УДК 51: 101.8

АНАЛИЗ УСЛОВИЙ ПРИМЕНЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ ПО КОЛМОГОРОВУ

Резников Владимир Моисеевич
к.филос.н., доцент
Институт философии и права СО РАН, Новосибирск, Россия

Анализируются известные подходы к объяснениям требований Колмогорова для вероятностей, изучаемых в приложениях. Обсуждаются наиболее обоснованные объяснения, предлагаются и исследуются новые подходы к объяснению связанности этих требований

Ключевые слова: ВЕРОЯТНОСТЬ, ПРИНЦИП КУРНО, ТЕОРЕМА БЕРНУЛЛИ, ЗАВИСИМОСТЬ ТРЕБОВАНИЙ, НЕЗАВИСИМОСТЬ ДАННЫХ, ТЕОРЕТИЧЕСКАЯ ВЕРОЯТНОСТЬ, МАЛАЯ ВЕРОЯТНОСТЬ

UDC 51: 101.8

ANALYSIS OF APPLICATION CONDITIONS OF KOLMOGOROV' PROBABILITY THEORY

Reznikov Vladimir Moiseevich
Cand.Philos.Sci., associate professor
Institute of Philosophy and Law of SB RAS, Novosibirsk, Russia

The article analyses well-known approaches to the explanations of Kolmogorov's conditions for probabilities studied in applications. It discusses better-founded explanations, offers and looks into new approaches to the explanations of connectedness of the conditions

Keywords: PROBABILITY, COURNOT' S PRINCIPLE, BERNOULLI THEOREM, DEPENDENCY OF CONDITIONS, DATA INDEPENDENCE, MATHEMATICAL PROBABILITY, IMPROBABLE

В 1936 году вышла книга Колмогорова, в которой была предложена аксиоматика теории вероятностей [1]. Эта аксиоматика представляла собой частичное решение шестой проблемы Гильберта и была принята математическим сообществом. В целом сугубо математическая, книга Колмогорова является разноплановой. В частности, в ней уделено внимание философским проблемам, связанным с математикой. Во-первых, исследовалась значимость понятия «независимость» для создания и развития теории вероятностей вплоть до последнего времени. Колмогоров пишет: «Исторически независимость испытаний и случайных величин явилась тем математическим понятием, которое придало теории вероятностей своеобразный отпечаток» [4, с. 18]. Во-вторых, Колмогоров отмечает значимость философского анализа для исследования предпосылок, при которых реальные явления могут считаться независимыми. Он писал: «Соответственно этому одной из важнейших задач философии естественных наук, после разъяснения пресловутого вопроса о сущности самого понятия вероятности, является выяснение и уточнение тех предпосылок, при которых можно какие-либо данные действительные рассматривать как независимые» [4, с. 13]. В-третьих, Колмогоров сформулировал требования к вероятностям событий, которые изучаются в приложениях теории вероятностей. Эти требования предполагают выполнение двух условий:

«А. Можно практически быть уверенным, что если комплекс условий S будет повторен большое число раз n и если при этом через m обозначено число случаев, при которых событие A наступило, то отношение m/n будет мало отличаться от P(A).

B. Если P(A) очень мало, то можно практически быть уверенным, что при однократной реализации условий S событие A не будет иметь места» [4, с. 13].

Условие А утверждает, что при проведении большого числа экспериментов частота изучаемого события будет близка к теоретической вероятности. Условие В запрещает появление маловероятного события при реализации единственного эксперимента. В известной литературе запрет маловероятных событий связывают с А. Курно. Он писал: «Следовательно, физически невозможное событие – это такое событие, математическая вероятность которого бесконечно мала» [6, с. 89].

Как утверждают Г. Шейфер и В. Вовк, идея Курно о невозможности реализации маловероятного события по предложению Г. Фреше получила название принципа Курно в 1949 году [11].

Отметим, что, по мнению Шейфера и Вовка, в отличие от аксиоматики Колмогорова, которую знают все математики, требования Колмогорова к вероятностям событий, изучаемых в приложениях, в известных работах не используются. [11, p. 54].

