений случайной величины с нормальным распределением (гауссово распределение со средним значением, равным нулю, и дисперсией, равной 1). После вызова функции FK_GAUSS массив g_arr портится (элементы массива упорядочиваются по возрастанию).

$P_{xi} = PN_XI2(XI2, n)$
integer *4 n
real *8 PN_XI2, Pxi, XI2
Вспомогательная функция вычисления уровня доверия по критерию χ^2 для n степеней свободы.

$P_{xi1} = P1_XI2(XI2)$
real *8 P1_XI2, Pxi1, XI2
Подпрограмма к PN_XI2 — вычисление уровня доверия по критерию χ^2 для одной степени свободы.

$F_p = FP_HIST(n, kh1, kh2, dist)$
integer *4 n, kh1(n), kh2(n)
real *8 FP_HIST, Fp, dist
Вычисление уровня доверия по критерию Колмогорова, что две гистограммы с числом каналов n , заданные массивами чисел $kh1(n)$ и $kh2(n)$, соответствуют одной и той же функции распределения вероятностей. В переменной $dist$ при этом возвращается абсолютная величина максимального отличия двух интегральных функций распределения.

$P_{chi} = CHI2HI(n, kh1, kh2, chi2)$
integer *4 n, kh1(n), kh2(n)
real *8 CHI2HI, Pchi, chi2
Вычисление уровня доверия P_{chi} по критерию χ^2 Пирсона, что две гистограммы распределений $kh1(n)$ и $kh2(n)$ статистически согласуются с одним и тем же распределением вероятностей (предполагается, что количество событий в каждом канале гистограммы достаточно велико, чтобы можно было применять критерий χ^2). При вызове подпрограммы возвращается также величина $chi2 = \chi^2$, подсчитанная по формуле (20).

6. Заключение

В работе рассмотрен известный критерий согласия Колмогорова статистической выборки с гипотезой о распределении вероятностей. Проведена оценка отличия истинного распределения вероятности по параметру Колмогорова от асимптотической формулы, по которой рассчитаны таблицы степени согласия по критерию Колмогорова и которые используются на практике.

На нескольких практических примерах проведено сравнение мощности критериев Колмогорова и χ^2 . Несмотря на то, что нельзя однозначно отдать предпочтение одному из критериев, всё же рассмотренные примеры использования демонстрируют сравнимую мощность этих критериев. Как правило, критерий χ^2 имеет большую мощность. Применение критерия χ^2 ограничивается тем, что он применим только к величинам, имеющим гауссово распределение. Применение критерия Колмогорова сомнительно в тех случаях, когда производится оптимизация параметров теоретической модели по экспериментальным данным (для критерия χ^2 есть удобное правило уменьшения числа степеней свободы на количество оптимизируемых параметров). На практике удобно применять оба критерия для проверки согласия теории с экспериментом.

Автор благодарен Е.В.Пахтусовой за полезные обсуждения.

Литература

- [1] CERN Program Library Long Writeup Q121. PAW — Physics Analysis Workstation. An Introductory Tutorial. CERN, Geneva, 1995
- [2] K. Pearson. On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling. Phil. Mag. (5), 50 (1900) 157.
- [3] Maurice G. Kendall and Alan Stuart. The advanced theory of statistics. 4th ed., Macmillan, 1977
М.Кендалл, А.Стьюарт. Статистические выводы и связи. Пер. с англ. Л.И.Гальчука, А.Т.Терехина, под редакцией А.Н.Колмогорова. М.: "Наука", 1973.
- [4] А.К.Митропольский. Техника статистических вычислений. М.: "Наука", 1971

Оглавление

1	Определение критерия Колмогорова	3
2	Вычисление функции $K(\lambda)$	4
3	Вычисление распределения вероятностей для конечного числа испытаний	8
4	Примеры использования критерия Колмогорова	10
4.1	Нормальное распределение	10
4.2	Использование распределения вероятностей, полученного в виде гистограммы	14
5	Описание подпрограмм, использованных для получения результатов	17
6	Заключение	19

А.Д. Букин

Критерий согласия Колмогорова

A.D. Bukin

Kolmogorov confidence level

ИЯФ 98-81

Ответственный за выпуск А.М. Кудрявцев

Работа поступила 5.11. 1998 г.

Сдано в набор 6.11.1998 г.

Подписано в печать 6.11.1998 г.

Формат бумаги 60×90 1/16 Объем 1.5 печ.л., 1.2 уч.-изд.л.

Тираж 110 экз. Бесплатно. Заказ № 81

Обработано на IBM PC и отпечатано на

ротаприфте ИЯФ им. Г.И. Будкера СО РАН

Новосибирск, 630090, пр. академика Лаврентьева, 11.