15

Объединение выборок в анализе данных с интервальной неопределенностью

© А.Н. Баженов,[1,2] А.Ю. Тельнова[1]

[1] Физико-технический институт им. А.Ф. Иоффе РАН,
194021 Санкт-Петербург, Россия
[2] Санкт-Петербургский политехнический университет Петра Великого (СпбПУ),
Институт прикладной математики и механики,
194064 Санкт-Петербург, Россия
e-mail: bazhenov_an@spbstu.ru

Поступило в Редакцию 1 июля 2024 г.
В окончательной редакции 26 ноября 2024 г.
Принято к публикации 27 ноября 2024 г.

Приведен практический пример использования методов интервальной статистики в области ядерной физики. Показано, как эти методы позволяют объединить две выборки оптимальным образом. В качестве параметра коррекции одной из выборок является мультипликативная поправка на фон в одной из спектральных областей. Из физических соображений известна оценка величины фона. Рабочим инструментом при оптимизации является комбинированная мера совместности, обеспечивающая получение одновременно внешних и внутренних оценок. Оптимальность в данном случае подразумевает увеличение размера максимальной клики в интервальной выборке. Тем самым объем данных для расчета константы фундаментальной ядерной реакции увеличивается, что делает результат вычислений более надежным. По результатам объединения выборок рассчитано новое значение физической величины — циркулярной поляризации гамма квантов в реакции захвата поляризованного нейтрона протонами.

Ключевые слова: статус измерения, вполне совместные измерения, коэффициент Жаккара, мера совместности, интервальный анализ, интервальная статистика, накрывающие и ненакрывающие измерения и выборки, информационное множество, интервальная мода.

DOI: 10.61011/JTF.2025.04.60020.218-24

Введение

Публикация продолжает тематику применения интервального анализа и интервальной статистики для обработки физического эксперимента [1,2]. Фундаментально методологические вопросы интервальной статистики рассмотрены в книге [3].

Задача определения постоянной величины — базовая задача анализа данных. Под постоянной понимают величину, остающейся неизменной в процессе измерений. В эту категорию попадают и мировые физические постоянные, которые в современном естествознании полагаются неизменными во всей Вселенной как в пространстве, так и во времени. Таким образом, спектр применений простирается от самых рутинных ситуативных измерений до измерения фундаментальных характеристик материи.

Для оценивания постоянной величины используют различные меры положения и рассеяния. В теоретико-вероятностной математической статистике применяют как параметрические оценки на основе теоретических распределений, так и непараметрические оценки, например, медианные методы, оценки Тьюки, и др.

В случае интервальных данных разработаны свои специфические методы оценки выборки [3]. Далее мы кратко представим необходимые понятия анализа данных с интервальной неопределенностью. Затем будет рассмотрена задача определения одного из параметров базовой реакции ядерной физики, радиационного захвата нейтрона протоном $(np \to d\gamma)$. На основе опубликованных ранее данных предлагается способ расширения выборки данных с использованием конструкций [1,2].

1. Теоретические основы интервального анализа

1.1. Основные определения

Интервальный анализ — раздел математики, в котором объекты описываются интервалами. В настоящей работе используются определения, методология и терминология из [4] и современная система обозначений из [5]. Поскольку в интервальной арифметике содержится вещественная арифметика как предельная форма, требуется различать обозначения объектов разных видов арифметики при написании. В связи с этим все интервальные объекты — скаляры, векторы, матрицы — обозначаются жирным шрифтом [5].

Интервалом $[a, b]$ вещественной оси $\mathbf{R}$ называется множество всех чисел, расположенных между заданными числами a, b, включая a, b, т.е.:

$$[a, b] := \{x \in \mathbf{R} | a \le x \le b\}, \qquad (1)$$

при этом a и b называются концами интервала $[a, b]$.

Другие базовые определения интервального анализа содержатся в [1,2]. Капитальное изложение интервального анализа можно найти в книге С.П. Шарого „Конечномерный интервальный анализ“ [4].

Приведем необходимые сведения из относительно нового раздела анализа данных — интервальной статистики или анализа данных с интервальной неопределенностью [3]. Ключевым в данном подходе является понятия накрывающего измерения и накрывающей выборки. Накрывающим измерением назовем такое интервальное измерение, в котором содержится истинное значение измеряемой величины. Накрывающей выборкой будем считать выборку, в которой преобладают накрывающие измерения. На практике понятие накрытия не всегда можно проверить, поскольку истинное значение величины известно далеко не всегда. В практических целях требование накрытия заменяют на требование совместности выборки, которое и будет использоваться далее.

1.2. Информационное множество и статус измерений

При решении конкретных задач с выборками интервальных данных важнейшую роль имеет понятие информационного множества. Согласно [3], имеется в виду множество значений интересующих нас величин, которые совместны с данными измерений в рамках выбранной модели их обработки. Применительно к измерению постоянной величины совместность данных с моделью можно выразить в виде условий на отношение интервалов параметров модели $\mathbf{\Omega}$ и интервала конкретного измерения $\mathbf{x}_i$. Способы нахождения $\mathbf{\Omega}$ могут быть различны (мода, медиана), а в качестве отношения могут выступать пересечения или минимум по включению. Для различных измерений могут иметь место различные варианты отношений $\mathbf{\Omega}$ и $\mathbf{x}_i$.

