ОБЪЕКТ ЭМПИРИЧЕСКОЙ СТАТИСТИЧЕСКОЙ ТЕОРИИ КАК МЕТОДОЛОГИЧЕСКАЯ ПРОБЛЕМА

В.М. Резников

Проблема объекта теории является методологической многоаспектной проблемой. *Во-первых*, это проблема развития теоретического знания, потому что в определенные моменты времени жизни теории ее развитие в той или иной степени опирается на учет ее объекта. Проблема объекта не возникает, как правило, при решении конкретных теоретических проблем. Эта проблема не возникает постольку, поскольку считается, что исследуемый объект совпадает с объектом теории.

Проблема объекта в контексте развития теоретического знания подобна проблеме определения сущностных признаков понятия в классической формальной логике. Решение этой проблемы предполагает проведение формального и содержательного анализа теоретического знания с целью определения фундаментальных результатов. Далее фиксируются свойства объекта, требующиеся для получения признанных результатов теории. Анализ избыточности используемых свойств, а также анализ полноты найденных свойств имеют первостепенное значение для развития теоретического знания.

Во-вторых, проблема объекта теории – это проблема корректного применения теории. В наибольшей степени теория адекватна для анализа собственно объектов теории. Проблема корректного применения теории предполагает решение ряда задач. Это качественные и количественные проблемы определения адекватности приложений. Качественный аспект проблемы предполагает определение допустимости репрезентации теоретических объектов с помощью исследуемых данных. Количественный аспект проблемы предполагает определение степени соответствия исследуемого объекта объекту теории.

Наша работа посвящена количественным аспектам проблемы объекта в байесовской концепции, стандартных частотных концепциях, теории вероятностей и метрологической концепции в контексте корректного применения этих теорий. Проблема определения сущностных свойств теорий излагается в той степени, в какой это требуется для анализа проблемы корректных приложений.

Проблема анализа сущностных свойств объектов вероятностных и статистических теорий

Сущностные свойства вероятностных и статистических теорий нами названы базовыми свойствами [1]. Базовые свойства удовлетворяют двум требованиям. Во-первых, они используются при доказательстве фундаментальных результатов. Во-вторых, они логически невыводимы на основе использования других свойств объектов теории. Такими свойствами являются концепция вероятности, закон распределения вероятностей в любой вероятностной концепции и свойство независимости. Для статистических теорий дополнительно к вышеназванным свойствам добавляется свойство однородности данных. Принятие определенной концепции вероятности детерминирует в целом вероятностную теорию. Значимость свойства независимости для развития теории вероятностей убедительно показал А.Н. Колмогоров [2].

По отношению к вероятностным и статистическим теориям выделяются три варианта количественного анализа объекта теории. В первом анализируется степень близости (удаленности) теоретических объектов. Во втором варианте анализ применяется для определения близости (удаленности) теоретического и полуэмпирического объектов. И наконец, в последнем варианте оба исследуемых объекта являются полуэмпирическими. Под теоретическими объектами понимаются вероятности и функции распределения вероятностей, а под полуэмпирическими – эмпирические средние, дисперсии и эмпирические функции распределения вероятностей.

Проблема объекта для определения степени близости полуэмпирического и теоретического объектов

Это наиболее общий случай из трех выделенных вариантов анализа, так как он включает все возможные варианты объектов. В теории вероятностей степень близости между теоретическим и полуэмпирическим объектами определяется на основе так называемых фундаментальных теорем, в частности теоремы закона больших чисел. Для целей статьи достаточен простой вариант теоремы закона больших чисел.

Теорема. Пусть вероятность успеха в каждом независимом испытании одинакова и равна p. Среди n испытаний оказалось m успешных. Тогда при неограниченном росте числа испытаний n оказывается, что

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

Недостатки теоремы для анализа проблемы объекта

1. Теорема не является универсальной. Она доказана для независимых и слабо связанных событий.

2. Теорема исследует вероятность отклонения вероятности успеха от частоты успеха при увеличении объема данных. Отклонение частоты успеха от вероятности успеха определяется с помощью следующего выражения:

$$|m/n - p \qquad (2)$$

Если многократно проведенные исследования отклонения частоты от вероятности с помощью формулы (2) показывают, что это отклонение регулярно меньше допустимой погрешности эксперимента, то определение вероятности отклонения не имеет смысла. Очевидно, что вероятность малых отклонений будет высокой.

