УДК 53.088, 519.23

СРАВНЕНИЕ ГИСТОГРАММ В ФИЗИЧЕСКИХ ИССЛЕДОВАНИЯХ

С.И. Битюков*, А.В. Максимушкина, В.В. Смирнова***
**** ФГБУ ГНЦ ИФВЭ НИЦ «Курчатовский институт», 142281, Московская обл., г. Протвино, Площадь науки, 1***
***** Обнинский институт атомной энергетики НИЯУ МИФИ 249040, Калужская обл., г. Обнинск, Студгородок, 1***

Рассмотрены основные подходы к методам сравнения гистограмм в физических исследованиях. Слово «гистограмма» впервые использовано знаменитым статистиком К. Пирсоном как «обобщенная форма графического представления» [1]. Гистограммы весьма полезны в этом их каноническом применении для визуализации данных. Однако сегодня гистограммы часто рассматривают как чисто математический объект. Гистограммы стали необходимым инструментом в различных предметных областях. Помимо научного анализа данных в экспериментальных исследованиях гистограммы играют важную роль в сопровождении баз данных, обработке изображений и компьютерном «зрении» [1]. Соответственно, цели и методы обработки гистограмм меняются в зависимости от области применения. В данной работе гистограммы рассматриваются как один из элементов системы обработки информации, используемый при анализе данных, получаемых в исследованиях на экспериментальных физических установках.
Приведены некоторые методы сравнения гистограмм и результаты сравнения трех методов (статистического сравнения гистограмм (SCH), Колмогорова-Смирнова (KS) и Андерсона-Дарлинга (AD)) для определения возможности сравнения гистограмм в оценке различимости выборок, при обработке которых были получены гистограммы.

Ключевые слова: гистограмма, метод Монте-Карло, поток событий, тестовая статистика.

ВВЕДЕНИЕ

Пусть существует набор непересекающихся интервалов. Гистограмма представляет собой эмпирическое распределение его заселенности значениями реализаций некоторой случайной величины, построенное по данным выборки конечного объема. Эти интервалы принято называть бинами. Реализация случайной величины называется событием.

Анализ гистограмм зависит от процедуры заполнения гистограммы. Например, предельным случаем является распределение яркостей фотоснимка. Событием является акт съемки. Одно событие – один фотоснимок и, следовательно, одна гистограмма. Другим предельным случаем является построение гистограммы, если событием является акт измерения случайной величины с занесением полученного значения в гистограмму. Заполнение гистограммы – это цепь независимых измерений случайной величины с постепенным заполнением гистограммы, т.е. одна выборка – одна гистограмма.

В физических экспериментах обычно рассматривается второй подход к построению гистограмм. Так в физике высоких энергий событие определяется возникновением условий, позволяющих зафиксировать проявления взаимодействий частиц в детекторах, получения соответствующей информации с регистрирующей электроники в цифровом виде и возвращения установки в исходное состояние для реакции на возникновение следующего события. Поток регистрируемых событий сохраняется в виде нескольких наборов выборок для последующей обработки. Соответственно, содержимое бина гистограммы называется числом событий в бине. Сумма числа событий во всех бинах является объемом гистограммы.

При определении гистограмм существует ряд проблем общего характера, решение которых часто также зависит от решаемой задачи. Такими проблемами являются выбор оптимального биннинга в гистограмме и выбор модели распределения ошибок для наблюденной величины в бине гистограммы.

СРАВНЕНИЕ ГИСТОГРАММ

Пусть даны две гистограммы. Как оценить, подобны они или нет? И что означает «подобны»? Существует несколько подходов к решению этой задачи.

Предположим, что известна эталонная гистограмма. Часто близость эталонной и тестируемой гистограмм измеряется с помощью некоторой тест-статистики, которая обеспечивает количественное выражение «расстояния» между гистограммами [2]. Чем меньше это расстояние, тем более подобны гистограммы. В литературе существует несколько определений таких расстояний, например, расстояние по Колмогорову [3], расстояние Кульбака-Лейблера [4], полная вариация функции [5], хи-квадрат-расстояние [6]. Обычно, это тест-статистики, распределение которых можно задать формулами или построить методом Монте-Карло. Другой путь – это преобразование гистограмм в функции плотности вероятности и проведение сравнения уже плотностей. Этот подход основан на предположении, что гистограммы получены при измерении случайных переменных, которые обеспечивают основу для оценки эмпирического распределения плотности вероятности. Вычисление расстояния между двумя плотностями можно рассматривать подобно вычислению байесовской вероятности. Например, для расстояния между двумя статистическими совокупностями используют расстояние Бхаттачария [7] или Хеллингера [8]. Следует сказать, что расстояния по Колмогорову [2], по Андерсону-Дарлингу [9], по Кульбаку-Лейблеру [3] также позволяют сравнивать исходные выборки без их представления в виде гистограмм. Но это уже несколько иная задача.

