

## МЕТОДОЛОГИЧЕСКИЕ ПРОБЛЕМЫ КОРРЕКТНОГО ПРИМЕНЕНИЯ ОБЪЕКТИВИСТСКИХ СТАТИСТИЧЕСКИХ КОНЦЕПЦИЙ

*В.М. Резников*

Показано, что современные статистические теории не являются ни эмпирическими, ни адекватными для приложений. Они не являются эмпирическими из-за использования принципа Курно, предполагающего априорное и детальное знание теоретических вероятностей, включая самые маленькие вероятности. Теории неадекватны из-за сложности решения проблемы объекта теории.

**Ключевые слова:** статистическая теория, объект теории, принцип Курно

В настоящее время имеется несколько объективистских эмпирических вероятностных и статистических концепций. Среди них фундаментальные вероятностные концепции Р. Мизеса, Г. Рейхенбаха, статистические концепции Р. Фишера и Ю. Неймана – Е. Пирсона. В основе объективистских стохастических концепций лежит представление о частоте, с которой наблюдаются изучаемые явления. Частота это естественное свойство объективного мира. В научной практике при изучении частотных закономерностей популярны методы математической статистики. Популярность статистики приводит к естественному вопросу: являются ли ее основания адекватными или эта популярность связана с модой на применение статистики в практике научных исследований? Проанализируем адекватность частотных объективистских концепций практике научных исследований.

Основными разделами математической статистики являются проверка гипотез и оценивание параметров. В основе фальсификации гипотез, исследуемой в стандартном статистическом анализе в разделе проверки гипотез, лежит принцип Курно. В статистике используется сильный вариант этого принципа. Согласно сильному варианту в единственном проведенном эксперименте невозможно появление события с ничтожной вероятностью. Корректной является слабая версия принципа Курно. По этой версии отличительная особенность маловероятных событий заключается в том, что они бу-

дуд редко происходить. Слабая версия допускает появление маловероятных событий, но у нее нет эффективной реализации. Принцип Курно в сильной форме не имеет, по сути, вероятностного духа, в лучшем случае он имеет вероятностно-детерминистский характер.

Принцип Курно не согласуется с эмпиризмом. Согласно Д. Гиллесу, формализация в рамках эмпирической концепции предполагает обязательное изучение естественных объектов, которые являются прообразами формальных объектов [1]. Например, для Р. Мизеса такими объектами были феномены статистической физики, данные демографических исследований и успешная практика функционирования игорных домов в Германии.

В рамках эмпирической традиции вероятностное теоретическое описание объекта неизвестно априори. Оно получается на основе формального описания свойств, которыми обладают эмпирические объекты. Теоретический объект выполняет функцию замещения полупэмпирических объектов в формальных рассуждениях. Применение принципа Курно в статистике основано на априорном и детальном знании теоретического объекта. Принцип Курно предполагает точное знание предельного распределения критериальной статистики вплоть до самых ничтожных вероятностей. По сути это вероятностно-детерминистская форма принципа всезнания Лапласа.

Корректное использование принципа Курно предполагает осуществление большой работы по исследованию событий с малыми вероятностями. Так, для того чтобы корректно определить вероятность события порядка  $10^{-n}$ , необходимо проведение хотя бы  $10^n$  экспериментов, но для надежной оценки требуется проведение порядка  $10^{n+1}$  экспериментов. В связи со сложностью точного определения вероятностей для маловероятных событий имеет смысл модифицировать стандартную аргументацию, обосновывающую фальсификацию гипотезы. Например, гипотеза опровергается, так как не произошло событие с достаточно большой и корректно определенной вероятностью. Или гипотеза не может быть принята, так как для ее принятия требуется знание вероятности события, которая не может быть корректно определена.

Формирование гипотез в статистическом анализе предполагает оценивание параметров для различных теоретических распределений. В основе раздела статистики под названием «оценивание параметров» лежит принцип максимального правдоподобия. Согласно этому принципу известным результатам эксперимента приписывает-

ся максимальная совместная плотность распределения. Это, по сути, реализация вероятностно-детерминистского принципа: то, что произошло, должно было произойти с максимальным правдоподобием. При применении объективистской концепции оценивания неизвестных параметров для одних и тех же данных двумя разными специалистами, естественно, должны быть получены одни и те же результаты. При обработке похожих данных должны получаться близкие результаты. В статистике это свойство называется робастностью. Однако современные методы оценивания параметров не являются робастными. Они обеспечивают неограниченную точность результатов, если данные адекватны используемой модели. Неробастность заключается в том, что если данные даже в минимальной степени отклоняются от требований модели, то ошибка вычислений является непредсказуемой и неограниченной.

