УДК 519.211 DOI 10.12737/5695

О корректности модификаций формулы Байеса*

А. И. Долгов

Утверждается недопустимость сравнения и совместной обработки условных вероятностей, вычисляемых с использованием известных модификаций формулы Байеса, учитывающих вполне определённое количество накопленных свидетельств. Данное утверждение обосновывается следующим образом. Ввиду неодинакового нормирования получаются результаты, приводящие к выводам, не соответствующим теоретико-вероятностным закономерностям. На примере конкретных исходных данных показано, что в случае использования традиционных формул отсутствует мультипликативный эффект: уменьшение апостериорных условных вероятностей каждой из гипотез при мультипликативном накапливании свидетельств оказывается неосуществимым. Предложены модификации формулы Байеса, учитывающие вполне определённое количество последовательно накапливаемых свидетельств, построенные с применением общего делителя, равного сумме значений всех нормируемых вероятностей, которые подлежат сравнению либо совместной обработке. В отличие от известных формул, ввиду корректного нормирования, получаются результаты, существенно более адекватные изначальным статистическим данным.

Ключевые слова: *условные вероятности, гипотезы, накопленные и накапливаемые свидетельства, нормирование.*

Введение. Формула Байеса широко используется в теории вероятностей и статистике [1–9]. Весьма актуальным направлением применения формулы Байеса является разработка модификаций, обеспечивающих определение апостериорных условных вероятностей гипотез с учётом накапливания свидетельств.

Наиболее известны по крайней мере три различающиеся между собой модификации формулы Байеса, предназначенные для оценки апостериорных условных вероятностей гипотез с учётом накопленных свидетельств [2–4]. Две из них [3, 4] являются модификациями формулы Байеса с изменениями, обеспечивающими учёт мультипликативно накопленных свидетельств, а одна [2] соответствует рекуррентному применению формулы Байеса с пошаговым получением соотношений, эквивалентных другой модификации [3].

Как показывает анализ, традиционно используемые модификации формулы Байеса, учитывающие мультипликативно накопленные свидетельства, не применимы для определения апостериорных условных вероятностей гипотез при мультипликативно накапливаемых свидетельствах, ввиду некорректности получаемых результатов.

Для устранения выявленного недостатка решается задача построения модификаций формулы Байеса, обеспечивающих учёт накапливаемых свидетельств с применением корректного нормирования.

Исходные положения. В основу излагаемых далее утверждений положим известное определение вероятности: «Чтобы количественно сравнивать между собой события по степени их возможности, очевидно, нужно с каждым событием связать определённое число, которое тем больше, чем более возможно событие. Такое число называется вероятностью события» [5].

Судя по приведённому определению, понятие вероятностей введено, в основном, для сравнения между собой различных событий по степени возможности, количественно выражаемой соотношениями измеренных значений их вероятностей.

С учётом сказанного целесообразно сформулировать представляющий объективную основу теории вероятностей, но не нашедший необходимого отражения в современной учебной и на-

* Работа выполнена в рамках инициативной НИР.

учно-технической литературе принцип сохранения отношений изначальных вероятностей: корректная обработка вероятностей осуществима лишь при нормировании, выполненном с применением одного и того же нормирующего делителя, обеспечивающего равенство отношений нормированных вероятностей отношениям нормируемых изначальных априорных вероятностей.

Изначальные вероятности — вероятности, характеризующие степень возможности всех входящих в генеральную совокупность рассматриваемых событий или их комбинаций, совместно определённые априорно на основе объективных данных при условии учёта одних и тех же факторов.

Из сформулированного принципа следует, что:

— отношения нормированных вероятностей событий или их комбинаций, образующих полную группу, в любых случаях получаются с использованием одного и того же нормирующего делителя и должны быть равны отношениям нормируемых изначальных вероятностей;

— сравнение и совместная обработка нормированных изначальных вероятностей событий или их комбинаций, принадлежащих разным полным группам, недопустимы, ввиду того, что такие вероятности получены с использованием разных нормирующих делителей.

При нарушении сформулированного принципа искажаются сведения о степени возможности рассматриваемых событий или их комбинаций, и получаемые на основе искажённых сведений результаты и принимаемые решения оказываются неадекватными реальным статистическим данным.