Существует новый метод максимума среднего различия [10]. Еще при сравнении гистограмм используются методика, основанная на ранжировании или перестановках (метод Манна-Уитни [11]), и в ряде случаев векторный подход. Гистограммы рассматриваются как векторы с заданной размерностью числа бинов, а расстояние между ними оценивается в метриках Эвклида или Минковского [12]. Иногда вводят меру похожести (*similarity*) в некоторой логической схеме, например, такой подход рассматривается в работе [13], основанной на логике Лукасиевича.

Важной задачей при сравнении гистограмм является проверка их совместимости или, наоборот, различимости. Под совместимостью понимается утверждение о том, что обе гистограммы являются результатом обработки независимых выборок, которые получены из одного потока событий (или, что то же самое, взяты из одной генеральной совокупности событий). В работе [14] предлагается метод, позволяющий оценить различимость гистограмм и, соответственно, различимость исходных потоков событий по собранным в них выборкам. Метод основан на статистическом сравнении гистограмм; в качестве расстояния между гистограммами предлагается использовать многомерную тестовую статистику. Модификация данного метода для регистрации изменений в параметрах информационных потоков в задачах беспроволочной передачи данных приводится в [15].

Если целью сравнения гистограмм является проверка на их совместимость, то задача сводится к проверке гипотез, где основной гипотезой $H0$ будет утверждение, что гистограммы получены при обработке независимых выборок, взятых из одного и того же потока событий, а альтернативной гипотезой $H1$ будет утверждение, что гистограммы получены при обработке выборок, взятых из различных потоков событий. Выбор основной и альтернативной гипотез зависит от решаемой задачи. Определив критическую область для принятия решения и сделав выбор между $H0$ и $H1$, можно оценить вероятности совершить ошибки первого рода (α) и второго рода (β). Ошибка первого рода – это вероятность сделать выбор в пользу гипотезы $H1$, в то время как верна гипотеза $H0$. Ошибка второго рода – это вероятность сделать выбор в пользу гипотезы $H0$, в то время как верна гипотеза $H1$. Выбор уровня значимости критерия a позволяет оценить мощность критерия $1 - \beta$. Обычно уровень значимости критерия устанавливается на уровне 1, 5 или 10%. Если гипотезы равнозначны, то можно использовать другие комбинации α и β. Например, в задаче о различимости потоков событий можно использовать величину относительной неопределенности в принятии решения $(\alpha + \beta)/(2 - (\alpha + \beta))$. При использовании теста равных «хвостов» работает средняя ошибка принятия решения $(\alpha + \beta)/2$. Это связано с тем, что при работе с дискретными распределениями обычно сложно получить строгое равенство между α и β.

Существуют и другие цели сравнения гистограмм. Так поиск аномальных структур в тестируемой гистограмме, которых нет в эталонной гистограмме, – очень важная задача в физике частиц. Возможным решением такой задачи является побиновое сравнение гистограмм. При этом вычисляется вероятность того, что средние значения в бинах совпадают и на основании этого определяется наличие или отсутствие аномальных структур в гистограмме.

Сравнение гистограмм обычно разбивается на сравнение нормализации гистограмм и сравнение формы гистограмм. Часто сравнение формы гистограмм зависит от нормализации, поэтому используют комбинацию двух тестов. В простейшем случае нормализация оценивается из общих соображений. Это может быть отношение объемов сравниваемых выборок, поправленное на дополнительные знания (например, эффективности регистрации событий при наборе выборок), или отношение времен сбора сравниваемых выборок при постоянном потоке событий. При сравнении формы гистограмм обычно используют методы сравнения распределений.

Проверка гипотез о совместимости или о различимости гистограмм требует знания распределения тестовых статистик как для одной из гипотез, так и для другой. На основании сравнения этих распределений и вычисленного значения тестовой статистики делаются выводы. Знание распределений тестовых статистик не всегда позволяет оценить достоверность сделанных выводов.