Даже в работах математиков, близких к школе А.Н. Колмогорова, эти условия формулируют не полностью. Только в известной монографии Г. Крамера, посвященной математической статистике, эти условия формулируются в полном объеме [11, p. 54, 5, с. 170-171]. Если говорить более точно, то во многих руководствах по теории вероятностей, в том числе и в известных: Б.В. Гнеденко, А.А. Боровкова и др., приводятся примеры, поясняющие требование А, однако условие В не отмечается и не иллюстрируется [2, с. 22; 3, с. 14].

Как известно, в настоящее время одной из проблем прикладной теории вероятностей является разработка строго обоснованных требований и норм, определяющих условия применения теории вероятностей, в частности это относится к описаниям событий, изучаемых в приложениях. До конца эта проблема еще не решена, однако традиционно предполагается, что описания событий являются независимыми. Действительно, нет особого смысла в использовании связанных утверждений о вероятностях событий, например, в случае, если какие-то утверждения выводимы на основе других посредством некоторых теорем, тогда достаточно использовать независимые утверждения, так как в случае необходимости, зависимые будут выведены на основе независимых утверждений и необходимых для вывода теорем. Что касается требований А.Н. Колмогорова, то, как отмечали его современники, в том числе известные математики, такие как Э. Борель, П. Леви и другие – колмогоровские требования оказываются связанными.

Если говорить более точно, то одно из условий Колмогорова зависит от другого, а именно условие А зависит от условия В, так как А выводимо на основе В и теоремы Бернулли. Несмотря на давнюю известность факта связанности требований Колмогорова в зарубежной литературе, в отечественных публикациях это не было отражено. Отметим, что в известной литературе, вплоть до недавнего времени, не было работ, посвященных объяснению зависимости условия А от условия В. В 2001–2006 гг. в ряде работ Шейфера и Вовка было предложено несколько рациональных объяснений этой зависимости [11-13].

Данная статья является продолжением и развитием наших ранее опубликованных работ, которые посвящены, во-первых, исследованию требований А.Н. Колмогорова, в частности принципу А. Курно, а во-вторых, философскому и методологическому анализу предложенных Шейфером и Вовком объяснений зависимости условий Колмогорова. В этих работах была обоснована значимость условий Колмогорова в контексте приложений математической статистики и теории вероятностей [7-8,10]. Кроме того, были сформулированы достоинства и недочеты интересного анализа требований Колмогорова к вероятностям, проведенного в работах Шейфера и Вовка [11-13].

Исследование условий Колмогорова является актуальным по ряду причин. Во-первых, его условие B представляет собой сильный принцип Курно, на основе которого осуществляется проверка гипотез в математической статистике и достигается наилучшая аппроксимация вероятности с помощью частот в теореме Бернулли. Отметим, что в соответствии с сильным принципом Курно проверяемая статистическая гипотеза отвергается, когда происходит событие с ничтожной вероятностью, если это событие является следствием изучаемой гипотезы.

Во-вторых, согласно Г. Фреше, А.Н. Колмогорову и другим математикам, знание, как применять математику, приводит к пониманию математических объектов. В-третьих, как условия Колмогорова оказались непопулярными, так и их интересный анализ в работах Шейфера и Вовка оказался фактически незамеченным, и нам неизвестны работы, где дан анализ предложенных Шейфером и Вовком объяснений зависимости требований Колмогорова к вероятностям.

Прежде чем исследовать объяснения Шейфера и Вовка, сделаем ряд замечаний. На первый взгляд, возникает несколько вопросов к факту выводимости условия A на основе условия B и теоремы Бернулли.

Во-первых, вопрос об адекватности этих условий для использования в теории вероятностей и, в частности, как применить теорему Бернулли к условию В, чтобы получить А. Проблема возникает, так как условия А и В заданы неформально. Условие А говорит о близости частоты события и его вероятности, однако оно не говорит, насколько близки частота и вероятность этого события. Условие В утверждает, что в единственном эксперименте запрещено появление маловероятного события. Однако оно не говорит о том, насколько ничтожной должна быть вероятность события, для того чтобы это событие не произошло. Тем не менее, данные обстоятельства не являются препятствием для применения теоремы Бернулли, так как при ее использовании исследователь задает, какова близость частоты и вероятности, и насколько малой должна быть вероятность события, для того чтобы его считать невозможным. Отметим, что если условия Колмогорова рассматривать как неформальные условия, которые и не предполагалось формализовывать, тогда вполне объяснимо, почему практически все математики, читавшие книгу Колмогорова, не заметили, что эти условия оказываются связанными с помощью теоремы Бернулли.