В публикациях [4,5] были введены понятия, характеризующие влияние отдельных измерений на информационное множество. Мы будем использовать термины внутренние, внешние и граничные измерения. Согласно [4,5], измерение внутреннее, если оно не изменяет информационного множества. Внутренние измерения xi совместны с моделью $\mathbf{\Omega}$ и не содержат границ модели $\delta\mathbf{\Omega}$:

$$x_i \in \Omega, \quad \delta\Omega \notin x_i.$$

Таким образом, такие измерения „покрывают“ модель „с запасом“ и их влияние на модель практически нулевое. Граничные измерения содержат одну или более границ информационного множества. Именно эти измерения формируют информационное множество. Внешние измерения изменяют информационное множество.

Может создаться впечатление, что внутренние измерения вообще не значимы, что на наш взгляд не вполне справедливо. После рассмотрения практического примера мы дополнительно обсудим термины описания статуса измерений.

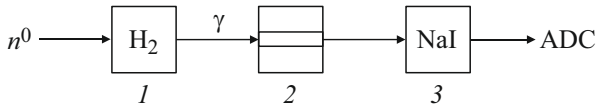


Рис. 1. Схема установки для измерения циркулярной поляризации гамма-квантов в реакции $np \to d\gamma$.

2. Практический пример

2.1. Исходные данные

Данные по циркулярной поляризации γ-квантов в фундаментальной ядерной реакции $np \to d\gamma$ с поляризованными нейтронами важны для ядерной физики и ядерного нуклеосинтеза. Наиболее точные измерения циркулярной поляризации, выполненные в лаборатории исследования слабых взаимодействий акад. В.М. Лобашева ПИЯФ РАН в 1988–1991 гг., приводятся в работе [6]. Подробности эксперимента содержатся в публикациях [7–10]. Конкретно в [7] представлена система управления спином нейтронов, в [8] описана твердая параводородная мишень и установка для производства чистого параводорода, в [9] — система регистрации эксперимента и в [10] — описание исследования влияния магнитного поля на детектор гамма-квантов.

На рис. 1 приведена упрощенная схема установки для измерения циркулярной поляризации гамма-квантов в реакции $np \to d\gamma$ [7–10]. Пучок холодных поляризованных нейтронов реактора ВВР-М ПИЯФ РАН подавался на твердую параводородную мишень *1*. Гамма-кванты из реакции $np \to d\gamma$ имели энергию 2.22 MeV. Для измерения поляризации использовался анализатор *2*, представлявший намагниченный ферромагнетик и имевший длину поглощения около 70 mm. Эффективность анализатора составляла примерно 5 %. Для регистрации гамма-квантов использовался сцинтилляционный детектор NaI(Tl) диаметром 150×100 mm. Система электронной регистрации ADC позволяла проводить измерения в трех диапазонах энергии: фотопике 2.22 MeV, в высокоэнергетической части комптоновского рассеяния и фоновой области с энергией выше 2.22 MeV.

Теоретические расчеты проводились с точки зрения ядерной физики [11], где в широком контексте обсуждается физика дейтрона, и [12], где даны подробности расчета циркулярной поляризации, а также с точки зрения нуклеосинтеза во время Большого Взрыва [13]. Более того, тонкие эффекты в структуре дейтрона могут повлиять на анализ прецизионных измерений A_γ-асимметрии при захвате неполяризованных нейтронов [14,15]. В этом эксперименте используется методики из работ [6,8,9], но с неполяризованными нейтронами.

Измерение круговой поляризации γ-квантов основано на зависимости их сечения комптоновского рассеяния от поляризации электронов в намагниченном ферромагнетике. Экспериментальный эффект вычислялся

по формуле (2):

$$\delta_0 = 2\,\frac{N^+ - N^-}{N^+ + N^-}, \qquad (2)$$

где N^+, N^- — отсчеты γ-квантов при различных поляризациях нейтронов. Число измерений было $N \approx 10^{10}$ в каждой серии, так что статистическая ошибка имела порядок $\approx 10^{-5}$.

Обработка данных эксперимента в [6] была проведена с учетом примерно половины данных. В табл. 1 приведена часть данных из [6]. В эксперименте не было прямой возможности измерить фон в комптоновской области квантов с энергией 2.22 MeV, захваченных детектором. Также в то время вычислительные мощности были недостаточны для расчета детекторов с реалистичной геометрией. Такой расчет и в наши дни не дает полного решения вопроса из-за того, что высокоэнергетичный фон возникает одновременно с полезными сигналами из-за рассеяния нейтронов на веществе мишени [16].

В настоящей работе мы представляем математический подход, который вместе с обобщенными метриками совместимости обеспечивает использование большего количества экспериментальных данных для расчетов эффектов.