Рассмотрим метод, который количественно позволяет оценивать надежность принятого решения. Как уже упоминалось, тестовые статистики можно построить методом Монте-Карло. Рассмотрим простой случай пособытийного накопления гистограмм. Число событий в каждом бине гистограммы можно рассматривать как реализацию случайной переменной с параметром «ожидаемое число событий в данном бине для данной выборки». При розыгрыше Монте-Карло нужно либо точно знать параметр разыгрываемой случайной переменной, либо использовать экспериментальную оценку этого параметра. Если для обеих гистограмм известны точно значения параметров «ожидаемое число событий в данном бине гистограммы для данной выборки», то они либо полностью совпадают, либо полностью различаются. Расстояние между гистограммами в этом случае уже не имеет смысла. Поэтому неопределенность в оценке параметров для каждого бина (по крайней мере, для одной из гистограмм) следует извлекать из полученных величин измеренных значений случайных переменных. В общем случае это не простая задача, однако есть класс распределений, позволяющий однозначно связать эти неопределенности, – это статистически дуальные распределения. В частности, для самодуальных распределений оценка параметра «ожидаемое число событий в данном бине для данной выборки» равна наблюденному числу событий в бине и является несмещенной. При этом плотность распределения доверия к величине параметра совпадает с распределением ошибки измерения числа событий в бине гистограммы. Пусть, для простоты, в бинах достаточно событий, чтобы аппроксимировать распределение ошибки нормальным распределением с нулевым средним и дисперсией, равной квадратному корню из числа событий в бине. Тогда для каждой гистограммы можно методом Монте-Карло построить имитационную модель совокупности гистограмм, которые мог бы породить тот поток событий, из которого была извлечена соответствующая выборка. Данная модель учитывает все неопределенности в оцененных параметрах для каждого бина. Эту процедуру по аналогии с генерацией повторной выборки в методике бутстрепа можно назвать генерацией повторной гистограммы. Подобная техника используется в работах [16, 17].

Так как в рассматриваемом методе число событий в каждом бине рассматривается как реализация независимой случайной величины, то при сравнении гистограмм сравниваются измеренные значения числа событий для соответствующих бинов. Удобной характеристикой для такого побинового сравнения является величина «значимость различия». Выбор конкретного представления этой величины зависит от решаемой задачи. Важно, что если сравниваемые значения являются реализациями одной и той же случайной величины, то «значимость различия» для них является реализацией случайной величины, близкой к стандартной нормальной величине. Таким образом, если сравниваемые гистограммы совместимы, то распределение полученных значений «значимостей различия» в каждом бине также должно быть близко к стандартному нормальному распределению.

СРАВНЕНИЕ МЕТОДОВ

В работе [14] в качестве расстояния между гистограммами использовались два момента распределения значимостей различия в бинах сравниваемых гистограмм – среднее и среднее квадратическое, т.е. тестовая статистика являлась двумерной.

Пусть в результате обработки двух выборок объемом N_1 и N_2 получены две гистограммы с числом бинов, равным M:

$$hist1\!: n_{11} \pm \sigma_{11}, n_{21} \pm \sigma_{21}, \ldots, n_{M\,1} \pm \sigma_{M\,1} \;\text{ и }\; hist2\!: n_{12} \pm \sigma_{12}, n_{22} \pm \sigma_{22}, \ldots, n_{M2} \pm \sigma_{M2}.$$

Сравнив эти гистограммы, нужно принять решение о том, принадлежат или нет потоки событий $G1$ и $G2$ (из которых были извлечены обработанные выборки) одной и той же генеральной совокупности, а также оценить вероятность того, что решение о том, что они принадлежат разным генеральным совокупностям, правильное. Введем «нормализованную значимость различия» в соответствующих бинах гистограмм

$$S_i = (n_{i1} - K{\cdot}n_{i2}) \,/\, (\sigma_{i1}^2 + K^2{\cdot}\sigma_{i2}^2)^{1/2}.$$

В данном случае n_{ik} – наблюденное число событий в бине i гистограммы k; σ_{ik} – соответствующее стандартное отклонение; K – некоторый коэффициент нормализации. Обычно, в зависимости от задачи, K равно либо отношению объемов выборок, либо отношению длительностей временных интервалов набора выборок. В качестве расстояния между гистограммами используется не одномерная величина, как в других методах, а многомерная. В рассматриваемом примере – двумерная $SRMS = (S^{cp}, RMS)$, где S^{cp} есть среднее значение распределения «нормализованных значимостей различия», а RMS – среднее квадратическое отклонение этого распределения.

$SRMS$ имеет ясную интерпретацию:

– если $SRMS = (0, 0)$, то две гистограммы идентичны;

– если $SRMS \approx (0, 1)$, то $G1 = G2$ (если $RMS < 0$, то выборки частично перекрываются, т.е. они не независимы);

– если вышеупомянутые условия не выполняются, то $G1 \neq G2$.

