ПРОБЛЕМА СТАТИСТИЧЕСКОЙ НЕЗАВИСИМОСТИ И ПРИНЦИП МАКСИМАЛЬНОЙ ЭНТРОПИИ*

Е.Е. Витяев, В.М. Резников

В отличие от стандартного статистического анализа, в котором корректность применяемого аппарата обосновывается формальным образом, в байесовизме формальные описания обосновываются с помощью интуитивных и содержательно-рациональных аргументов. В статье показано, что принцип максимальной энтропии, используемый в байесовизме для вычисления совместных вероятностей независимых событий, не имеет универсального характера. Он адекватен для решения проблем, описываемых с помощью неполных графов с небольшим числом вершин, когда максимизация энтропии имеет смысл. Для решения реальных задач, в которых учитывается неопределенность знания, в области ответственных приложений адекватным является комбинирование логического программирования и статистического анализа данных.

Ключевые слова: независимость, марковское свойство, энтропия, принцип максимальной энтропии

Стохастические модели в наибольшей степени эффективны, когда имеются основания для интерпретации моделей как независимых экспериментов. Независимость значима как для анализа теоретических моделей, так и для применения прикладных моделей в практике научных исследований. А.Н. Колмогоров убедительно показал значимость независимости для возникновения аналитического аппарата теории вероятностей, представляющего собой мост между миром исследуемых явлений и миром вероятностной математики. В теории вероятностей таким мостом служат фундаментальные теоремы теории вероятностей, в частности теорема закона больших чисел. А.Н. Колмогоров отмечал, что первые результаты, относящиеся к закону больших чисел и центральной предельной теореме, были получены с использованием предпосылки о независимости исследуемых данных. Он говорил, что в современных исследованиях с целью обобщения классических результатов отказываются от

* Работа выполнена при финансовой поддержке РФФИ (проект N 08-07-00272-а) и Междисциплинарного интеграционного проекта СО РАН № 47.

модели независимых испытаний, однако где-то в глубине идея независимости обнаруживается [1].

Некоторые крупные математики, в частности Г. Крамер, недооценивали значимость независимости. В своей знаменитой книге Крамер отмечал, что центральная предельная терема верна для случая, когда случайные величины не обязательно являются независимыми [2]. Однако это не означает, что они произвольно зависимые. На самом деле все известные доказательства центральной предельной теоремы получены не для произвольно зависимых случайных величин, а для слабозависимых случайных величин. К слабозависимым случайным величинам относятся марковские процессы (описывающие зависимость вероятности от факторов, действующих в настоящий момент времени, и независимость этой вероятности от прошлого) и мартингалы (описывающие зависимость математического ожидания от факторов, действующих в настоящий момент времени, и независимость от будущих факторов). По Колмогорову независимость важна не только для развития теории вероятностей, но и для того, чтобы эта теория стала самостоятельной математической дисциплиной. Именно благодаря особому интересу к анализу независимости теория вероятностей отделилась от абстрактной математической дисциплины – теории меры и стала самостоятельной наукой.

Независимость значима не только для развития и применения теории вероятностей, но и для корректного и эффективного использования моделей математической статистики. Действительно, идея независимости является центральной в математической статистике. Так, самый мощный и универсальный принцип оценивания параметров – это принцип максимального правдоподобия, предполагающий независимость наблюдений. Знаменитые критерии проверки гипотез – хи-квадрат Пирсона, критерий Колмогорова – Смирнова предполагают независимость данных. Однако существуют принципиальные сложности для введения и описания независимости в статистических моделях.

В отличие от теории вероятностей, где независимость вводится однозначно, в математической статистике не существует средств единообразного описания независимости. Независимость определяется контекстуально. Так, например, для описания и нормально распределенных данных, и линейно зависимых данных адекватен коэффициент корреляции. Для описания переменных, имеющих совместную невысокую вероятность, адекватно распределение Пуассона. Отсутствие единообразия ставит проблему объективности статистического метода. Корректное определение независимости предполагает формальное описание независимости и формальную проверку адекватности полученного формального описания.