Во-вторых, возникает вопрос о необходимости условия В вообще, так как условие А является заключением теоремы Бернулли, и при этом условие А выводится в теореме Бернулли без использования В. Однако условие В имеет значение, потому что когда оно учитывается, то условие А оказывается верным на любой конечной типичной выборке, то есть условие В приводит к получению А на выборках меньшего объема по сравнению с непосредственным выводом А на основе теоремы Бернулли.

Первая часть работы состоит в классификации объяснений Шейфера и Вовка по степени обоснованности, от менее обоснованных к более обоснованным:

1) Зависимость не является существенным недостатком, так как в книге Колмогорова условия применения теории вероятностей представлены после его аксиоматики, следовательно теорема Бернулли еще не была получена и, поэтому, вывод требования А на основе требования В и теоремы Бернулли не мог быть осуществлен. По нашему мнению это объяснение является формальным и не вполне серьезным.

2) Колмогоров мог просто не обращать внимания на то, что его требования к вероятностям оказались зависимыми. Это предположение основано на том, что Колмогоров нигде не дал каких-либо объяснений факта зависимости его требований к вероятностям. Однако трудно предположить, что Колмогоров и многочисленные читатели его книги не замечали формальной связанности этих требований, скорее всего условия Колмогорова воспринимались неформально. Однако, если требования к применению математики являются неформализуемыми, то они оказываются вне математики, и не могут быть строго проанализированы и использованы в приложениях.

К вполне серьезным и обоснованным относятся следующие объяснения:

3) Условие А хотя и выводимо на основе условия В, имеет самостоятельное значение, так как является частотным, а Колмогоров отмечал, что его требования к применению теории вероятностей близки к Мизесу, одному из основателей частотной интерпретации теории вероятностей. А.Н. Колмогоров писал: «В изложении необходимых предпосылок для приложимости теории вероятностей к миру действительных событий автор в значительной мере следует выводам Мизеса» [4, с.12].

4) Существуют трудности верификации свойства независимости в исследуемых данных, когда они представлены выборкой большого объема. Проверка независимости имеет значение для корректного применения

теоремы Бернулли, так как первоначально теорема была доказана в предположении независимости данных.

Во второй части нашей работы предлагается усиление двух последних объяснений, а также обобщение четвертого объяснения Шейфера и Вовка. Под обобщением объяснения понимается привлечение новых аргументов, которые аналогичны предложенным Шейфером и Вовком. Под усилением объяснения понимается введение новых аргументов, не являющихся аналогичными аргументам, использованным Шейфером и Вовком.

Так, усиление третьего объяснения связано с привлечением новых доводов, основанных на эпистемологической, логической и операциональной критике принципа Курно, представленного условием B.

1. Философские недостатки принципа Курно

Во-первых, как известно принцип Курно осуществляет связь теории вероятностей и математической статистики с миром исследуемых в этих науках явлений посредством определения условий фальсификации суждений, описывающих вероятности изучаемых событий. Однако, строго говоря, вероятностные суждения, не верифицируемы и не фальсифицируемы эмпирически. Во-вторых, принцип запрещает реализацию событий, имеющих низкую вероятность, поэтому он имеет не вероятностный, а в лучшем случае вероятностно-детерминистский характер, так как значимая особенность случайности связана с реализацией маловероятных событий. В-третьих, принцип Курно апеллирует к вероятностям, которые являются теоретическими величинами, однако теоретические величины априори неизвестны, тем более это относится к малым, ничтожным вероятностям.