Как уже говорилось, распределения тест-статистик можно получать с помощью моделирования. Рассмотрим результаты эксперимента Монте-Карло для сравнения трех методов: статистического сравнения гистограмм (SCH), Колмогорова-Смирнова (KS) и Андерсона-Дарлинга (AD). Цель – определить возможности трех методов сравнения гистограмм в оценке различимости выборок, при обработке которых были получены гистограммы.

Фактически будет проверяться чувствительность методов к различиям в потоках событий, из которых получены выборки. Для этого моделируются две пары независимых потоков выборок (эталонная и тестовая), состоящих из реализаций случайных переменных (каждая реализация есть событие). Объем каждого потока равен 5000 выборок. Первый поток каждой пары потоков является эталонным с объемом выборки 1000 событий, полученных при моделировании случайной переменной, подчиняющейся нормальному закону $N(300, 50)$ (рис. 1а). Выбор данных параметров для эталонного распределения был задан тестовой программой. Второй поток из первой пары потоков – также эталонный с выборками объемом 2000 событий (см. рис. 1а). Второй поток второй пары потоков – тестовый с выборками объема 2000 событий с реализацией случайной переменной $N(X,W)$, где X меняется от значения 300 до 310 и W меняется от 42 до 58.

Распределения реализованных значений в бинах гистограмм для эталонной и тестовой выборок из второй пары потоков показано на рис. 1б.

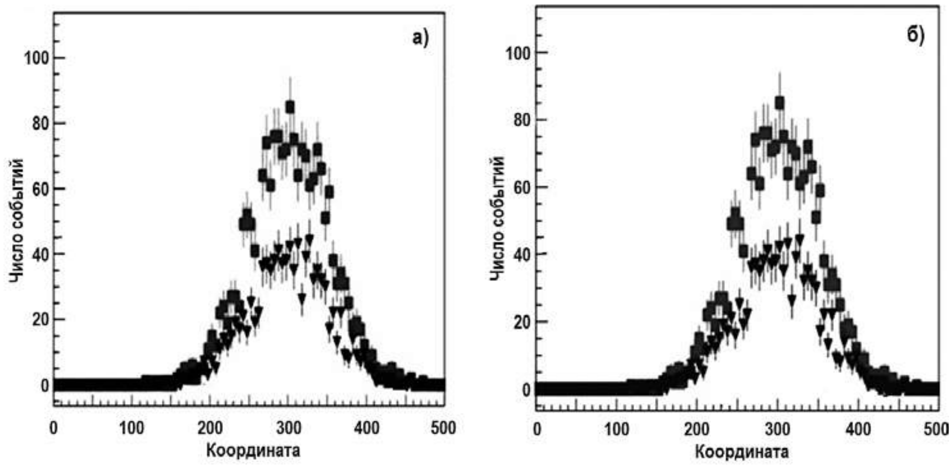

Рис. 1. Примеры распределений реализаций случайной переменной в выборках из двух пар потоков: а) –эталонная пара распределений (1000 событий и 2000 событий для $N(300,50)$; б) – тестовая пара (1000 событий для $N(300,50)$ и 2000 событий для $N(310,50)$)

При использовании критериев Андерсона-Дарлинга (AD) и Колмогорова-Смирнова (KS) для сравнения гистограмм вычисляется тест-статистика. Затем она преобразуется в п-величину или п-значение (*p-value*). П-величина – величина, используемая при тестировании статистических гипотез. Обычно п-величина равна вероятности того, что случайная величина с данным распределением (распределением тестовой статистики при нулевой гипотезе) примет значение, не меньшее, чем фактическое значение тестовой статистики. В рассматриваемом случае п-величина имеет равномерное распределение на отрезке [0, 1], если выборки получены из одного и того же потока событий. Если эталонная и тестовая выборки получены из разных потоков событий, то распределение п-величин концентрируется в малых значениях (рис. 2). В каждой паре потоков выборок проводилось 5000 сравнений пар выборок из разных потоков.