Иными словами, исследование отклонения теоретической величины от полуэмпирической величины с помощью другой теоретической величины не имеет прагматической значимости.

3. В теореме вероятность успеха задана априори. Поэтому теорема предназначена только для верификации или фальсификации известной заранее вероятности, а не для ее определения.

Достоинства теоремы

1. Если выполняются требуемые в теореме условия, то она позволяет точно определить предельную вероятность отклонения частоты от вероятности.

2. В недавно предложенном варианте теоремы закона больших чисел для игровой концепции теории вероятностей эта теорема

обнаружила новые теоретические и прагматические достоинства [3]. В стандартном случае теорема закона больших чисел выполняется всюду, за исключением множества меры нуль. В игровом варианте она выполняется всегда. Игровой вариант теоремы обнаружил адекватность для принятия решений в экономических моделях. Оказалось, что игроки, выбирающие ставки, в соответствии с теоремой в подавляющем числе случаев оказывались победителями.

Проблема объекта в контексте статистических теорий Фишера и Неймана – Пирсона

Качественный вариант проблемы объекта. Качественный вариант проблемы объекта связан с гипотезой о генеральной совокупности. Согласно этой гипотезе сформированные результаты экспериментов представляют собой выборку, взятую из гипотетически существующей генеральной совокупности. Принятие гипотезы о генеральной совокупности означает, что выборка является репрезентативной, для того чтобы выводы, полученные с ее помощью, оказались верными для популяции в целом. Кроме того, принятие гипотезы означает, что данные являются однородными. Если данные не являются однородными, то это обстоятельство необходимо учитывать путем нивелирования неоднородностей, например с помощью методов рандомизации. Как убедительно показано Дж. Уорреллом [4], рандомизация уменьшает влияние неоднородности, но не позволяет оценить степень остаточной неоднородности.

Количественные подходы к проблеме объекта. Проблема объекта является методологической проблемой, а не математической, поэтому эта проблема не выделена в рамках математической статистики в самостоятельное направление. Тем не менее ряд разделов математической статистики, например оценивание параметров и проверка гипотез, в определенной степени адекватны для определения близости полуэмпирических и теоретических величин, а значит, для анализа проблемы объекта.

Критика методов оценивания параметров

Наиболее мощным методом оценивания параметров является метод максимального правдоподобия. Он основан на принципе максимального правдоподобия. Согласно этому принципу имеющиеся данные представляют собой независимые случайные величины и их совместная плотность вероятности является максимальной.

1. Принцип адекватен не для вероятностных экспериментов, а для вероятностно-детерминистских экспериментов, поскольку все, что произошло, имело максимальную вероятность возникновения.

2. В связи с этим принципом метод максимального правдоподобия не является универсальным, он адекватен только для независимых экспериментов.

3. Критерии качества оценивания, например, состоятельность, являются асимптотическими. В соответствии с асимптотическим критерием при неограниченном числе экспериментов найденная оценка по вероятности сходится к искомому параметру распределения. Критерий не позволяет определить степень близости оценки параметра и искомого параметра для выборок конечного объема данных.

4. Метод максимального правдоподобия не является робастным. При небольшом отклонении данных от модельных предположений получаемые оценки будут сильно отличаться от искомого параметра.

5. Метод не позволяет сравнивать полученные оценки для разных выборок с целью определения наилучшей оценки.

Достоинство методов оценивания параметров

Если данные соответствуют модельным предположениям(требованиям о независимости, однородности данных и др.), то метод обеспечивает получение значений параметра с практически неограниченной точностью.

Критика методов проверки гипотез

1. Современные методы проверки гипотез построены на принципе Коорнота. Согласно этому принципу в однократно проведенном эксперименте маловероятное событие считается физически невозможным. С логических позиций принцип некорректен. Маловероятные события являются редкими, а не невозможными, и они могут осуществиться в результате любого эксперимента, в том числе и первого.