С философско-эпистемологических позиций проверка гипотез не является в полной мере корректной. Проверка гипотез основана на принципе Курно. Существуют два варианта этого принципа: сильный и слабый. В математической статистике получил развитие сильный принцип Курно. По сильному принципу маловероятные события считаются физически невозможными. Сильный принцип Курно отвергает случайность и не согласуется с индетерминизмом. Учитывая значимость независимости и сложности ее описания и проверки, А.Н. Колмогоров полагал, что полностью определить условия, при которых необходимо принимать модель независимых экспериментов невозможно, оставаясь в рамках математики, и что эта проблема относится к области философии.

Для того чтобы подчеркнуть значимость и сложность формального описания и проверки независимости формальным образом, нами введено понятие базового свойства математической дисциплины [3]. Свойство называется базовым, если оно удовлетворяет следующим условиям. Во-первых, оно используется для получения фундаментальных результатов определенной математической дисциплины. Во-вторых, это свойство в данных не может быть получено на основании других свойств или требует знания полного теоретического распределения данных. В-третьих, это свойство в данных не может быть получено унифицированным образом [4].

А.Н. Колмогоров только отметил значимость философского анализа для определения независимости, но не указал принципы и методы философии, адекватные для описания независимости. Можно предположить, что один из этих принципов предполагает апеллирование к интуиции. Действительно, во-первых, уже начиная с работ Р. Декарта роль интуиции в математике трудно переоценить. Во-вторых, интуиция играет ключевую роль в созданном Я. Брауэром и другими математиками интуиционизме. В-третьих, другие современные направления математики и философии математики, в частности структурализм, являются вполне интуитивными.

Ввиду описанных нами сложностей, возникающих при формальном описании и проверке независимости, в практике научных исследований независимость часто вводится на основе интуитивно-содержательного анализа. Например, считают, что если в интуитивно-содержательном плане результаты экспериментов являются независимыми и наблюдения осуществляются при контроле над фоновыми условиями, то адекватна модель независимых наблюдений. Однако полный контроль фоновых условий даже в физике не всегда осуществим, а за пределами физики он практически невозможен [5]. Кроме того, в стохастической математике множество результатов имеют статус неинтуитивных. К ним относятся парадокс агрегации Симпсона, парадокс о зависимости в вероятностном смысле непересекающихся множеств. Вполне парадоксальным считается результат о принципиальном различии (более чем на порядок) дисперсии суммы независимых случайных величин от дисперсии суммы малозависимых случайных величин, имеющих коэффициент корреляции 0,01, когда число переменных достигает 1000 [6]. В математике факт вероятностной парадоксальности известен, несколько монографий посвящено этой проблеме [7]. Независимость играет важную роль практически во всех известных вероятностных интерпретациях: от частотной эмпирической интерпретации Р. фон Мизеса до чисто субъективистской интерпретации Б. де Финетти. Кроме того, свойство независимости является значимым для современных высоких статистических технологий, а также для экспертных систем.

Современные экспертные системы основаны на байесовской статистической концепции. Аналитический аппарат байесовизма, используемый в экспертных системах, – это синтез теории графов и теории вероятностей. Граф – это совокупность двух множеств: V и E. V является множеством вершин. Вершины интерпретируются как некоторые события. E является множеством связей. Мы ограничимся анализом направленных графов. В этих графах движение между вершинами осуществляется по выделенным направлениям. Разрешимые направления изображаются дугами со стрелками, например $A \to B$. Для описания графов естественно использование семейных отношений. Так, например, если имеется непосредственное ребро, идущее от вершины A к вершине B, то A называется родителем, а B – ребенком. Если имеется путь (последовательность ребер), ведущий от вершины A к вершине B, то A называется предком, а B – потомком. Если нет пути, ведущего от A к B, то A и B не составляют отношения прародители – потомки. Тогда множество B представляет неродственников для множества A. Независимость на графах определяется следующим образом. Граф обладает свойством марковости (независимости), если любая его вершина в вероятностном смысле условно независима от множества неродственников, где в качестве условия взяты родители данной вершины [8]. Вероятностная независимость в байесовской концепции основана ни на формальном анализе, ни на основе эмпирических измерений, она вводится на основе интуитивных и рациональных рассуждений.