Следует исходить из того, что любые представляющие интерес для рассмотрения полные группы событий или их комбинаций выбираются из их генеральной совокупности **субъективно.** При этом в анализируемой предметной области (как в «дискретном куске действительности») рассматриваются только такие ситуации, при которых в результате опыта появляется хотя бы одно из событий или их комбинаций, входящих в выбираемую полную группу.

При определении условных вероятностей традиционно рассматриваются простые и комбинированные события.

1) Простые (отдельно рассматриваемые) образуют полную группу $\{H_1, H_2, ..., H_n\}$ несовместных гипотез. При этом события характеризуются изначальными априорными вероятностями $P(H_1), P(H_2), ..., P(H_n)$, сумма которых равна 1.

2) Комбинированные — в виде мультипликативных комбинаций, представляемых произведениями $H_kE_i, H_kE_iE_j, ..., H_kE_1E_2...E_m$ простых событий — учитываемых гипотез H_k $(k = 1,...,n)$ и свидетельств E_i $(i = 1,...,m)$, подтверждающих гипотезы. При этом вероятности комбинаций событий определяются на основе предварительно полученных изначальных вероятностей — в частности, вероятность $P(H_k E_1)$ комбинации простых событий H_k и E_1 определяется как произведение известных изначальных априорных вероятностей $P(H_1)$ и $P(E_1 | H_1)$, так как $P(H_kE_1) = P(H_1)\,P(E_1 | H_1)$ согласно теореме умножения вероятностей [5].

Анализ известных соотношений. Как известно, формула Байеса [5] может быть записана в виде

$$P(H_k|E) = \frac{P(H_k)P(E|H_k)}{\sum\limits_{k=1}^{n} P(H_k)P(E|H_k)}, \qquad (1)$$

где $P(H_k)$ — априорная вероятность гипотезы (события-следствия) H_k ($k = 1,...,$n); $P(E | H_k)$ — априорная условная вероятность наличия свидетельства (события-причины) E при возникновении события-следствия H_k.

Нетрудно видеть, что выражением (1) представлена вероятностная доля $P(H_k)P(E | H_k)$ среди всех возможных комбинаций $P(H_k)P(E | H_k)$, соответствующих $k = 1,...,n$, указанных в знаменателе рассматриваемого соотношения в соответствии с так называемой [5] формулой полной вероятности $P(E)=\sum_{k=1}^{n}P(H_k)P(E\,|\,H_k)$.

Формула, использованная в публикации Нейлора [2], записывается в виде, более простом с точки зрения количества указываемых вычислительных операций:

$$P(H_k\,|E)=\frac{P(H_k)\,P(E\,|H_k)}{P(H_k)\,P(E\,|H_k)+P(неH_k)\,P(E\,|неH_k)}, \tag{2}$$

где $P(E\,|\,неH_k)$ — априорная вероятность наличия события-причины E при невозникновении события-следствия H_k.

В рамках традиционных допущений из формулы полной вероятности вида $P(E\,|(H_k+неH_k))=P(H_k)P(E\,|\,H_k)+P(неH_k)P(E\,|\,неH_k)$ с учётом того, что при рассматриваемой полной группе несовместных гипотез $неH_k=\sum_{i=1(i\neq k)}^{n}H_i$, следует, что

$$P(H_k)P(E\,|H_k)+P(неH_k)P(E\,|\,неH_k)=\sum_{k=1}^{n}P(H_k)P(E\,|\,H_k).$$

Таким образом, соотношением (2) также представлена вероятностная доля $P(H_k)P(E\,|\,H_k)$ среди указанных в знаменателе всевозможных комбинаций $P(H_k)P(E\,|\,H_k)$, соответствующих $k=1,\ldots,n$. Другими словами, при традиционных допущениях в случае одного свидетельства E как в знаменателях, так и в числителях обеих формул (1) и (2) представлены эквивалентные выражения.

Следует отметить, что в отличие от (1) в соотношении (2) используется значение $P(E\,|\,неH_k)$, которое при условии эквивалентности знаменателей может быть определено из уравнения:

$$\sum_{k=1}^{n}P(H_k)P(E\,|\,H_k)=P(H_k)P(E\,|H_k)+P(неH_k)P(E\,|\,неH_k), \tag{3}$$

при этом

$$P(E\,|\,неH_k)=\frac{\sum_{i=1(i\neq k)}^{n}P(H_i)P(E\,|H_i)}{P(неH_k)}=\frac{\sum_{i=1(i\neq k)}^{n}P(H_i)P(E\,|H_i)}{1-P(H_k)}. \tag{4}$$

В сущности, формула (2) была использована в публикации [2] как модификация формулы Байеса, весьма удобная для рекуррентных вычислений. Как отмечает автор публикации [2][*], в случае накапливания свидетельств, вычислив апостериорную вероятность гипотезы H_k для одного учитываемого свидетельства E, «мы забываем об этом, за исключением того, что априорная вероятность $P(H)$ заменяется на $P(H\,|\,E)$», а затем продолжается выполнение вычислений, «но с учётом постоянной коррекции значения $P(H)$ по мере поступления новой информации» (свидетельств).