2. Логические недостатки принципа Курно

Существуют структуры данных, для которых принцип Курно неверен. Приведем известный из литературы пример [10]. Предположим, что Джон – американец. Следовательно, маловероятно, что он член конгресса США. Однако он член конгресса США. Тогда, по логике проверки статистических гипотез, гипотеза, что он американец, опровергается, хотя очевидно, что это верная гипотеза. Данный пример не является единственным, приведем собственный пример. Майк профессиональный хирург и специалист в области лечения определенной болезни. Следующий его пациент практически безнадежный больной. Однако если операцию делает Майк, то вероятность благополучного исхода мала, но больше нуля. Он делает операцию, и пациент остался жить. Произошло маловероятное событие, тогда, по логике проверки гипотез, предположение, что Майк профессиональный хирург, должно быть отвергнуто [8].

3. Практические слабости принципа Курно

Во-первых, вопреки принципу Курно маловероятные события действительно реализуются в единственном эксперименте. Так, например, в генетике известны реализации событий с чрезвычайно ничтожными вероятностями. Во-вторых, принцип Курно не учитывает интеллектуальные и другие затраты, необходимые для его использования, в частности он не учитывает количество испытаний, необходимых для определения вероятностей экспериментальным образом, так для того чтобы определить вероятность порядка 10^{-n}, необходимо проведение порядка 10^{n+1} экспериментов, по скромной оценке Ю.И. Алимова [1, с. 20]. Обобщение четвертого объяснения основано на том, что теорема Бернулли верна и для малозависимых наблюдений, поэтому к указанным Шейфером и Вовком трудностям проверки на независимость добавляются проблемы проверки на свойство Маркова, мартингальную зависимость и т.д. Усиление четвертого объяснения основано на использовании эмпирической, частотной интерпретации теоремы Бернулли и на демонстрации требуемых интеллектуальных затрат для ее применения путем определения необходимого количества испытаний для вычисления вероятностей, входящих в заключение теоремы Бернулли. Отметим, что в эмпирической интерпретации теоремы Бернулли предполагается определение всех теоретических вероятностей, входящих в заключение теоремы, посредством частот для конечных выборок данных.

Использование конечной частотной интерпретации связано с тем, что Колмогоров критиковал частотную интерпретацию Мизеса, так как в ней вероятности определялись для последовательностей данных бесконечной мощности. В простейшем случае теорема Бернулли формулируется так.

Теорема. Проводится n независимых испытаний события A, и в m экспериментах произошло событие A. Известно, что теоретическая вероятность появления события A в каждом эксперименте равняется p(A), m/n – это частота события A, ε – это точность вычислений. Тогда, при бесконечном числе экспериментов выполняется следующее равенство:

$$\lim_{n\to\infty} P(\,|\,m/n - p(A)| < \varepsilon) = 1 \qquad (1)$$

Иногда утверждают, что теорема имеет операциональное и эпистемологическое значение по следующим основаниям. Во-первых, так как теорема устанавливает близость теоретической вероятности и эмпирических частот. Во-вторых, еще более важно то, что она определяет вероятность для этой близости. По-нашему мнению, трудно говорить об эпистемологической значимости, так как теорема предполагает заранее известной теоретическую вероятность p(A), хотя обычно она неизвестна. Во-вторых, при естественной эмпирической интерпретации требуется намного больше вычислений, чтобы определить внешнюю вероятность P в выражении (1) по сравнению с внутренней вероятностью p(A).

Для того чтобы оценить сложность определения внешней вероятности P по сравнению с внутренней вероятностью p, вначале оценим количество необходимых экспериментов с целью эмпирического определения вероятности p. Для этого проведем следующий мысленный эксперимент, который будет состоять в бросании правильной монеты, при этом успехом будет считаться выпадение герба. Событие, состоящее в выпадении герба, будем обозначать символом A.

Предполагается проведение k серий экспериментов, где каждая серия состоит из n экспериментов по наблюдению за выпадением герба монеты, тогда всего необходимо осуществить $k \times n$ экспериментов. Отметим, что при проведении каждой серии экспериментов будет определена одна частота события A. Очевидно, что при реализации k серий – будет получено k частот события A. При реалистичной, эмпирической интерпретации теоретические величины являются неизвестными, поэтому, несмотря на то, что в теореме считается известной теоретическая вероятность события A – p(A); она в предлагаемом эксперименте не является известной и определяется посредством частотных характеристик события A, следующим образом. А именно, если эти частоты реально близки друг к другу, т. е. попадают в интервал, который меньше или равен точности наблюдений – ε, то в качестве вероятности возьмем любую из частот. Если частоты более-менее близки, то путем усреднения частот получим искомую вероятность. Если же частотные характеристики очень вариабельны, то хорошей частотной оценки не существует. Будем считать, что в нашем случае хорошая частотная оценка существует.