Для применения методов интервальной статистики сформируем интервальные данные и их выборку. Данные в виде интервалов образуются из данных, находящихся в столбцах 4 и 5 табл. 1 в виде [3]:

$$\mathbf{x}_i^p = [x_i^p - \Delta x_i^p,\, x_i^p + \Delta x_i^p]. \qquad (3)$$

Выборка интервальных данных $\mathbf{X}^p$ (4) состоит из значений $\mathbf{x}_i^p$ (3):

$$\mathbf{X}^p = \{\mathbf{x}_i^p\}^n, \quad i = 1, 2, \ldots, n, \quad n = 15. \qquad (4)$$

Применительно к данным (3) ленточная диаграмма рассеяния выборки интервальных данных $\mathbf{X}^p$ в области фотопика и концы интервалов массива $\mathbf{Z}$ представлены на рис. 2.

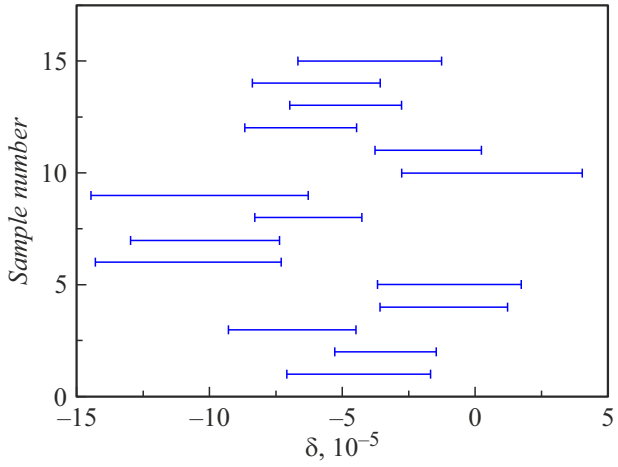


Рис. 2. Диаграмма рассеяния выборки интервальных данных $\mathbf{X}^p$ и концы интервалов массива $\mathbf{Z} = \{\mathbf{z}_i^p\}$ в области фотопика, умноженные на 10^5.

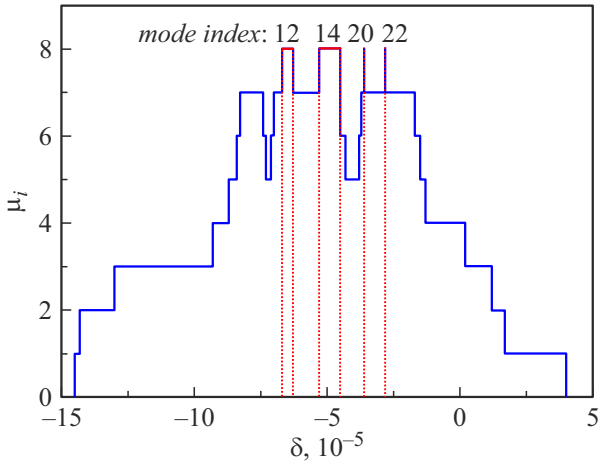


Рис. 3. График частот элементарных подынтервалов $\mathbf{Z}$.

Выборка интервальных данных $\mathbf{X}^p$ не обладает свойством совместности, поскольку в рамках классической интервальной арифметики пересечение всех измерений пусто: $\cap \mathbf{x}_i^p = \varnothing$.

В связи с пустотой пересечения всех измерений в качестве информационного множества будем использовать интервальную моду, которая обобщает обычное понятие моды выборки в статистике. Для вычисления интервальной моды по методике из книги [3] необходимо создать массив элементарных подынтервалов измерений $\mathbf{Z}$. Он состоит из упорядоченных по возрастанию значений концов интервалов данных $\mathbf{x}_i^p = [\underline{\mathbf{x}}_i^p, \overline{\mathbf{x}}_i^p]$:

$$\mathbf{Z}^p = \{\mathbf{z}_i^p\}^n, \quad i = 1, 2, \ldots, N, \quad N = 2n - 1 = 29. \qquad (5)$$

Для нахождения интервальной моды вычисляют частоты μ_i вхождения элементарных подынтервалов в интервальные данные с учетом условия (6):

$$\mathbf{z}_i \cap \mathbf{x}_i^p \neq \varnothing, \quad j = 1, 2, \ldots, n, \quad i = 1, 2, \ldots, N. \qquad (6)$$

Максимум графика μ_i соответствует интервальной моде. График частот μ_i представлен на рис. 3. Сверху указаны номера элементарных подынтервалов $\mathbf{z}_i$, образующих моду $\mathbf{X}^p$.

Как видно из графика на рис. 3, размер максимальной клики составляет 8 измерений из 15. Мода распределения является мультиинтервалом (7):

$$\operatorname{mode}\mathbf{X}^p = \mathbf{z}_{12} \cup \mathbf{z}_{14} \cup \mathbf{z}_{20} \cup \mathbf{z}_{22} = [-6.7, -6.3]$$
$$\cup\, [-5.3, -4.5] \cup [-3.6, -3.6] \cup [-2.8, -2.8]. \qquad (7)$$

Интервалы $\mathbf{z}_{20}$ и $\mathbf{z}_{22}$ являются вырожденными. Причиной данного факта является способ представления данных в [6] с малым числом значащих цифр после десятичной точки.