В метрологической статистической концепции формальный и эмпирический подходы доминируют над интутивно-содержательными соображениями. Например, если факт независимости установлен формальным образом, то допустимы осторожные предположения о содержательной независимости. В байесовской концепции имеет место примат рационального и интуитивного подходов по отношению к формальным рассуждениям. Вероятностная независимость вводится на основе геометрической независимости. Действительно несвязанные геометрические структуры интуитивно представляются условно несвязанными и в вероятностном плане, где в качестве условия берутся родители исследуемой структуры. Однако, как мы уже говорили, в стохастической математике много неинтуитивных результатов. Попытки обосновать и использовать независимость во многих случаях объясняются тем, что независимость принципиально упрощает вычисления.

Пусть A_1, A_2, …, A_n – вершины графа, описывающие события a_1, a_2, …, a_n. Для случая, когда эти события являются зависимыми, совместная вероятность $P(a_1, a_2, …, a_n)$ определяется следующим образом:

$$P(a_1,...,a_n) = P(a_1)\cdot P(a_2 / a_1)\cdot P(a_3 / a_1, a_2)\cdot ...\cdot P(a_n / a_1, ..., a_{n-1}). \quad (1)$$

Для независимых событий вычисления упрощаются:

$$P(a_1,...,a_n) = P(a_1)\cdot P(a_2 / PAR_2)\cdot P(a_3 / PAR_3)\cdot ...\cdot P(a_n / PAR_n). \quad (2)$$

Здесь PAR_i – родители вершины A_i, $P(a_i/PAR_i)$ – условная вероятность переменной a_i, где в качестве условия выступают родители этой вершины.

Во многих случаях вычисления по формуле (2) осуществляются на основе применения принципа максимальной энтропии. Пусть $P_i = P(a_i/PAR_i)$, а H – это обозначение функции энтропии. На основе введенных обозначений энтропия определяется следующим образом:

$$H = \sum H_i . \quad (3)$$

Здесь H_i – это энтропия независимых частей графа, i=1, …, n:

$$H_i = -\sum \Big(\prod_{A_j \in Anc_i} P_j \cdot \log(P_i) \Big), \qquad (4)$$

где Anc_i – множество предков для вершины A_i [9].

Перейдем к анализу обоснованности принципа максимальной энтропии в байесовских цепях. Существует по крайней мере четыре группы аргументов в пользу обоснованности этого принципа.

Корректное применение принципа максимальной энтропии в специальных технических науках: теории связи, теории информации, термодинамике и др.

Впервые принцип максимальной энтропии был разработан и использован в теории связи К. Шенноном. Принцип заключается в максимизации функции энтропии. В самом простом случае формула энтропии имеет следующий вид:

$$H = -\sum P \cdot \log(P) . \qquad (5)$$

Содержательно функция энтропии определяет меру неопределенности. По Э. Джейнсу, уникальность функции энтропии и адекватность ее использования в теории информации, теории связи, термодинамике обусловлено сочетанием следующих свойств [10]:

1) функция энтропии является непрерывной функцией аргументов;

2) функция является возрастающей функцией числа аргументов;

3) вычисление функции от n аргументов сводится к взвешенной сумме от этой же функции для меньшего числа аргументов.

Однако существует ряд контраргументов к уникальности функции энтропии относительно вышеперечисленных условий и к обоснованности энтропии.