Теперь проведем анализ требуемого количества испытаний для оценивания внешней вероятности P. Внешняя вероятность вычисляется на основе частот w, с которыми выполняется следующее неравенство:

$| m/n-p(A)|<\varepsilon$ (2)

Отметим, что, несмотря на то, что определение частотных характеристик w для события, определенного выражением (2), сложнее, чем определение частоты m/n события A, тем не менее в рамках теоремы Бернулли для определения частоты w предполагается проведение того же самого эксперимента, который использовался для определения частоты события A. Тогда, согласно формуле (2) частоты w будут вычисляться на основе частот появления события A и уже оцененной вероятности события A. Предположим, что проведена i-я серия из n наблюдений и определена некоторая частота m_i/n события A, тогда, зная p(A) и ε, осуществляется верификация выполнимости неравенства (2). Факт выполнимости неравенства (2) будем обозначать – числом один, а невыполнимость – числом ноль. Ранее проведение серии из n экспериментов приводило к получению частоты события A, теперь та же серия экспериментов приводит к получению сингулярной оценки, определяющей выполнимость второго неравенства. Если мы хотим определить частотную характеристику w с той же точностью, с какой определялись частоты события A, то необходимо определять частоту w на основе n характеристик, описывающих выполнимость неравенства (2). Для того чтобы получить n свидетельств, определяющих выполнимость неравенства (2), необходимо провести n серий экспериментов, где каждая серия состоит из n бросаний монеты, тогда для получения одной частотной характеристики w необходимо проведение n^2 экспериментов. Так как вероятность p(A) определялась на основе k частотных характеристик m_i/n, i=1, k, поэтому для определения вероятности выполнимости (2) на основе k частотных характеристик требуется проведение $k\times n^2$ экспериментов.

Однако в действительности для получения частоты более сложного события, определенного выражением (2), надо, чтобы соответствующая серия экспериментов состояла из N наблюдений, где N>>n, и потребуется провести намного больше серий наблюдений K, где K >>k. Здесь выражение X>>Y означает, что X намного больше Y.

В результате вместо $n \times k$ наблюдений для определения вероятности p(A) необходимо провести $N \times K$ наблюдений с целью получения вероятности P, где $N \times K >> k \times n^2$.

Итак, предложенные Шейфером и Вовком объяснения зависимости требований Колмогорова на основе применения теоремы Бернулли и предпочтения Колмогоровым частотной интерпретации Мизеса являются рациональными, но неполными. Во-первых, согласно Шейферу и Вовку применение теоремы Бернулли сопряжено с трудностями верификации свойства независимости для выборок большого объема данных, так как теорема верна для независимых наблюдений. Однако теорема верна и для малозависимых наблюдений, поэтому к отмеченным трудностям добавляются проблемы верификации свойства Маркова, мартингальной зависимости и другие. Во-вторых, нами подтверждается самостоятельная ценность условия A посредством описания разнообразных слабостей как принципа Курно, так и теоремы Бернулли, на основе которых выводится условие A. В результате утверждается методологическая значимость требований Колмогорова, несмотря на их формальную связанность.

Список литературы

1. Алимов Ю.И. Альтернатива методу математической статистики. М.: Знание, 1980. – 84 с.
2. Боровков А. А. Теория вероятностей. М.: Наука, 1986.– 432 с.
3. Гнеденко Б. В. Курс теории вероятностей. М.: Наука, 1988. – 451 с.
4. Колмогоров А. Н. Основные понятия теории вероятностей. М.: Наука, 1974. – 120 с.
5. Крамер Г. Математические методы статистики. М: Мир, 1975. – 648 с.
6. Курно А. Основы теории шансов и вероятностей. М: Наука, 1970. – 384 с.
7. Резников В. М. Методологический анализ приложений теории вероятностей у Колмогорова // Вест. Новосиб. ун-та. Сер. Философия. Новосибирск. – 2009. – Вып. 1, Т. 7. – С. 26-31.