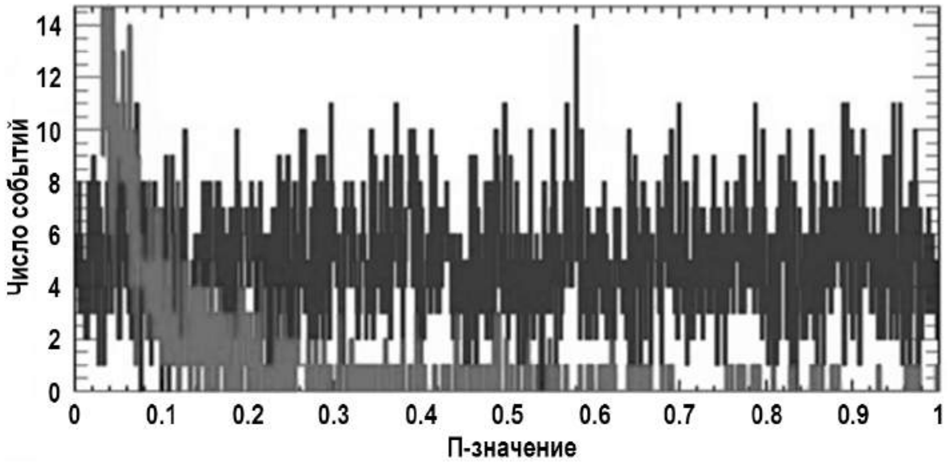

Рис. 2. Распределения значений п-величин (AD-критерий) для двух пар потоков: верхнее – 5000 сравнений для первой пары потоков (эталонные выборки 1000 и 2000 событий для реализаций $N(300,50)$); нижнее – 5000 сравнений для второй пары потоков (эталонная выборка 1000 событий для $N(300,50)$ и тестовая выборка 2000 событий для $N(306,50)$)

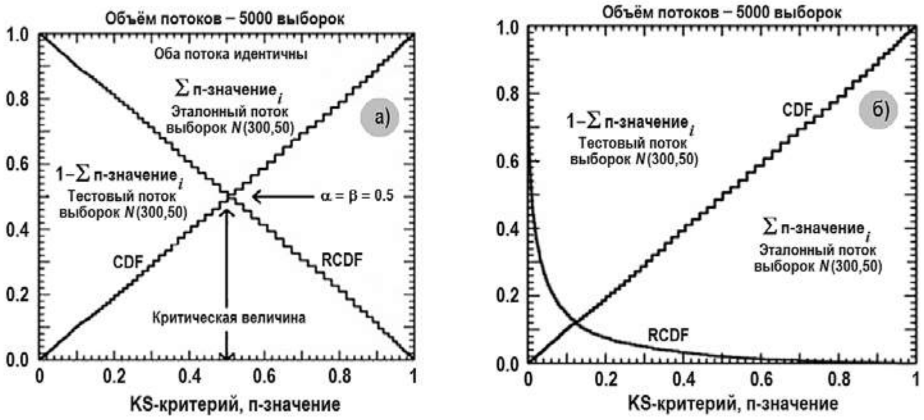

Рис. 3. Критерий Колмогорова-Смирнова. CDF и RCDF для п-величин: а) – сравнение выборок двух эталонных потоков (случайные переменные ~ $N(300,50)$); б) – сравнение выборок из эталонного потока и выборок из тестового потока (случайная переменная ~ $N(306,50)$)

Затем для первого из полученных распределений (эталонное) строилась эмпирическая кумулятивная (CDF) функция распределения (рис. 3). Для второго распределения (тестовое) строилась эмпирическая обратная кумулятивная (RCDF) функция (см. рис. 3). Критическая величина для теста равных хвостов определяется точкой пересечения линий CDF и RCDF. Следовательно, ошибки первого рода (α) и второго рода (β) при проверке гипотезы о совместимости гистограмм против альтернативы о том, что гистограммы построены при обработке выборок из разных потоков событий, примерно равны ($\alpha \approx \beta$). Это позволяет характеризовать величиной $(\alpha + \beta)/2$ различие эталонного и тестового потоков.

В случае метода статистического сравнения гистограмм (критерий SCH) для каждой выборки строится гистограмма (см. рис. 1). Затем для соответствующих пар выборок содержимое гистограмм используется для побинового вычисления «нормализованных значимостей различия» S_i, (i = 1, ..., M), определения среднего значения $S^{ср}$ и среднего квадратического значения RMS значимостей различия. Для каждой пары потоков выборок (эталонной и тестовой) строится ($S^{ср}$, RMS) – двумерное распределение полученных значений (рис. 4). Используя тест равных хвостов, находится критическая линия и вычисляются ошибки первого и второго рода, характеризующие вероятность принять неправильное решение при выборе основной или альтернативной гипотезы. Для сравнения данного критерия с критериями AD и KS также используется средняя ошибка при принятии решения $(\alpha + \beta)/2$.