Таблица 1. Данные (1), умноженные на 10^5 ((−) данные не включены в [6]), СКО — среднеквадратическое отклонение

№ серии измерений	Комптон — среднее	Комптон — СКО	Фотопик	Фотопик — СКО	Фон	Фон — СКО
(1)	(2)	(3)	(4)	(5)	(6)	(7)
	x^C	Δx^C	x^p	Δx^p	x^b	Δx^b
1	−3.1	2.7	−4.4	2.7	4.2	6.7
2	−0.2	2.1	−3.4	1.9	−3.2	4.8
3	−4.0	2.1	−6.9	2.4	12.1	9
4	−2.1	2.5	−1.2	2.4	12.4	7.2
5	−3.7	1.9	−1.0	2.7	9.4	5.1
6	−1.7	3.7	−10.8	3.5	1	12.4
7	−5.7	2.8	−10.2	2.8	−0.6	6.1
8	−2.8	1.9	−6.3	2	3.9	4.3
9	−8.0	4.0	−10.4	4.1	10.3	10
10	−2.1	3.9	0.6	3.4	−4.8	10.6
11	−3.6	2.6	−1.8	2	4.6	4.2
12	−7.2	2.5	−6.6	2.1	−5.7	4.6
13	(−)	(−)	−4.9	2.1	13	3
14	(−)	(−)	−6.0	2.4	8.4	4.6
15	(−)	(−)	−4.0	2.7	10.6	5.5
Среднее	−3.5	0.7	−4.8	0.8	5.8	1.7

Результат вычислений в виде мультиинтервала в форме (6) затрудняет интерпретацию результата исследований. Более естественным выглядит

$$\text{mode}\,\mathbf{X}^p = \mathbf{z}_{12} \cup \mathbf{z}_{14} = [-6.7, -6.3] \cup [-5.3, -4.5]. \quad (8)$$

Максимальная клика K_p включает 8 измерений

$$K_p = \{1, 3, 8, 9, 12, 13, 14, 15\}, \quad (9)$$

что лишь ненамного превышает половину выборки $\mathbf{X}^p$.

2.2. Дополнительные данные и их включение в исходную выборку

Для увеличения представительности данных были привлечены измерения в области комптоновского рассеяния, находящиеся в столбцах 2 и 3 табл. 1. Они образуют выборку интервальных данных $\mathbf{X}^c$, сформированную аналогично $\mathbf{X}^p$ (2) по данным 1 и 2 столбца.

По внешнему виду диаграммы рассеяния выборки интервальных данных $\mathbf{X}^p$ и $\mathbf{X}^c$ на рис. 4 видно, данные комптоновского рассеяния и фотопика попарно совместны для большинства измерений. Таким образом, можно предположить, что в вполне могут быть использованы в обработке после необходимой коррекции.

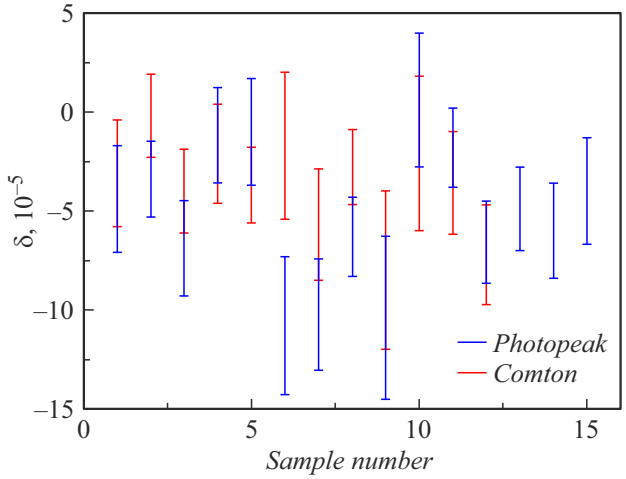


Рис. 4. Диаграмма рассеяния выборки интервальных данных $\mathbf{X}^p$ и $\mathbf{X}^c$.

Зададимся самой простой моделью коррекции. Будем считать, что относительный вклад фона во всех измерениях был постоянной величиной и коэффициент k_{BG}, учитывающий этот факт, также можно полагать константой. Полагаем, что истинная величина в области

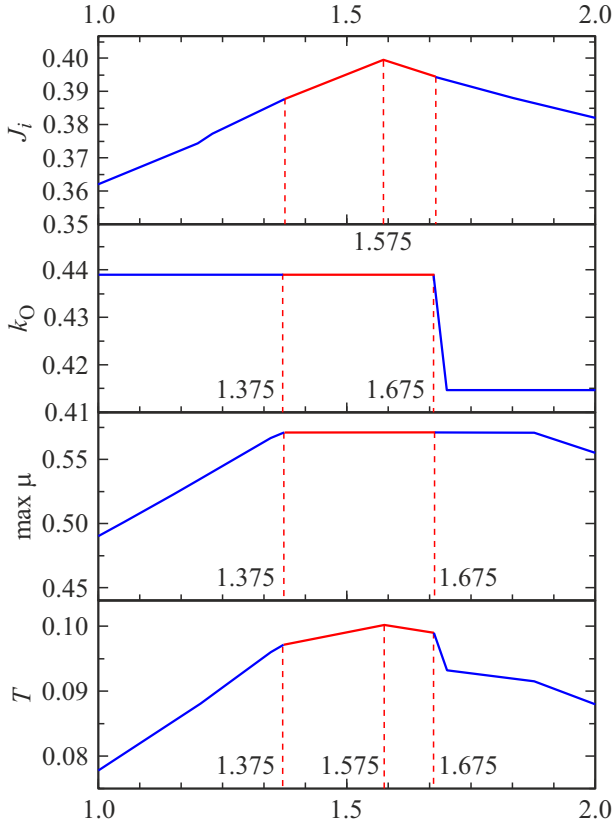


Рис. 5. Индекс Жаккара, коэффициент Оскорбина k_O, размер максимальной клики от k_{BG}, комбинированная мера совместности T по [2].