8. Резников В. М. Об адекватности математической статистики // Материалы Всероссийской конференции с международным участием «Знания-Онтологии-Теории». –
Новосибирск: Изд-во РИЦ Прайс-Курьер, 2013, Т. 2., с. 106–111.

9. Cohen J. The Earth is Round (p<0,05) // American Psychologist. – 1994. – Vol. 49, № 12. – P. 997–1003.

10. Reznikov V. On Kolmogorov's analysis of applicability of probability theory // Сборник материалов международной конференции «Мальцевские чтения». – Новосибирск: Изд-во Института Математики СО РАН, 2009. - с. 221.
Сайт института математики СО РАН. [Электронный ресурс]. URL: http://www.math.nsc.ru/conference/malmeet/09/Abstracts/abstracts-09 (дата обращения: 09.03.2013).

11. Shafer G., Vovk V. The Game-Theoretic Probability and Finance Project, Working Paper № 4. 2003. P. 11.
[Электронный ресурс]. URL: http://www.probabilityandfinance.com (дата обращения: 03.02.2013).

12. Shafer G., Vovk V. Probability and Finance It's Only a Game! N.Y.: A Wiley-Interscience Publication, 2001. – 414 p.

13. Shafer G., Vovk V. The Sources of Kolmogorov's Grundbegriffe // Statistical Science. – 2006. – Vol. 21, № 1. P. 70–98.

References

1. Alimov Yu.I. Alternativa metody matematicheskoj statistiki. M.: Znanie, 1980. – 84 s.

2. Borovkov A. A. Teorija verojatnostej. M.: Nauka, 1986. – 432 s.

3. Gnedenko Bo.V. Kurs teorii verojatnostej. M.: Nauka, 1988. – 451 s.

4. Kolmogorov A. N. Osnovnye ponjatija teorii verojatnostej. M.: Nauka, 1974. – 120 s.

5. Kramer G. Matematicheskie metody statistiki. M: Mir, 1975. – 648 c.

6. Kurno A. Osnovy teorii shansov i verojatnostej. M.: Nauka, 1970. – 384 s.

7. Reznikov V. M. Metodologicheskij analiz prilozhenij teorii verojatnostej u Kolmogorova // Vest. Novosib. un-ta. Ser. Filosofija. Novosibirsk. – 2009. – Vyp. 1, T. 7. – S. 26-31.

8. Reznikov V. M. Ob adekvatnosti matematicheskoj statistiki // Materialy Vserossijskoj konferencii s mezhdunarodnym uchastiem «Znanija – Ontologii - Teorii». – Novosibirsk : Izd-vo RIC Prajc – Kur'er, 2013, T. 2., s. 106–111.

9. Cohen J. The Earth is Round (p<0,05) // American Psychologist. – 1994. – Vol. 49, № 12. – P. 997–1003.

10. Reznikov V. On Kolmogorov's analysis of applicability of probability theory // Sbornik materialov mezhdunarodnoj konferencii «Mal'cevskie chtenija». – Novosibirsk : Izd-vo Instituta Matematiki SO RAN, 2009. - s. 221.
Sajt Instituta Matematiki SO RAN [Jelektronnyj resurs]. URL: http://www.math.nsc.ru/conference/malmeet/09/Abstracts/abstracts-09 (data obrashhenija: 09.03.2013).

11. Shafer G., Vovk V. The Game-Theoretic Probability and Finance Project, Working Paper № 4. 2003. P. 11.
[Jelektronnyj resurs]. URL: http://www.probabilityandfinance.com (data obrashhenija: 03.02.2013).

12. Shafer G., Vovk V. Probability and Finance It’s Only a Game! N.Y.: A Wiley-Interscience Publication, 2001. – 414 p.

13. Shafer G., Vovk V. The Sources of Kolmogorov’s Grundbegriffe // Statistical Science. – 2006. – Vol. 21, № 1. P. 70–98.