2. Корректная проверка вероятности события предполагает определение частоты этого события. Для маловероятных событий эта проверка

является чрезвычайно сложной. Так, для определения частоты события порядка 10^{-n} требуется проведение порядка 10^{n+1} экспериментов.

3. Выбор статистического критерия не может быть обоснован. Применение разных критериев может приводить к противоречащим результатам.

Достоинство методов проверки гипотез

Если проверяемая гипотеза задана корректно и не предполагает оценивания параметров, то гипотеза будет принята практически безошибочно.

От анализа проблемы объекта, исследуемой путем анализа полуэмпирической и теоретической величин, перейдем к анализу этой проблемы для теоретических величин.

Проблема объекта в контексте анализа проблемы подтверждения гипотез

В байесовском анализе исследуется проблема подтверждения гипотез. Возникает вопрос об анализе проблемы подтверждения теоретических гипотез в контексте проблемы объекта. В случае успешной интерпретации проблемы объекта с помощью аппарата, используемого для анализа проблемы объекта, будет обнаружена связь между проблемами подтверждения и фальсификации гипотез. Так как проблема подтверждения гипотез является типичной проблемой байесовского анализа, а проблема фальсификации гипотез – характерной проблемой стандартных частотных концепций, тем самым будет определена связь между байесовской и частотными парадигмами, считающимися непримиримыми оппонентами.

В байесовском подходе степень подтверждения гипотез определяется с помощью различных вероятностных мер. На примере одной меры покажем определение степени подтверждения гипотезы.

Пусть H – обозначение гипотезы, а D – это суждение, описывающее результат эксперимента. Пусть $P(H)$ – это априорная вероятность правильности гипотезы, $P(D)$ – вероятность появления данных, а $P(D/H)$ – это вероятность появления экспериментального результата при условии справедливости гипотезы. Тогда

$$P(D/H) = P(D/H) \cdot P(H))/P(D).$$

Данные в наибольшей степени согласуются с гипотезой, если они представляют объект теории. В этом случае $P(D/H)$ будет близко к величине $P(D)$, тогда величина $P(D/H)$ будет близка к величине $P(H)$. В рассматриваемом случае степень подтверждения гипотезы определяется следующим образом:

$$\mathrm{S_t} = P(H/D) - P(H). \tag{4}$$

Гипотеза будет подтверждена, если выражение (4) принимает минимальное значение.

Сложность использования формализаций в контексте анализа проблемы объекта связана с множественностью мер. В настоящее время наиболее обоснованными и популярными в байесовском анализе [5] являются следующие пять мер:

$$d(H, D/K) = P(H/D \wedge K) - P(H/K); \tag{5}$$

$$r(H, D/K) = \log(P(H/D \wedge K))/\log P(H/K); \tag{6}$$

$$l(H, D/K) = \log((P(E/H \wedge K))/P(E/\neg H \wedge K)); \quad 7)$$

$$s(H, D/K) = P(H/D \wedge K) - P(H/\neg E \wedge K); \tag{8}$$

$$\tau(H, D/K) = P(H \wedge D \wedge K)P(K) - P(H \wedge K)P(E \wedge K); \tag{9}$$

Обозначения H и D определены нами ранее, а K – это суждение, описывающее фоновые условия. Для корректного решения проблемы многообразия мер, а следовательно, и проблемы объекта существует несколько подходов. Первый подход основан на доказательстве адекватности применения определенной меры для решаемой задачи. Как правило, такого рода доказательство не удается получить. Второй подход базируется на идее инвариантности мер. Инвариантность обосновывается путем анализа свойств меры, которая применяется для анализа проблемы объекта. Если используются свойства, которыми обладают все меры, то это и будет доказывать нечувствительность проблемы объекта к используемой мере. Как показано в ряде работ [6], решение многих проблем, предполагающих решение проблемы объекта, связано с использованием таких свойств мер, которыми другие меры не обладают.

Так, например, в реконструированном Д. Гиллесом [7] варианте совместной работы К. Поппера и Д. Миллера [8]. доказательство невозможности индуктивной вероятности основано на свойстве аддитивности меры d:

$$d(H, D/K) = d(H \vee D, D/K) + d(H \vee \neg D, D/K). \qquad (10)$$

Возникает вопрос: является ли доказательство Гиллеса инвариантным по отношению к используемой мере? Ответ на этот вопрос отрицательный, потому что в работах М. Редхеда и Б. Фительсона показано, что соответственно мера r и мера l не обладают свойством аддитивности.