Дух объективизма несовместим с отсутствием формальных предписаний относительно выбора конкретных критериев, так как выбор на основе полусодержательных соображений не является однозначным, потому что множество критериев адекватно для анализа конкретной задачи. Но применение различных критериев приводит к различным результатам [2]. Современная статистическая теория не может считаться объективистской из-за асимптотического характера критериев математической статистики. Асимптотическая теория не определяет минимальный объем данных, обеспечивающий сходимость распределения критериальной статистики к известным распределениям. Предлагаемые математической статистикой подходы к решению задач не имеют унифицированного характера. В наибольшей степени это относится к критериям, основанным на анализе группированных данных. Например, группирование используется в знаменитом критерии Пирсона хи-квадрат. Существует несколько принципов группирования: разбиение данных на интервалы равной длины или на равновероятные интервалы, в которые с одинаковой вероятностью попадают реализации случайной величины с заданным распределением. Теория группирования не является детально разработанной и не детерминирует выбор способа группирования для конкретных ситуаций. Рекомендации по выбору оптимального числа интервалов, минимального числа данных в интервале, способа оценивания параметров на основе группированных данных являются вариабельными. По мнению байесовистов, это сильный аргу-

мент против объективного характера стандартных статистических теорий Фишера и Неймана – Пирсона [3].

В математическом отношении частотные теории Фишера и Неймана – Пирсона являются корректными. Проблемы возникают в связи с применением этих теорий. Одним из важных понятий статистической теории оценивания является понятие минимального доверительного интервала. Минимальный доверительный интервал – это наименьший из возможных интервалов со случайными концами, в который с заданной вероятностью попадает искомое значение параметра распределения. Интервал со случайными концами является абстрактным понятием. В реальной ситуации границы области научного поиска описываются вещественными числами. В этом случае понятие произвольной вероятности попадания в доверительный интервал теряет смысл. Параметр или обязательно попадает в конкретный интервал, можно сказать, с единичной вероятностью, или обязательно не попадает в этот интервал, т.е. попадает в него с нулевой вероятностью.

Современные математические теории не являются универсальными, – они являются эффективными для анализа определенных данных, близких к теоретическому объекту. Это приводит к проблеме объекта теории. Проблема заключается в определении того, обладают ли результаты экспериментов важнейшими базовыми свойствами теоретических объектов. Самыми главными свойствами теоретических объектов являются закон распределения и (или) независимость данных. Математические методы теоретической статистики из-за множества нерешенных философско-методологических проблем, таких как проблема индукции, недоопределенность теоретического знания с помощью конечного объема эмпирических данных, а также вследствие неробастности, неунифицированности статистических процедур и асимптотического характера критериев не обеспечивают решение проблемы объекта.

Возможно ли решение этой проблемы объекта вне математической статистики? Иногда фундаментальным теоремам теории вероятностей, например теореме закона больших чисел, приписывается эпистемологический статус, так как считается, что они обеспечивают решение проблемы объекта, обосновывая близость частот и теоретических вероятностей. Эвристическая значимость теоремы переоценивается. Теорема определяет близость частотных характеристик данных к априори известной теоретической вероятности. В лучшем случае она адекватна для верификации близости частот-

ных характеристик и теоретической вероятности. Теорема утверждает о близости этой вероятности к эмпирическим частотам, доказывая, что предельная вероятность небольших отклонений априорно заданной вероятности от эмпирических частот равна единице. Практическое применение теоремы к выборкам конечного объема данных предполагает частотное оценивание вероятности сложного события, описывающего отклонение априорной вероятности от эмпирических частот. Для определения качества найденных эмпирических частот, т.е. близости этих частот и априорно известной (или вычисленной некоторым образом) вероятности сложного события, необходимо использовать уже третью теоретическую вероятность. Теорема построена на увеличении числа теоретических величин, и поэтому она не может иметь прагматическую значимость.