Во-первых, сам создатель теории информации К. Шеннон, первым применивший для описания неопределенностей функцию энтропии, не считал ее уникальной и единственной функцией, удовлетворяющей выше перечисленным условиям. И действительно, существуют другие функции, описывающие эти условия. По мнению Шеннона, значимость функции энтропии определяется не ее особыми логическими свойствами, а адекватностью для применений.

Во-вторых, понятие энтропии не является вполне корректным, в частности это относится к понятию условной энтропии. Понятие условной энтропии получается на основе понятия обычной энтропии при замене обычной вероятности в формуле (5) на условную вероятность. Имеем

$$H(x/y) = -\sum P(y)\sum P(x/y)\log(P(x/y)) .$$

Из теории Шеннона получается, что всегда

$$H(x/y) \le H(x) , \qquad (6)$$

т.е. при получении новой информации энтропия не может возрастать. Это соотношение не всегда является верным, в работе Дж. Уффинка приведен пример, опровергающий универсальность (6). Пусть вероятность того, что первый встреченный кролик будет черным, против вероятности того, что он будет белым, определяется как 0,99:0,01. Однако если у кролика мать белого цвета, то вероятность того, что и он будет белым, равна 0,5 [11]. В данном случае выходит, что при получении новой информации неопределенность не уменьшается.

Принцип максимальной энтропии как универсальный принцип описания вероятностей

Первоначально применение принципа энтропии ограничивалось его использованием для определенного круга задач в определенных научных дисциплинах. Впервые попытка использования этого принципа в качестве универсального статистического, а вообще говоря, с учетом широкой области приложений математической статистики, – в качестве общенаучного принципа была осуществлена Э. Джейнсом. По Джейнсу, принцип максимальной энтропии является развитием принципа недостаточной причины, получившего со временем название принципа безразличия Кейнса.

Использование принципа безразличия предполагает отсутствие информации о предпочтениях по отношению к n положений дел. В случае отсутствия предпочтений назначается одна и та же вероятность $1/n$ для каждого из n выделенных положений. Принцип максимальной энтропии по Джейнсу, усиливающий принцип безразличия Кейнса, должен, с одной стороны, учитывать известную эмпирическую и логическую информацию, а с другой стороны, по отношению к неизвестным параметрам, фоновым признакам и т.д. обеспечить максимальную свободу в назначении исходных вероятностей исследуемых положений дел. Детально учет эмпирической информации при использовании принципа максимальной энтропии был описан Дж. Уильямсоном.

В байесовизме предполагается, что исследователь обладает знанием объективного шанса – объективной сингулярной вероятностью, и на ее основе определяется степень веры. Связь объективного шанса и степени веры определяется с помощью ментально-физического принципа измерений. В самом простом случае ментально-физический принцип измерений формулируется следующим образом: «Если агент знает шанс $P^*(u)$ изучаемого положения дел u, то степень веры $P(u)$ равняется шансу $P^*(u)$» [12]. С целью широкого применения принцип был разработан не только для случая, когда шанс известен точно, но и для случая, когда он известен приблизительно.

Во-первых, принцип измерений учитывает информацию в дизъюнктивной (альтернативной) форме. Например, $P^*(u) = r \vee s$. Иногда считают, что альтернативное знание не является знанием. Ведь степень веры не может равняться ни одному из значений r, s (так как нам неизвестно, какое значение правильнее), ни одновременно обоим значениям. Однако в действительности дизъюнктивное знание имеет смысл. Даже в самом неопределенном случае, когда r и s – это вещественные числа, равные соответственно нулю и единице, для степени веры имеет смысл осторожная оценка 0,5. Однако наибольшее значение принцип калибровки имеет место, когда r и s являются близкими числами. В этом случае естественной оценкой степени веры служит замкнутый интервал $[r, s]$.