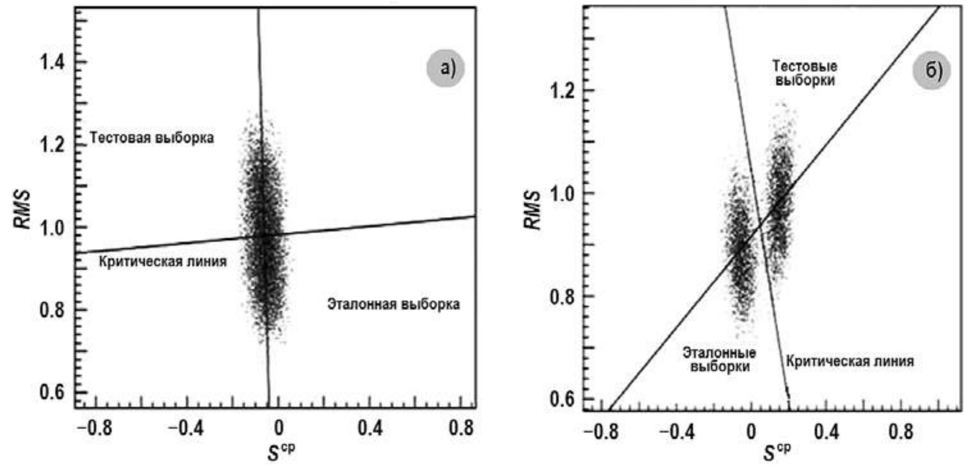

Рис. 4. Распределения ($S^{ср}$, RMS) для 5000 сравнений пар гистограмм: а) – пятно результатов для эталонных гистограмм ниже критической линии, пятно результатов сравнения тестовых выборок (изменен центр распределения, N(308, 50)) с эталонными выше критической линии; б) – пятно результатов для эталонных гистограмм слева от критической линии, пятно результатов сравнения тестовых выборок (изменена ширина распределения, N(300, 44)) с эталонными справа от критической линии

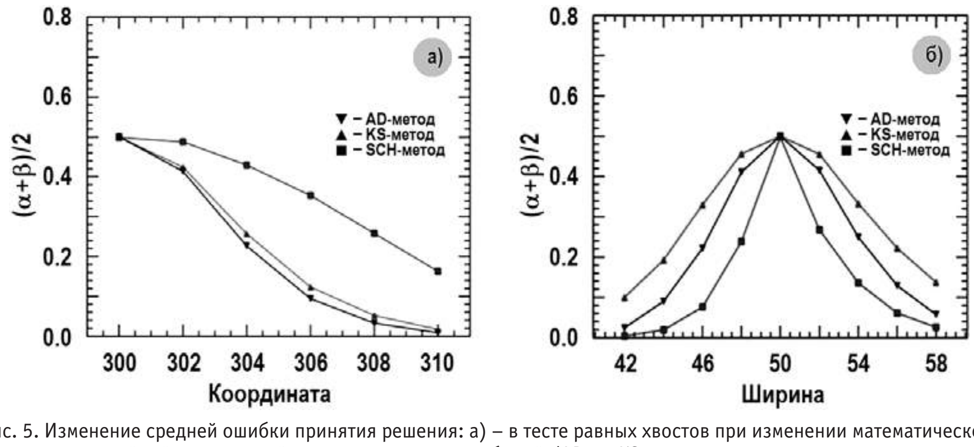

Рис. 5. Изменение средней ошибки принятия решения: а) – в тесте равных хвостов при изменении математического ожидания случайной переменной в тестовом потоке событий (AD- и KS-критерии позволяют различать потоки событий с большей вероятностью); б) – при изменении ширины распределении случайной переменной (здесь критерий SCH работает лучше)

На рисунке 5 приведены результаты сопоставления трех методов сравнения гистограмм. Результаты исследования показали, что критерии Андерсона-Дарлинга и Колмогорова-Смирнова лучше различают потоки, в которых случайные переменные имеют различие в математическом ожидании. В то же время метод статистического сравнения гистограмм лучше различает потоки, в которых случайные переменные имеют различие в ширине распределения. Однако метод статистического сравнения гистограмм [14] многомерный и дает возможность включать в качестве дополнительных размерностей любые из одномерных тестовых статистик, обычно используемых при сравнении гистограмм. Например, включение тестовой статистики Андерсона-Дарлинга в качестве третьей компоненты уже трехмерной тестовой статистики в методе статистического сравнения гистограмм решает проблемы с разделимостью потоков событий, имеющих различие в значении математического ожидания. Это является серьезным преимуществом данного метода по отношению к другим методам, рассмотренным в обзоре.