Неадекватность теоремы для приложений отмечают ряд исследователей, например А.Н. Колмогоров, А.И. Орлов, Ю.И. Алимов. Однако если утверждение не соответствует закону больших чисел, то оно вообще не может быть исследовано с помощью вероятностной математики. Поэтому минимально обоснованное применение математики предполагает адекватность данных предельным теоремам, но для реального применения предельные теоремы неадекватны из-за их неконструктивного характера.

Невозможность решения проблемы теоретическими средствами привела к появлению статистических концепций, альтернативных математической статистике и не использующих свойства теоретических объектов. Примером является метрологическая концепция статистики. Метрологическая концепция была разработана Ю.И. Алимовым и П.С. Эльясбергом [4]. В концепции выделяются критическая и позитивная составляющие. Критика направлена на принятие формальных базовых свойств моделей на основе неформальных рассуждений, таких как контроль фоновых условий проводимых экспериментов, а также интуитивные соображения, например привлечение максимально большого объема данных – по принципу «кашу маслом не испортишь» с целью получения точной статистической оценки. Однако надежный контроль условий экспериментов за пределами физических исследований практически невозможен. Даже в физике периодически возникают проблемы с контролем фоновых условий. Известно, что при открытии радона в лаборатории Резерфорда показания приборов вели себя хаотическим образом. Дело в том, что не сразу догадались, что новое вещество является газооб-

разным. После этого условия проведения эксперимента были рафинированы, и тогда приборы стали вести себя адекватно.

Принцип «чем больше данных, тем лучше» на первый взгляд кажется совершенно естественным и адекватным статистике из-за асимптотического характера статистических критериев. Однако этот принцип адекватен исключительно для идеальных, корректных данных. Критерий до сих пор сохранил научную респектабельность в тех областях знания, где критерий, по существу, является абстрактным, так как в этих областях много данных не бывает. Например, это относится к биологическим исследованиям. Принцип «кашу маслом не испортишь» наиболее критикуем в технических науках, где получение новых данных не представляет особой проблемы. Так, например, при работе с новой, еще не отлаженной техникой проблема заключается в выявлении оптимального минимального числа экспериментов для корректного определения статистических характеристик прибора. Аналогично проблема установления минимального оптимального объема данных возникает при работе с некачественными данными. При превышении некоторого критического объема данных точность не увеличивается, а может только ухудшиться.

При работе с большими объемами данных интуитивно корректные и вполне адекватные решения для данных небольшого объема оказываются некорректными и приводят к грубым ошибкам. Так, например, во многих науках анализ данных осуществляется на основе корреляционного анализа. Невысокие корреляции, скажем порядка одной сотой, обычно не учитывают. Однако для выборок большого объема получение объективных результатов невозможно при игнорировании даже самых низких корреляций. В работе П.С. Эльясберга [5] проведен сравнительный статистический анализ двух выборок, каждая объемом 1000. В первом случае все данные считались некоррелированными, а во втором учитывался коэффициент корреляции 0,01. В обоих случаях полагалось, что данные связаны линейной зависимостью. Статистические оценки для этих двух случаев оказались в существенной степени разными. Так, для некоррелированных данных величина дисперсии оказалась меньшей на порядок по сравнению с дисперсией для минимально коррелированных данных.

Особенностью позитивной части метрологической концепции является доминирование полуэмпирических величин над теоретическими величинами. Теоретические величины не существуют априори, – они существуют, если соответствующие полуэмпирические ве-

личины являются устойчивыми. Определение устойчивости оценок – приоритетная проблема для корректного использования статистических методов в технических науках, где ценится точность результатов. Кроме того, одной из приоритетных задач является получение точных прогнозных решений, и чем более устойчивыми оказываются статистические оценки, полученные на большом экспериментальном материале, тем больше оправданно использование данных с целью получения обоснованных прогнозных решений.