Кроме свойства аддитивности мера d обладает рядом других свойств, которыми обладают не все меры. Пусть на основе гипотезы H получена корректность факта D. Тогда верными по отношению к мере d являются следующие соотношения:

$$d(H \wedge X, D/K) = P(X/H \wedge K)\, d(H, D/K); \qquad (11)$$

$$d(H \wedge X, D/K) < d(H, D/K). \qquad (12)$$

Этими свойствами не обладают меры l и s. В свою очередь, мера r обладает свойством, которым обладают не все из пяти рассмотренных мер. Пусть на основе гипотезы H получена справедливость факта D. Тогда имеет место следующее равенство:

$$r(H, D/K) = r(H \wedge X, D/K), \text{ для любого } X. \qquad (13)$$

Как доказано Р. Розенкранцем [9], свойством (13) не обладают ни мера d, ни мера l. Другие меры тоже обладают уникальными свойствами. Поэтому попытки обосновать инвариантность любой из пяти рассмотренных мер в контексте решения проблемы объекта оказались неудачными.

Третий подход заключается в построении обобщенной меры, обладающей всеми свойствами, которыми обладают любые релевантные меры. Мера, в наибольшей степени обладающая свойствами обобщенной меры, и будет наилучшей.

Наибольший вклад в развитие этого направления был сделан П. Милном, Р. Карнапом и И. Гудом [10]. Милн показал, что среди известных мер в качестве обобщенной меры адекватна мера r. Сформулированные им требования формально корректны, но не все имеют методологическую убедительность. Самым неубедительным оказалось следующее требование:

$$\text{Если } P(D/H \wedge K) = P(D/H1 \wedge K), \text{ то } c(H \wedge D/K) = c(H1 \wedge D/K). \quad (14)$$

Как показано Б. Фигельсоном [11], это требование, вопреки предположению Милна, не связано с принципом правдоподобия, а кроме того, этому требованию удовлетворяет только единственная мера r. Поэтому требования Милна не являются убедительными, и соответственно мера r не есть наилучшая.

Р. Карнап обосновывал тезис, согласно которому на роль обобщенной меры подходит его собственная мера τ. Подход Карнапа базируется на предположении, что искомая мера является симметричной. Симметричные ситуации часто встречаются, но ими невозможно ограничиться, потому что несимметричные ситуации еще более часты. Пусть гипотеза H означает, что из колоды вытащили карту черной масти. Выбор карты описан с помощью суждения D, означающего, что вынули девятку треф. Суждение D подтверждает H. Для симметричных мер переход от свидетельства к противоположному суждению является не менее информативным. В данном случае переход к суждению «не девятка треф» оказывается непродуктивным. Являясь симметричной, мера Карнапа не является универсальной.

Последний, четвертый, подход связан с И. Гудом. Сформулированной им системе требований адекватна мера l. Одно из требований Гуда – это независимость обобщенной меры от априорных знаний об объекте. Как показано в работе Б. Фигельсона [12], это требование оказалось невыполненным. Тем не менее мера l в настоящее время считается наиболее серьезным претендентом на роль обобщенной меры. Фительсон показал, что наибольшее число проблем, связанных с решением проблемы подтверждения знания, удалось решить на основе использования меры l [13].

В наибольшей степени в концепции проверки гипотез используется мера d. Если применение этой меры является обоснованным в решении задачи из области подтверждения гипотез, то в этом случае можно говорить о сущностной связи проблем фальсификации и подтверждения гипотез. По отношению к остальным вероятностным мерам сущностную связь проблем фальсификации и подтверждения, по крайней мере пока, не удалось установить.

Философская мотивация поиска наиболее обоснованной меры не является состоятельной. Предположим, что удалось сконструировать меру, с помощью которой можно решить все известные проблемы, связанные с подтверждением гипотез. Это не гарантирует того, что данная мера адекватна для анализа новых проблем. Кроме того, существующие меры не учитывают влияние фоновых условий и поэтому не способны определить степень подтверждения гипотезы в реалистичных ситуациях.