Во-вторых, использование интервалов имеет эвристическую значимость для знания, представленного в конъюнктивной форме – $P^*(u) = r \wedge s$, если $|r - s| < \varepsilon$, где ε – точность измерений. Необходимо отметить, что неправильно буквально понимать равенство шанса $P^*(u)$ и степени веры $P(u)$ для всех потенциально мыслимых положений дел. Например, пусть $P^*(u) \neq 0{,}5$ и нет никакой другой информации, иными словами, используется неправильная монета. Чему в этом случае равна степень веры? Если неизвестно, больше ли шанс, чем 0,5, или он меньше, чем 0,5, то естественно принять в качестве оценки степени веры именно 0,5. Чтобы избежать прямого нарушения принципа равенства шанса степени веры, имеет смысл определить степень веры равной $0{,}5 \pm \varepsilon$.

Существуют принципиально полярные позиции относительно значимости принципа калибровки. В первом случае считается, что принцип предъявляет слишком сильные требования к обязательному равенству объективного шанса и степени веры. Эта позиция основана на убеждении, что описание (измерение) степени веры с помощью шанса является второстепенной задачей, а главная задача состоит в определении предсказания. Однако в работе К. Корба и Э. Николсона [13] показана как самостоятельная значимость проблемы калибровки, так и связь корректной калибровки и последующего предсказания. Во втором случае принцип калибровки считается слишком слабым, так как он базируется на знании шанса, а обоснованное определение шанса представляет собой проблему. Действительно, определение шанса будет эпистемологическим, а не прагматическим, а потому связано с трудностями. Во-первых, знание шанса является точным, только если нам известна вся картина мира в фиксированный момент времени. Во-вторых, при определении шанса возникает проблема референтного класса, о которой мы писали ранее [14].

Принцип безразличия начиная с момента его создания подвергался критике вследствие многочисленных парадоксов, возникающих при его применении. Однако в последние годы все известные парадоксы дискретной природы удалось разрешить. Для непрерывных описаний ситуация не столь оптимистичная, однако даже для частного случая самого тяжелого парадокса о смеси воды и вина было найдено приемлемое решение. Принцип максимальной энтропии свободен от парадоксов, присущих принципу безразличия; кроме того, принцип максимальной энтропии учитывает известную эмпирическую и логическую информацию. Однако при всех своих достоинствах этот принцип не является универсальным. Существует огромное множество ситуаций, когда принцип максимизации энтропии не имеет никакого смысла вообще. Заметим, что энтропия принимает максимальное значение, когда все вероятности в формуле энтропии оказываются равными или приблизительно равными. На основании этого нетрудно привести множество примеров, когда максимизация энтропии неприемлема. Например, пусть p_1 – вероятность осудить невиновного человека, p_2 – вероятность наказать преступника. Очевидно, что p_2 должно быть в огромной степени больше p_1.

Принцип максимальной энтропии и симметричные принципы common sense.

Всплеск интереса к принципу максимальной энтропии связан с необходимостью учета неопределенностей в предпосылках при осуществлении рассуждений (reasoning) в экспертных системах. Как было показано в многочисленных работах Ж. Париса и А. Венковской, принцип максимальной энтропии, оказываясь усовершенствованным вариантом принципа безразличия, согласуется с множеством симметричных принципов [15]. Выделено семь общих принципов (common sense), имеющих в той или в иной степени симметричную интерпретацию:

1) *принцип переименования*. Рассуждения, обосновывающие вероятность выпадения грани, на которой написано число шесть, при бросании правильной шестигранной кости, не изменяются при замене этой грани на другую, где представлено другое число;

2) *иррелевантный информационный принцип*. Знание, не адекватное решаемой проблеме, не должно приниматься во внимание. Так, данные о погоде в России за прошлые годы не помогут предсказать, кто победит в чемпионате России по футболу;

3) *принцип избыточности, неэффективности*. Информация, подтверждающая ранее сделанные оценки тщательно изученных положений дел, не изменяет эти оценки;