В целях анализа корректности соотношений, традиционно применяемых для определения апостериорных условных вероятностей гипотез при накапливании свидетельств, вычислим с использованием формулы (2) апостериорные условные вероятности $P(E_j\,|\,H_k)$ двух гипотез при каждом одном свидетельстве, а также вероятности $P(H_k\,|E_1E_2)$ при двух свидетельствах на конкретном примере следующих исходных данных, характеризующих рассматриваемые значения априорных вероятностей:

---

[*] Цитируются фрагменты в переводе с английского языка на русский.

$$P(H_1) = 0{,}3;\ P(E_1 \mid H_1) = 0{,}6;\ P(E_2 \mid H_1) = 0{,}5;$$

$$P(H_2) = 0{,}7;\ P(E_1 \mid H_2) = 0{,}8;\ P(E_2 \mid H_2) = 0{,}9.$$

С учётом исходных данных, запишем используемые далее промежуточные значения:

$$P(H_1)P(E_1 \mid H_1) = 0{,}18;\ P(H_1)P(E_2 \mid H_1) = 0{,}15;\ P(H_1)P(E_2 \mid H_1)P(E_1 \mid H_1) = 0{,}09;$$

$$P(H_2)P(E_1 \mid H_2) = 0{,}56;\ P(H_2)P(E_2 \mid H_2) = 0{,}63;\ P(H_2)P(E_1 \mid H_2)P(E_2 \mid H_2) = 0{,}504;$$

$$\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k) = 0{,}74;\ \sum_{k=1}^{2} P(H_k)P(E_2 \mid H_k) = 0{,}78;\ \sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)P(E_2 \mid H_k) = 0{,}594.$$

Для упрощения вычислений будем исходить из того, что значение $P(E_j \mid неH_k)$ является вычисляемым, приводящим формулу (2) к соотношению, тождественно эквивалентному формуле (1) Байеса:

$$P(H_k \mid E_j) = \frac{P(H_k)P(E_j \mid H_k)}{P(H_k)P(E_j \mid H_k) + P(неH_k)P(E_j \mid неH_k)} = \frac{P(H_k)P(E_j \mid H_k)}{\sum_{k=1}^{2} P(H_k)P(E_j \mid H_k)}, \tag{5}$$

и, следовательно,

при $k = 1$ $P(H_1 \mid E_1) = \dfrac{0{,}18}{0{,}74} = 0{,}243$ и $P(E_1 \mid H_2) = \dfrac{0{,}15}{0{,}78} = 0{,}192$,

при $k = 2$ имеем $P(H_2 \mid E_1) = \dfrac{0{,}56}{0{,}74} = 0{,}757$ и $P(H_2 \mid E_2) = \dfrac{0{,}63}{0{,}78} = 0{,}808$.

Теперь определим апостериорную условную вероятность гипотезы H_1 при двух свидетельствах с учётом того, что $P(неH_1) = 1 - P(H_1)$ и в формуле (5) в порядке её рекуррентного использования при $k = 1$ априорная вероятность $P(H_1)$ заменяется на $P(H_1 \mid E_1)$:

$$P(H_1 \mid E_1E_2) = \frac{P(H_1 \mid E_1)P(E_2 \mid H_1)}{P(H_1 \mid E_1)P(E_2 \mid H_1) + (1 - P(H_1 \mid E_1))P(E_2 \mid неH_1)}.$$

Подстановка вычисляемого в соответствии с (4) значения

$$P(E_2 \mid неH_1) = \frac{\sum_{i=1(i\neq 1)}^{2} P(H_i)P(E_2 \mid H_i)}{\sum_{i=1(i\neq 1)}^{2} P(H_i)} = \frac{P(H_2)P(E_2 \mid H_2)}{P(H_2)} = P(E_2 \mid H_2) = 0{,}9,$$