Таблица 2. Данные (1), умноженные на 10^5 (в таблице по отношению к табл. 1 изменены столбцы (2) и (3))

№ серии измерений	Комптон — среднее	Комптон — СКО	Фотопик	Фотопик — СКО
(1)	(2)	(3)	(4)	(5)
	x^C	Δx^C	x^p	Δx^p
1	−4.87	4.24	−4.4	2.7
2	−0.31	3.3	−3.4	1.9
3	−6.28	3.3	−6.9	2.4
4	−3.3	3.92	−1.2	2.4
5	−5.81	2.98	−1.0	2.7
6	−2.67	5.81	−10.8	3.5
7	−8.95	4.4	−10.2	2.8
8	−4.4	2.98	−6.3	2
9	−12.56	6.28	−10.4	4.1
10	−3.3	6.12	0.6	3.4
11	−5.65	4.08	−1.8	2
12	−11.3	3.92	−6.6	2.1
13			−4.9	2.1
14			−6.0	2.4
15			−4.0	2.7
Среднее	−5.5	1.1	−4.8	0.8

комптоновского рассеяния

$$x_C = k_{BG} x_{C_0}, \tag{10}$$

где x_C — скорректированные данные, x_{C_0} — данные из столбца 2 табл. 1. После корректировки данных $\mathbf{X}_C$, они были включены вместе с данными $\mathbf{X}^p$ (3) в единую интервальную выборку $\mathbf{X} = \{\mathbf{X}^p, \mathbf{X}^c\} = \{\mathbf{X}^p, \mathbf{k}_{BG}\mathbf{X}^{c0}\}$.

Величина k_{BG} полагается неизвестной. Для ее определения использовалась методика, впервые предложенная в [1] с вычислением комбинированной меры совместности [2]. Точечная и интервальная оценки вычисляются с помощью одномерной оптимизации в области значений от 1 (нет коррекции) до 2 (очень грубая оценка фона). Оценка по индексу Жаккара дает аргумент точечной оценки

$$k_{BG}^{opt} = \text{argmax}_{k_{BG}}\ Ji. \tag{11}$$

Интервальная оценка определяется значениями коэффициента Оскорбина k_O и максимальной частотой $\max \mu_i$.

Значение аргумента точечной оценки (8) (рис. 5):

$$k_{BG} = 1.57. \tag{12}$$

Комбинированная мера совместности дает интервальную оценку

$$k = [1.38, 1.68]. \tag{13}$$

Видно, что точечное значение k_{BG} находится надежно внутри интервальной оценки $k_{BG} \in k$, так что экстремум индекса Жаккара дает надежную точечную оценку.

Графики использованных мер совместности приводятся на рис. 5.

На рис. 5 красным цветом обозначена внешняя оценка (12) интервала мультипликативного множителя учета k_{BG}. Результирующая мера совместности T [2] дает исчерпывающее описание как внешней, так и внутренней оценки, причем оптимальная внутренняя оценка T совпадает с оптимумом меры интервальной Жаккара [1]:

$$\arg\max T = \arg\max Ji.$$

В табл. 2. приводятся модифицированные данные столбцов 2 и 3 табл. 1. с учетом оптимального множителя k_{BG} (11).

Будем теперь рассматривать 27 измерений как одну выборку

$$\mathbf{X} = \{\mathbf{X}^p, \mathbf{X}^c\} = \{\mathbf{X}^p, \mathbf{k}_{BG}\mathbf{X}^{c0}\}, \quad i = 1, 2, \ldots, 27. \tag{14}$$

Первые 15 измерений отвечают исходным данным $\mathbf{X}^p$. Следующие 12, начиная с 16-го, соответствуют добав-

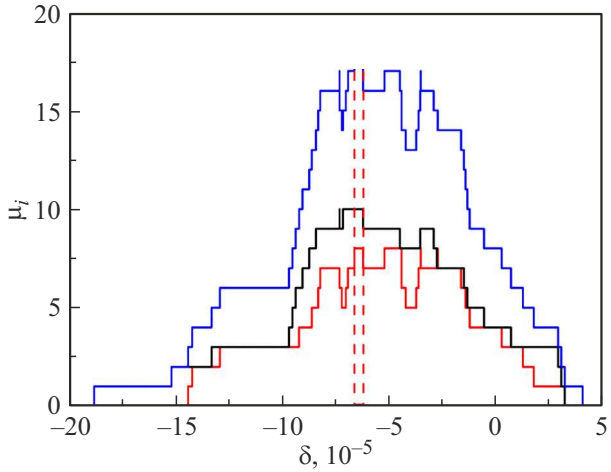


Рис. 6. График частот элементарных подынтервалов **Z** объединенной выборки **X** (11).