4) *принцип относительности*. Некоторая информация может приниматься во внимание при условии знания контекста. Предположим, что в конце длительного пути путешественники достигли городской автобусной остановки. У них есть информация, что во все дни недели, кроме воскресенья, автобусы ходят всю ночь. Если сегодня будний день, то имеет смысл ждать автобус; если сегодня воскресенье, то нет смысла ждать его. В одном из объявлений утверждается, что последний автобус в воскресенье может прибывать с задержкой. Однако эта информация является контекстуальной, она важна только в том случае, если день недели известен;

5) *принцип эквивалентности*. Одинаковым положениям дел назначаются одинаковые вероятности. Некто не склонен делать зарядку, однако обещает начать ее делать, если при бросании трех монет на каждой выпадет герб. Вероятность начать делать зарядку равна 1/8. Та же вероятность получится при следующих условиях: если

при первом бросании будет герб, то испытание продолжается, и далее, если при одновременном бросании двух монет выпадает два герба, то предполагается начать делать зарядку;

6) *принцип непрерывности*. Практически все игры в покер игрок N проигрывает, и он замечает, что у партнера оказываются чрезвычайно сильные, выигрышные комбинации карт. Это приводит к подозрению, что партнер нечестен. Одна игра N удалась. Однако одна ласточка не делает погоду, и этот выигрыш не устраняет сомнения в честности противника;

7) *принцип слабой независимости* [16].

Принцип максимальной энтропии обеспечивает выполнимость всех этих семи принципов. При определении вероятностей на основе принципа максимальной энтропии они определяются однозначно и согласуются со всеми сформулированными выше семью принципами. Связь принципа максимальной энтропии с семью принципами, в той или иной степени являющимися принципами симметричности, в строгой форме обнаружена Парисом и Венковской и сформулирована ими в виде следующей теоремы [17].

Теорема. Соблюдение семи принципов: переименования, иррелеватности, избыточности, относительности, эквивалентности, непрерывности и слабой независимости (все они могут рассматриваться как специальные случаи принципа симметрии) – однозначно определяет вероятности на основе имеющейся информации.

Комбинирование принципов марковости и максимальной энтропии

Для положений дел, описываемых независимыми множествами A_1, A_2, …, A_n на графе G, совместное распределение $P(a_1 / PAR_1, a_2 / PAR_2, ..., a_n / PAR_n)$ определяется путем максимизации функции энтропии, соответствующей формулам (3) и (4). Задача максимизации успешно решается для неполных графов с небольшим числом вершин с помощью методов оптимизации или вариационного исчисления.

Принцип максимальной энтропии весьма популярен в экспертных системах. Несмотря на определенные достоинства принципа, а именно то, что он учитывает знание эмпирического и логического характера и позволяет установить распределение вероятностей для независимых положений дел универсальным образом – путем максимизации функции энтропии, этот принцип не является универсальным методом определения вероятностей. Его применение ограничено ситуациями, когда максимизация энтропии в принципе имеет смысл. Техническое воплощение максимизации энтропии эффективно реализуется для неполных графов с небольшим числом вершин для решения проблем, в которых описываемые независимые множества имеют близкие вероятности и распределение вероятностей задано априори.

Для решения реальных больших задач в области ответственных приложений, например в медицине или экономике, в которых, как правило, распределение данных априори не известно, адекватным является чисто эмпирический подход [18].

Рассмотрим задачу восстановления неизвестного дискретного распределения $P(x_1, x_2, \ldots, x_n)$. Покажем, что она может быть решена на основе чисто эмпирического подхода в отличие от рассмотренных выше. Относительно этого распределения мы не делаем никаких априорных предположений. Нами разработан метод [19] обнаружения сети зависимостей, подобный методам обнаружения структур байесовской сети по данным [20]. Все зависимости обнаруживаются в виде множества вероятностных законов. По совокупности этих зависимостей может быть восстановлено распределение $P(x_1, x_2, \ldots, x_n)$. При этом зависимости в виде вероятностных законов ищутся так, что между собой они в определенном смысле независимы. Таким образом, исходное распределение декомпозируется на зависимые и независимые части.