приводит к принимающему соответствующее значение соотношению

$$P(H_1 \mid E_1E_2) = \frac{\dfrac{P(H_k)P(E_1 \mid H_k)}{\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)} P(E_2 \mid H_1)}{\dfrac{P(H_k)P(E_1 \mid H_k)}{\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)} P(E_2 \mid H_1) + \left(1 - \dfrac{P(H_k)P(E_1 \mid H_k)}{\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)}\right)P(E_2 \mid H_2)} =$$

$$= \frac{P(H_1)P(E_1 \mid H_1)P(E_2 \mid H_1)}{\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)P(E_2 \mid H_k)} = \frac{0{,}09}{0{,}594} = 0{,}152.$$

Аналогично $P(H_2 \mid E_1E_2) = \dfrac{P(H_1)P(E_1 \mid H_1)P(E_2 \mid H_1)}{\sum_{k=1}^{2} P(H_k)P(E_1 \mid H_k)P(E_2 \mid H_k)} = \dfrac{0{,}504}{0{,}594} = 0{,}848.$

Следует обратить внимание на то, что формула (2) в случае учёта m свидетельств приводит к довольно известной (см., например, [3, с. 284–285]) модификации формулы Байеса

$$P(H_k|E_1E_2...E_m)=\frac{P(H_k)P(E_1|H_k)P(E_2|H_k)\times...\times P(E_m|H_k)}{\sum\limits_{k=1}^{n}P(H_k)P(E_1|H_k)P(E_2|H_k)\times...\times P(E_m|H_k)}, \quad (6)$$

учитывающей множество накопленных свидетельств.

Из анализа результатов, полученных при каждом из двух рассматриваемых свидетельств и при их мультипликативной комбинации, имеем два следствия.

Во-первых, при традиционно применяемых методах вычисления апостериорных условных вероятностей $P(H_1|E_1)$, $P(H_1|E_2)$ и $P(H_1|E_1E_2)$ осуществляются с использованием разных нормирующих делителей, что исключает возможность их сравнения и совместной обработки.

Во-вторых, при мультипликативном накапливании свидетельств отсутствует мультипликативный эффект:

— апостериорная условная вероятность гипотезы H_1, равная при одном свидетельстве $P(H_1|E_1)$ = 0,243 либо $P(H_1|E_2)$ = 0,192, при переходе к двум свидетельствам уменьшилась до $P(H_1|E_1E_2)$ = 0,152,

— апостериорная условная вероятность гипотезы H_2, равная при одном свидетельстве $P(H_2|E_1)$ = 0,757 либо $P(H_2|E_2)$ = 0,808, при переходе к двум свидетельствам увеличилась до $P(H_2|E_1E_2)$ = 0,848.

Заключая сказанное, следует также отметить, что использование разных нормирующих делителей при изменяющемся количестве учитываемых свидетельств и несоблюдение мультипликативного эффекта присуще не только модификациям формулы Байеса (2) и (6), но и другой модификации, имеющей [4] следующий вид:

$$P(H_k|E_1E_2...E_m)=\frac{P(H_k)P(E_1|H_k)P(E_2|H_kE_1)\cdot...\cdot P(E_m|H_kE_1E_2\cdot...\cdot E_{m-1})}{\sum\limits_{k=1}^{n}P(H_k)P(E_1|H_k)P(E_2|H_kE_1)\cdot...\cdot P(E_m|H_kE_1E_2\cdot...\cdot E_{m-1})}.$$

Таким образом, традиционные методы оценки апостериорных условных вероятностей гипотез, учитывающие накапливание свидетельств, ввиду того, что при нормировании таких вероятностей используются разные нормирующие делители, приводят к некорректным результатам: при мультипликативном накапливании свидетельств уменьшение апостериорных условных вероятностей каждой из гипотез (связанное с мультипликативным эффектом, соответствующим теоретико-вероятностным закономерностям) оказывается неосуществимым.

Следует также обратить внимание на то, что с применением известных модификаций формулы Байеса оцениваются апостериорные условные вероятности гипотез при свидетельствах, накопленных в порядке их предварительно установленной нумерации, в то время как более реальными являются случаи, когда порядок поступления свидетельств является изменяемым. Это можно учесть при построении модификаций формулы Байеса с корректным нормированием.