Проблема объекта для полуэмпирических величин

Многочисленные сложности и отсутствие перспектив, связанных с решением проблемы объекта для полуэмпирических и теоретических величин в теории вероятностей и в стандартных частотных концепциях, показывают, что теоретические методы неадекватны для решения проблемы объекта теории. Неадекватность современных теоретических методов для проблемы объекта подтверждается также выявленными сложностями решения этой проблемы в субъективистской концепции вероятности на основе решения проблемы подтверждения гипотез. Неудачи теоретических методов послужили мотивацией к созданию концепции, связанной с решением проблемы объекта для полуэмпирических величин. Такой концепцией является метрологическая концепция.

Метрологическая концепция [14] решает вариант проблемы объекта для полуэмпирических устойчивых статистических характеристик. Эта концепция свободна от недостатков теоремы закона больших чисел. Она является универсальной, здесь не требуется знания базовых характеристик, например наличия независимости. Методологический подход дает средства для вычисления теоретической величины, а не только обеспечивает проверку близости существующей априори теоретической величины и величины эмпирической. В нем не привлекаются теоретические величины для исследования близости сравниваемых величин.

В отличие от методологии Фишера и Неймана – Пирсона в этом подходе не используются критикуемые нами асимптотические критерии качества оцениваемых параметров, принцип Коорнота, провозглашающий невозможность событий с малыми вероятностями, и предположения о знании распределения данных. Не используются также предположения о свойствах выборки.

Методологический подход свободен от проблемы множественности мер в субъективистской теории вероятностей, потому что используется хорошо мотивированная мера различия для полуэмпирических величин. Данный подход, в существенной степени решая проблему объекта, является пропедевтикой к корректному использованию других концепций.

Примечания

1. См.: *Резников В.М.* Вероятностные концепции: анализ оснований и приложений. – Новосибирск: Новосибирск. гос. ун-т, 2005.

2. См.: *Колмогоров А.Н.* Основные понятия теории вероятностей. – М.: Наука, 1974.

3. См.: *Shafer G., Vovk V.* Probability and finance: It is only a game. – N.Y.: Wiley, 2001.

4. См.: *Worrall J.* What is evidence in evidence-based medicine? // Philosophy of science. – 2002.

5. См.: *Fitelson B*. The plurality of Bayesian measures of confirmation and the problem of measure sensitivity // Philosophy of Science. – 1999. – No. 66. – P. 362–378.

6. См.: *Fitelson B*. The plurality of Bayesian measures of confirmation…; *Redhead M*. On the impossibility of inductive probability // The British Journal of Philosophy of Science. – 1985. – No. 36. – P. 185–191.

7. См.: *Gillies D*. In defense of the Popper – Miller argument // Philosophy of science. – 1986. – No. 54. – P. 110–113.

8. См.: *Popper K., Miller D*. The impossibility of inductive probability // Nature. – 1983. – V. 302. – P. 687–688.

9. См.: *Rosenkrantz R*. Foundations and applications of inductive probability. – Atascadero, Calif., 1981.

10. См.: *Milne P*. log[$p(h/eb)/p(h/b)$] is the one true measure of confirmation // Philosophy of Science. – 1996. – No. 63. – P. 21–26; *Carnap R*. Logical foundations of probability. – Chicago: Univ. of Chicago Press, 1962; *Good I*. The best explicatum for weight of evidence // Journal of Statistical Computation and Simulation. – 1984. – No. 19. – P. 294–299.

11. См.: *Fitelson B*. Studies in Bayesian Confirmation Theory. – Univ. of Wisconsin, 2001.

12. *Ibid*.

13. *Ibid*.

14. См.: *Алимов Ю.И.* Альтернатива методу математической статистики. – М., 1980.

Институт философии и права
СО РАН, г. Новосибирск

Reznikov, V.M. The object of empirical statistic theory as a methodological problem

The paper discusses quantitative aspects of the problem of object in Bayes conception, standard frequency conception, probability theory and metrological conception in respect to correct use of these theories. The problem of finding of essential properties of theories is stated to the extent which is necessary to analyze the problem of correct applications.