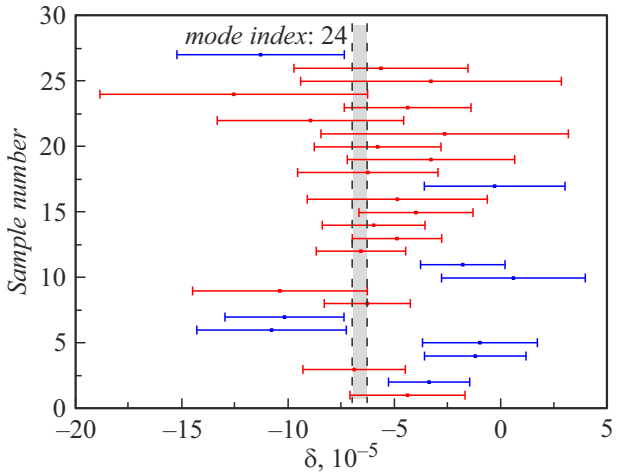


Рис. 7. Диаграмма рассеяния объединенной выборки (табл. 2).

ленным данным $\mathbf{X}^c$. Интервальные данные в $\mathbf{X}^c$ формируются аналогично формуле (3) по данным, находящимся в столбцах 2 и 3 табл. 1. Массив элементарных подынтервалов $\mathbf{Z} = \{\mathbf{z}_i\}$, $i = 1, 2, \ldots, 53$, образуется аналогично $\mathbf{Z}^p$ (4) по условию (5).

Проведем результаты вычисления интервальной моды, максимальной клики и коэффициента Оскорбина

$$\operatorname{mode}\mathbf{X} = \mathbf{z}_{24} = [-6.7, -6.3], \tag{15}$$

$$\max\mu_i = 18, \tag{16}$$

$$k_{\mathrm{O}} = 1.7. \tag{17}$$

На рис. 6 красным цветом дан график частот для выборки $\mathbf{X}^p$ (3), серым цветом — аналогичный график для $\mathbf{X}^c$, и синим цветом — для $\mathbf{X}$ (13).

Мода распределения $\mathbf{X}^c$ является мультиинтервалом, как и $\operatorname{mode}\mathbf{X}^p$ (6):

$$\operatorname{mode}\mathbf{X}^c = [-7.4, -7.4] \cup [-7.2, -6.3]. \tag{18}$$

Интервальная мода $\operatorname{mode}\mathbf{X}$ (14) всей выборки (13) является пересечением интервалов $\operatorname{mode}\mathbf{X}^p$ (6) и $\operatorname{mode}\mathbf{X}^c$ (17) и состоит из одного элементарного подынтервала $[-6.7, -6.3]$.

Ленточная диаграмма рассеяния объединенной выборки **X** (14) представлена на рис. 7. Максимальная клика K (18) выделена красным цветом.

Номера измерений для максимальной клики K выборки **X** (14) следующие:

$$K = K_p + K_C = \{1, 3, 8, 9, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26\}. \tag{19}$$

Здесь K_p (9) — массив для максимальной клики выборки $\mathbf{X}^p$ (4).

Следует заметить, что из 12 добавленных измерений 10 вошли в максимальную клику (14). Отметим заметное улучшение свойств однородности выборки **X** (12) по сравнению с выборкой $\mathbf{X}^p$ (3). Распределение частот элементарных подынтервалов стало унимодальным. Мультиинтервал интервальной моды $\operatorname{mode}\mathbf{X}^p$ (7) соответственно стал обычным интервалом $\operatorname{mode}\mathbf{X}$ (15).

Следует заметить, что измерения, не вошедшие в максимальную клику (19), выделенные на рис. 7 синим цветом, имеют многочисленные непустые пересечения с другими элементами интервальной выборки X. Таким образом, их не следует относить к невалидным и считать выбросами, они также имею ценность для характеризации всего объема данных [6]. В [3] содержится развернутое обсуждение несовместных выборок и способов их анализа.

Назовем такие измерения аутсайдерами (маргинальными) (англ. outliers), не вкладывая в этот эпитет негативного смысла. Обозначим множество аутсайдеров как K_{out}:

$$K_{\mathrm{out}} = \{2, 4, 5, 6, 7, 10, 11, 17, 27\}.$$

Построим график частот элементарных подынтервалов Z_{out} аутсайдеров объединенной выборки **X** (11). Как видно из рис. 8, распределение бимодально, и моды находятся левее и правее $\operatorname{mode}\mathbf{X}$ (14).

Множество значений $\mathbf{X}(K_{\mathrm{out}})$ более узкое, чем максимальной клики **X** (14):

$$\operatorname{wid}[-15.3, 4.0] = 19.3 < \operatorname{wid}[-18.9, 3.2] = 22.1.$$

Также можно заметить, что интервалы „левой“ моды входят в интервал $x_{24} = [-18.9, -6.3]$, а большинство интервалов „правой“ моды — в $x_{21} = [-8.5, 3.2]$. Этот факт отражает так называемый парадокс Хлебникова [3]: более широкие интервалы x_{24}, x_{21} входят в максимальную клику, а более узкие интервалы $\mathbf{X}(K_{\mathrm{out}})$ — нет.