Авторы благодарны В.А. Качанову, Ю.А. Коровину и Н.В. Красникову за поддержку работы, Н.А. Корнеевой за плодотворные дискуссии. Результаты работы получены при финансовой поддержке Министерства образования и науки РФ в рамках выполнения проекта, реализуемого по Соглашению от 17 октября 2014 г. № 14.610.21.0004, идентификатор ПНИЭР RFMEFI61014X0004.

Литература

1. Ioannidis Y., The history of histograms (abridged), Proceedings 2003 VLDB Conference, 2003, pp.19-30.

2. Cha S.-H., Srihari S.N., On measuring the distance between histograms, Pattern Recognition, 2002, vol. 35, №6, pp. 1355-1370.

3. Kolmogorov A.N., Confidence limits for an unknown distribution function, Ann.Math.Stat., 1941, vol. 12, №4, pp. 461-463.

4. Kullback S., Information Theory and Statistics, Wiley, New York, 1959.

5. Rosenthal J., Convergence rates for Markov chains, SIAM Rev., 1995, vol. 37, pp. 387-485.

6. Cochran W., The chi-square test of goodness of fit, Ann.Math.Stat., 1952, vol. 23, №3, pp. 315-342.

7. Kailath T., The divergence and Bhattacharyya distance measures in signal selection, IEEE Trans.Commun.Technol, 1967, vol.15, №1, pp. 52-60.

8. Hellinger E., Neue Begrundung der Theorie quadratischer Formen von unendlichvielen Veranderlichhen, Journal fur die reine und angewandte Mathematik, (in German), 1909, vol. 1909, №136, p. 210.

9. Anderson T.W., Darling D.A., Asymptotic theory of certain «goodness of fit» criteria based on stochastic processes, Ann. Math. Statist., 1952, vol. 23, №2, p. 193.

10. Gretton A., Borgwardt K., Rasch M.J., Scholkopf B., Smola A.J., A Kernel method for two-sample problem, arXiv:0805.2368, 2008.

11. Mann H.B., Whitney D.R., On a test of whether one of two random variables is stochastically larger than the other, Ann.Math.Stat, 1947, vol.18, №1, p.50.

12. Bandemer H., Nather W., Fuzzy data analysis, Kluwer academic publishers, Dordrecht, 1992.

13. Luuka P., Collan M., Modulo similarity in comparing histograms, Proc. of IFSA-EUSFLAT2015, Eds. Alonso J.M., Bustince H., M. Reformat, Atlantis Press, 2015.

14. Bityukov S., Krasnikov N., Nikitenko A., Smirnova V., A method for statistical comparison of histograms, arXiv:1302.2651, 2013.

15. Krupanek B., Bogacz R., Comparison algorithm of multimodal histograms from wireless transmission, Przeglad Electrotechniczny, 2014, №11, p.32.

16. Cao Y., Petzold L., Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems, J. of Computational Physics, 2006, vol. 212, №1, pp. 6-24.

17. Xu K.-M., Using the bootstrap method for a statistical significance test of differences between summary histograms, NASA Technical Reports Server, ID: 20080015431, 2006.

Поступила в редакцию 21.08.2015 г.

Авторы

Битюков Сергей Иванович, ведущий научный сотрудник, доктор физ.-мат. наук
E-mail: Serguei.Bitioukov@cern.ch

Максимушкина Анастасия Владимировна, старший преподаватель, кандидат техн. наук
E-mail: a.v.saenko@mail.ru

Смирнова Вера Васильевна, старший научный сотрудник, кандидат физ.-мат. наук
E-mail: Vera.Smirnova@ihep.ru

UDC 53.088, 519.23

COMPARISON OF HISTOGRAMS IN PHYSICAL RESEARCH

Bityukov S.I.*, Maksimushkina A.V.** , Smirnova V.V.*

* State Research Center - Institute for High Energy Physics
1 Ploschad' nauki, Protvino, Moscow reg., 142281 Russia
** Obninsk Institute for Nuclear Power Engineering,
National Research Nuclear University «MEPhI»
1 Studgorodok, Obninsk, Kaluga reg., 249040 Russia

ABSTRACT

The review of methods of histograms comparison is presented. Possible approaches for the comparative analysis of histogram are considered.

The term "histogram" was coined by the famous statistician Karl Pearson to refer to a "common form of graphical representation" [1]. Histograms are very useful in their canonical visual representation, but today histograms are considered as purely mathematical objects.