Другой особенностью метрологического подхода является минимальное требование к знанию базовых статистических свойств. Это требование связано с рядом обстоятельств. Во-первых, обычно базовые свойства неизвестны априори. Во-вторых, из-за неробастности интерес представляет только точное знание базовых свойств. Однако решение этой проблемы является более сложным, чем решение любой непосредственной проблемы. В-третьих, для ряда статистических проблем определение устойчивых статистических оценок и их верификация возможны на основе полуэмпирических методов без знания базовых теоретических свойств. Получение устойчивой статистической оценки предполагает ее многократное измерение. Поэтому в отличие от методологии максимального правдоподобия, предполагающей использование всех данных для получения точной оценки, метрологическая парадигма при каждом вычислении использует только часть имеющихся данных. При применении метрологического подхода следуют принципу повторяемости вычислений, производимых на разных данных. Для этого осуществляется разбиение данных объема  $n$  на  $k$  одинаковых по объему групп. Для каждой группы вычисляется нужная статистическая характеристика, например среднее. В результате определяются полуэмпирические минимальные и максимальные значения средней величины. Интервал, граничные точки которого составляют минимальное и максимальное средние представляет собой невероятностный эмпирический доверительный интервал.

В отличие от вероятностного доверительного интервала для невероятностного интервала граничные точки являются не случайными величинами, а обычными вещественными числами. Точность измерений определяется величиной интервала. Для проверки устойчивости найденной интервальной оценки известно два подхода, различающихся объемом доступных данных. При большом объеме  $N$  доступных для верификации данных если  $N \gg n$ , то тогда новые данные

разбиваются на число групп  $K \gg k$ . Если все новые средние попадают в старый доверительный интервал, то тем самым осуществлена убедительная верификация устойчивости доверительной оценки среднего. Если некоторые новые средние выходят за границы старого интервала, то осуществляется модификация границ. В случае если данных для верификации немного, то верификация осуществляется на основе совместного использования старых и новых данных, при этом минимальное число групп должно быть больше или равно двум. Несмотря на простоту используемых средств, данный подход адекватен в ряде областей, где неизвестны законы распределения данных, например в социологии. Кроме того, он адекватен в технических областях, где ошибку измерений принято представлять не мысленным доверительным вероятностным интервалом, а с помощью реального невероятностного доверительного интервала, полученного посредством измерений и (или) вычислений.

Итак, нами показано, что использование принципа Курно в стандартных статистических концепциях Фишера и Неймана – Пирсона не позволяет считать их эмпирическими в строгом смысле, – скорее это квазиэмпирическая теория. В связи же с непреодолимыми сложностями решения проблемы объекта на базе этих теорий нет оснований считать их адекватными для корректного применения. Сложности корректного применения теоретической статистики привели к разработке и использованию методов прикладной статистики, в частности метрологической теории. Применение измерительных подходов и методологии компьютерного моделирования в статистике знаменует этап доминирования реальных эмпирических подходов в статистике. Однако в последнее время паритет эмпирических подходов в статистике нарушен новой игровой концепцией Г. Шафера и В. Вовка [6]. Они сумели на основе игрового варианта теории мартингалов получить конструктивные доказательства фундаментальных теорем теории вероятностей, адекватные для применения.

### Примечания

1. См.: *Gillies D.* Philosophical theories of probability. – L.: Routledge, 2003.
2. См.: *Hacking I.* Logic of statistical inference. – Cambridge: Cambridge Univ. Press, 1965.
3. См.: *Howson C., Urbach P.* Scientific reasoning: The Bayesian approach. – Illinois: Open Court, 1996.



4. См.: Алимов Ю.И. Альтернатива методу математической статистики. – М.: Знание, 1980; Эльясберг П.С. Вычислительная информация: Сколько ее нужно? Как ее обрабатывать? – М.: Наука, 1986.

5. См.: Эльясберг П.С. Вычислительная информация...

6. См.: Shafer G, Vovk V. Probability and finance: It is only a game. – N.Y.: Wiley-Interscience Publication, 2001.

Институт философии и права  
СО РАН, г. Новосибирск

**Rezников, V.M. Methodological problems of correct application of objectivist statistical conceptions**

The paper proves that modern statistical theories are neither empirical nor valid for correct application. They are not empirical through the use of Cournot's principle which presupposes prior and detailed knowledge of theoretical probabilities including negligible ones. Theories are not valid because of difficulties in solving the theory object problem.

**Keywords:** statistical theory, object of theory