Примечания

1. См.: *Колмогоров А.Н.* Основные понятия теории вероятностей. – М.: Наука, 1974.
2. См.: *Крамер Г.* Математические методы статистики. – М.: Мир, 1975.
3. См.: *Резников В.М.* Методологические проблемы применения статистических критериев // Вест. Новосиб. ун-та. Сер. Философия. – Новосибирск, 2009. – Т. 7, вып. 4. – С. 32.
4. Там же.
5. См.: *Алимов Ю.И.* Альтернатива методу математической статистики. – М.: Знание, 1980.
6. См.: *Эльясберг П.С.* Вычислительная информация: Сколько ее нужно? Как ее обрабатывать? – М.: Наука, 1986.
7. См., например: *Секей Г.* Парадоксы в теории вероятностей и математической статистике. – М.: Мир, 1990.
8. См.: ***Neapolitan E.*** Learning Bayesian networks. – Illinois Univ. Press, 1990.

9. См.: *Williamson J.* Objective Bayesian nets // Essays in Honour of Dov Gabbay. – Kent: College Publications, 2005. – V. 2. – P. 713–730.
10. См.: *Jaynes E.* Probability theory – the logic of science. – Cambridge Univ. Press, 2003.
11. См.: *Uffink J.* Can the maximum entropy principle be explained as a consistency requirement? // Studies in History and Philosophy of Science. Part B: Studies In History and Philosophy of Modern Physics. –1995. – V. 26, is. 3. – P. 223–261.
12. См.: *Williamson J.* Bayesian nets and causality: philosophical and computational foundations. – Oxford Univ. Press, 2005. – P. 71.
13. См.: *Korb K., Nicholson A.* Bayesian Artificial Intelligence. – London.: CRC Press Company, 2004.
14. См.: *Резников В.М.* Проблема референтного класса и условные вероятности // Философия науки. – 2004. – № 1. – С. 83–93.
15. См., например: *Paris J.* Common sense and maximum entropy // Synthese. – 1999. – V. 117. – P. 75–93.
16. Ibid.
17. Ibid.
18. См.: *Kovalerchuk B., Vityaev E.* Data mining in finance: Advances in relational and hybrid methods. – N.Y.: Kluwer Academic Publishers, 2000; *Kovalerchuk B, Vityaev E., Ruiz J.* Consistent and complete data and «expert» mining in medicine // Medical Data Mining and Knowledge Discovery. – 2001. – P. 238–280; *Витяев Е.Е.* Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. – Новосибирск: Новосиб. гос. ун-т. – 2006.
19. См.: *Витяев Е.Е.* Извлечение знаний из данных…
20. См.: *Spirtes P., Glymour C., Scheines R.* An algorithm for fast recovery of sparse causal graphs // Social Science Computing Reviews. – 1991. V.9. – P. 62–72.

Институт математики СО РАН,
г. Новосибирск
vityaev@math.nsc.ru

Институт философии и права СО РАН,
г. Новосибирск
rvm@philosophy.nsc.ru

***Reznikov, V.M. and E.E. Vityaev.* The problem of statistical independency and the maximum entropy principle**

Unlike the standard statistical analysis where correctness of the apparatus applied is proved formally, bayesovism proves formal descriptions by means of intuitive and substantive-rational arguments. The paper shows that the maximum entropy principle which bayesovism use to calculate integrated probability of independent events is not universal. It is valid for solving problems described by means of incomplete graphs with a small number of vertices where entropy maximization makes sense. Concerning real problems where uncertainty of knowledge is taken into account, in the domain of responsible applications combination of logical programming and statistical data analysis is more valid.

Keywords: independency, Markov condition, entropy, maximum entropy principle