Histograms are used in different scientific fields. Besides physics data analyses, histograms play a very important role in databases, image processing and computer vision [1]. Correspondingly, goals and methods of treatment of histograms are varied in dependence to the area of application. In this paper histograms are considered in frame of tasks related to physical experiments.

The paper presents some of the methods and results of the comparison of histograms. A comparison was made of three methods of the comparison of histograms: the Kolmogorov-Smirnov (KS) method, the Anderson-Darling (AD) method and the method for statistical comparison of histograms (SCH).

The dependence of the mean error in hypotheses testing of distinguishability of the reference data set and test data set on the difference in position parameter of samples: the Anderson-Darling and Kolmogorov-Smirnov criteria give the better result than SCH method. The dependence on the width parameter of samples: the SCH criterion gives the better result than AD and KS criteria.

Nevertheless, the SCH is a multidimensional method. It allows to include the any one-dimensional test statistic as an additional component of multidimensional test statistic in the frame of the method. For example, the including of the AD test statistic into SCH method as third component of the three dimensional test statistic will allow to reach the better coordinate resolution in the example which was considered above.

Possible approaches for the comparative analysis of histogram are considered. As shown, there is no single best test for all applications. It means that before application any test must be checked with care. A good solution is a combined use of several tests.

Key words: a histogram, the Monte Carlo method, the event flow, the test statistic

REFERENCES

1. Ioannidis Y., The history of histograms (abridged), Proceedings 2003 VLDB Conference, 2003, pp.19-30.

2. Cha S.-H., Srihari S.N., On measuring the distance between histograms, Pattern Recognition, 2002, v. 35, no. 6, pp. 1355-1370.

3. Kolmogorov A.N., Confidence limits for an unknown distribution function, Ann.Math.Stat., 1941, v. 12, no. 4, pp. 461-463.

4. Kullback S., Information Theory and Statistics, Wiley, New York, 1959.

5. Rosenthal J., Convergence rates for Markov chains, SIAM Rev., 1995, v. 37, pp. 387-485.

6. Cochran W., The chi-square test of goodness of fit, Ann.Math.Stat., 1952, v. 23, no. 3, pp. 315-342.

7. Kailath T., The divergence and Bhattacharyya distance measures in signal selection, IEEE Trans.Commun.Technol, 1967, v.15, no. 1, pp. 52-60.

8. Hellinger E., Neue Begrundung der Theorie quadratischer Formen von unendlichvielen Veranderlichhen, Journal fur die reine und angewandte Mathematik, (in German), 1909, v. 1909, no. 136, p. 210.

9. Anderson T.W., Darling D.A., Asymptotic theory of certain «goodness of fit» criteria based on stochastic processes, Ann. Math. Statist., 1952, v. 23, no. 2, p. 193.

10. Gretton A., Borgwardt K., Rasch M.J., Scholkopf B., Smola A.J., A Kernel method for two-sample problem, arXiv:0805.2368, 2008.

11. Mann H.B., Whitney D.R., On a test of whether one of two random variables is stochastically larger than the other, Ann.Math.Stat, 1947, v.18, no. 1, p.50.

12. Bandemer H., Nather W., Fuzzy data analysis, Kluwer academic publishers, Dordrecht, 1992.

13. Luuka P., Collan M., Modulo similarity in comparing histograms, Proc. of IFSA-EUSFLAT2015, Eds. Alonso J.M., Bustince H., M. Reformat, Atlantis Press, 2015.

14. Bityukov S., Krasnikov N., Nikitenko A., Smirnova V., A method for statistical comparison of histograms, arXiv:1302.2651, 2013.

15. Krupanek B., Bogacz R., Comparison algorithm of multimodal histograms from wireless transmission, Przeglad Electrotechniczny, 2014, no. 11, p.32.

16. Cao Y., Petzold L., Accuracy limitations and the measurement of errors in the stochastic simulation of chemically reacting systems, J. of Computational Physics, 2006, v. 212, no. 1, pp. 6-24.

17. Xu K.-M., Using the bootstrap method for a statistical significance test of differences between summary histograms, NASA Technical Reports Server, ID: 20080015431, 2006.

Autors

Bityukov Sergey Ivanovich, Leading Scientist, Dr. Sci. (Phys.-Math.)
E-mail: Serguei.Bitioukov@cern.ch

Maksimushkina Anastasiya Vladimirovna, Senior Teacher, Cand. Sci. (Engineering)
E-mail: a.v.saenko@mail.ru

Smirnova Vera Vasil'evna, Senior Research Scientist, Cand. Sci. (Phys.-Math.)
E-mail: Vera.Smirnova@ihep.ru