2.3. Обсуждение результатов

Для сопоставления полученных результатов с расчетами физической величины δ [6] проведем вычисления

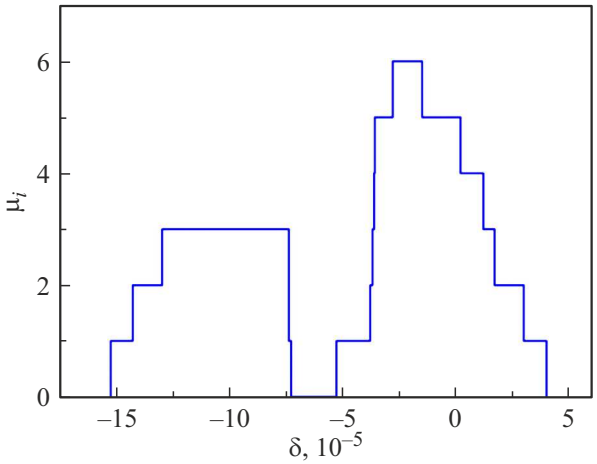


Рис. 8. График частот элементарных подынтервалов Z_{out} для объединенной выборки **X** (14).

методами теоретико-вероятностной математической статистики по данным табл. 2.

$$\delta_{(P)} = (-4.8 \pm 0.8) \cdot 10^{-5}\ [6],$$

$$P_{\gamma(P)} = (-1.5 \pm 0.3) \cdot 10^{-3}\ [6]; \qquad (20)$$

$$\delta_{(C+P)} = (-5.1 \pm 0.7) \cdot 10^{-5},$$

$$P_{\gamma(C+P)} = (-1.6 \pm 0.3) \cdot 10^{-3}, \qquad (21)$$

где индекс P указывает на вычисления с данными фотопика, а индекс $C + P$ соответствует таким же вычислениям с данными фотопика и комптоновского спектра. Неопределенность в оценке (21) приведена для уровня значимости, соответствующему одному стандартному отклонению, как и в публикации [6].

Результаты вычислений (20) показывают хорошее соответствие новых оценок с вычислениями в [6] при небольшом уменьшении оценок статистического разброса. Результаты (21) являются коррекцией [6] и вместе с обсуждением ядерно-физических аспектов задачи и различных расчетов будут опубликованы отдельно.

Развитием предложенной методики является более детальный способ коррекции данных в области комптоновского рассеяния. В настоящей работе для всех измерений использовался один и тот же коэффициент (9). При этом в публикации [6] приводятся данные по измерению фона для каждого измерения. С учетом формы спектров гамма-квантов [16] можно рассчитать вклад фона в эффект индивидуально для всех данных выборки.

Обсудим теперь вопрос о статусе интервальных измерений в более широком контексте анализа данных. Мы взяли за отправной пункт исследования базовую выборку 3 (табл. 1) и добавили к ней после необходимой коррекции дополнительные данные из этой же таблицы, но ранее не включенные в обработку. Точечные оценки и меры рассеяния (21) изменились ненамного по отношению к исходным (20).

Что же изменилось вместе с увеличением мощности выборки? Существенно возросло количество совместных между собой данных выборки, с 8 из 15, до 18 из 27 (а число аутсайдеров выросло незначительно, с 7 до 9). Таким образом, результат стал существенно более представительным. Этот факт свидетельствует в пользу того, что следует внимательнее рассмотреть статус добавленных измерений с точки зрения терминологии, представленной в разд. 1.2.

В разд. 1.2 совместные с моделью измерения назывались внутренними и не участвующими в формировании оценок информационного множества. Это действительно так, но вот в параметры мер совместности вклад таких измерений, напротив, высок. Здесь уже „активные“ граничные измерения не вносят определяющего вклада. В англоязычной терминологии это отражено в терминологической паре „inlier/outlier“. Именно внутренние измерения формируют интервальную моду и входят в максимальную клику интервальной выборки. Авторами книги [3] было предложено называть внутренние измерения „вполне совместными“.

Заключение

Применен оригинальный прием объединения выборок. Основой приема является применение комбинированной меры совместности для выборок с интервальной неопределенностью. В результате существенно увеличена максимальная клика выборки данных. Получено новое значение для параметра $P\gamma$ — базовой ядерной реакции захвата протона нейтроном. Наиболее часто $P\gamma$ используется в астрофизических расчетах при рассмотрении обратной реакции фоторасщепления дейтона с учетом поляризации гамма-квантов.

Благодарности

Авторы признательны участникам Всероссийского вебинара по интервальному анализу — С.И. Жилину, С.И. Кумкову, А.В. Пролубникову, Е.В. Чаусовой, С.П. Шарому — за творческое и конструктивное сотрудничество в области анализа данных с интервальной неопределенностью и выработки терминологии, О.М. Скрекель — за выполнение расчетов спектров гамма-квантов.

Финансирование работы

Раздел „Практический пример“ выполнен при поддержке госзадания РАН (государственное задание РФ FFUG-2024-0034. Раздел „Информационное множество и статус измерений“ выполнен при поддержке госзадания РАН (государственное задание РФ FFUG-2024-0028).

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Список литературы

[1] А.Н. Баженов, А.Ю. Тельнова. Измерительная техника, **12**, 15 (2022). https://doi.org/10.32446/0368-1025it.2022-12-15-22 [A.N. Bazhenov, A.Yu. Telnova. Meas. Tech., **65** (12), 882 (2023). http://dx.doi.org/10.1007/s11018-023-02180-2]

[2] А.Н. Баженов, С.И. Жилин, А.Ю. Тельнова. Измерительная техника, **11**, 17 (2023). http://doi.org/10.32446/0368-1025it.2023-11-17-25 [A.N. Bazhenov, S.I. Zhilin, A.Yu. Telnova. Meas. Tech., **11**, 1 (2024). https://doi.org/10.1007/s11018-024-02297-y]

[3] А.Н. Баженов, С.И. Жилин, С.И. Кумков, С.П. Шарый. *Обработка и анализ данных с интервальной неопределенностью* (НИЦ, Регулярная и хаотическая динамика РХД, Москва–Ижевск, 2024), ISBN 978-5-4344-1018-2 2024 г., 356 с.

[4] О.Е. Родионова. *Интервальный метод обработки результатов многоканальных экспериментов* (Докт. дисс. 01.04.01, Москва, 2008)

[5] А.Л. Померанцев, О.Е. Родионова. Журн. аналит. химии, **61**, 1032 (2006). DOI: 10.1134/S1061934806100030

[6] A.N. Bazhenov, L.A. Grigor'eva, V.V. Ivanov, E.A. Kolomensky, V.M. Lobashev, V.A. Nazarenko, A.N. Pirozhkov, Yu.V. Sobolev. Phys. Lett. B, **289** (1–2), 17 (1992).

[7] A.N. Bazhenov, V.M. Lobashev , A.N. Pirozhkov, V.N. Slusar. Nucl. Instrum. Methods Phys. Res., **332** (3), 534 (1993).

[8] А.Н. Баженов, В.И. Медведев, А.Н. Пирожков. Приборы и техника эксперимента, **4**, 63 (1992).

[9] A.N. Bazhenov, V.V. Ivanov. Nucl. Inst. Methods Phys. Res. A, **337** (2–3), 622 (1994).

[10] V.V. Ivanov, V.A. Solovey. Nucl. Inst. Methods Phys. Res. Section A, **311** (3), 569 (1992).

[11] I.B. Khriplovich. Phys. Atom. Nuclei, **64**, 516 (2001). https://doi.org/10.1134/1.1358476

[12] A.P. Burichenko, 1.B. Khriplovich. Nucl. Phys. A, **515** (1), 139 (1990). https://doi.org/10.1016/0375-9474(90)90327-I

[13] S.P. Shilpashree. Intern. J. Modern Phys. E, **22** (5), 1350030 (2013).

[14] W.M. Snow, A. Bazhenov, C.S. Blessinger, J.D. Bowman, T.E. Chupp, K.P. Coulter, S.J. Freedman, B.K. Fujikawa, T.R. Gentile, G.L. Greene, G. Hansen, G.E. Hogan, S. Ishimoto, G.L. Jones, J.N. Knudson, E. Kolomenski, S.K. Lamoreaux, M.B. Leuschner, A. Masaike, Y. Masuda, Y. Matsuda, G.L. Morgan, K. Morimoto, C.L. Morris, H. Nann, S.I. Penttilä, A. Pirozhkov, V.R. Pomeroy, D.R. Rich, A. Serebrov, E.I. Sharapov, D.A. Smith, T.B. Smith, R.C. Welsh, F.E. Wietfeldt, W.S. Wilburn, V.W. Yuan, J. Zerger. Nucl. Instrum. Methods Phys. Res. A, **440**, 729 (2000). doi: 10.1016/S0168-9002(99)01071-2

[15] D. Blyth, J. Fry, N. Fomin, R. Alarcon, L. Alonzi, E. Askanazi, S. Baeßler, S. Balascuta, L. Barrón-Palos, A. Barzilov, J.D. Bowman, N. Birge, J.R. Calarco, T.E. Chupp, V. Cianciolo, C.E. Coppola, C.B. Crawford, K. Craycraft, D. Evans, C. Fieseler, E. Frlež, I. Garishvili, M.T.W. Gericke, R.C. Gillis, K.B. Grammer, G.L. Greene, J. Hall, J. Hamblen, C. Hayes, E.B. Iverson, M.L. Kabir, S. Kucuker, B. Lauss, R. Mahurin, M. McCrea, M. Maldonado-Velázquez, Y. Masuda, J. Mei, R. Milburn, P.E. Mueller, M. Musgrave, H. Nann, I. Novikov, D. Parsons, S.I. Penttilä, D. Počanić, A. Ramirez-Morales, M. Root, A. Salas-Bacci, S. Santra, S. Schröder, E. Scott, P.-N. Seo, E.I. Sharapov, F. Simmons, W.M. Snow, A. Sprow, J. Stewart, E. Tang, Z. Tang, X. Tong, D.J. Turkoglu, R. Whitehead, W.S. Wilburn. Phys. Rev. Lett., **121**, 24 (2002).

[16] О.М. Скрекель. *Численное моделирование спектров гамма-квантов* (Частное сообщение. 2022).