Нетрадиционное нормирование. Сравнение апостериорных условных вероятностей гипотез при неодинаковом (изменяемом) количестве свидетельств осуществимо в случае корректного нормирования с использованием не разных нормирующих делителей, а одного общего делителя q, равного сумме альтернативных выражений, указанных в числителях всех используемых соотношений, а следовательно, и сумме всех ранее использованных нормирующих делителей.

Например, в случае рассмотренных априорных исходных данных, ввиду того, что нормируемыми являются вероятности $P(H_k)P(E_i|H_k)$, $i = 1$ или $i = 2$, и $P(H_k)P(E_1|H_k)P(E_2|H_k)$, $k = 1,...,2$, нормирующий делитель выбирается равным

$$q=\sum_{k=1}^{2}P(H_k)P(E_1|H_k)+\sum_{k=1}^{2}P(H_k)P(E_2|H_k)+\sum_{k=1}^{2}P(H_k)P(E_1|H_k)P(E_2|H_k).$$

Можно перейти к более простой записи нормирующего делителя. Если просуммировать слагаемые с сомножителями, представляющими вероятности альтернативных комбинаций свидетельств при гипотезе H_1, и раздельно просуммировать аналогичные слагаемые с вероятностями при гипотезе H_2 с вынесением одинаковых сомножителей $P(H_k)$ за скобки, и затем полученные выражения сложить, то (при $n = m = 2$)

$$q = \sum_{k=1}^{2}(P(H_k)[P(E_1 \mid H_k) + P(E_2 \mid H_k) + P(E_1 \mid H_k)P(E_2 \mid H_k)]) = \\ = \sum_{k=1}^{2}\left(P(H_k)\left[\sum_{i=1}^{2}P(E_i \mid H_k) + \prod_{i=1}^{2}P(E_i \mid H_k)\right]\right). \qquad (7)$$

С использованием общего нормирующего делителя, который при заданных конкретных исходных данных принимает значение

$$q = 0{,}74+0{,}78+0{,}594 = 2{,}114,$$

равное сумме ранее использованных знаменателей, естественно, получаем иные результаты:

$$P(H_1 \mid E_1) = \frac{0{,}18}{2{,}114} = 0{,}085, \qquad P(H_1 \mid E_2) = \frac{0{,}15}{2{,}114} = 0{,}071;$$

$$P(H_1 \mid E_1 E_2) = \frac{0{,}09}{2{,}114} = 0{,}043; \qquad P(H_2 \mid E_1) = \frac{0{,}56}{2{,}114} = 0{,}265;$$

$$P(H_2 \mid E_2) = \frac{0{,}63}{2{,}114} = 0{,}298; \qquad P(H_2 \mid E_1 E_2) = \frac{0{,}504}{2{,}114} = 0{,}238.$$

Результаты вычислений, выполненных для рассматриваемого примера конкретных априорных исходных данных, подтверждают, что в случае нетрадиционного (корректного) нормирования значения апостериорных условных вероятностей всех гипотез при мультипликативном учёте накапливаемых свидетельств могут только уменьшаться, что соответствует теоретико-вероятностным закономерностям. В частности, при заданных априорных исходных данных конкретно рассматриваемого примера апостериорные условные вероятности обеих гипотез при получении второго свидетельства (ввиду мультипликативного эффекта) уменьшились: апостериорная условная вероятность гипотезы H_1 уменьшилась с 0,085 либо с 0,071 до 0,043, а прежде увеличившаяся вероятность гипотезы H_2 уменьшилась с 0,265 либо с 0,298 до 0,238.

Таким образом, на основе обобщения (7) может быть предложено корректное соотношение для определения апостериорных условных вероятностей гипотез в случае изменяемого количества $\mu = 1,\dots,m$ мультипликативно накапливаемых свидетельств:

$$P\left(H_k \left| \prod_{i=i_1}^{i_\mu} E_i \right.\right) = \frac{P(H_k)P\left(E_{i_1} \middle| H_k\right)\dots P\left(E_{i_\mu} \mid H_k\right)}{\sum_{k=1}^{n}\left(P(H_k)\left[\sum_{i}P(E_i \mid H_k) + \sum_{i,j}P(E_i \mid H_k)P(E_j \mid H_k) + \dots + \prod_{i=1}^{m}P(E_i \mid H_k)\right]\right)},$$

где E_{i_1} — первое поступившее свидетельство, …, E_{i_μ} — μ-е поступившее свидетельство, причём $\left\{E_{i_1}\dots E_{i_\mu}\right\} \subseteq \left\{E_1\dots E_\mu\right\}$, и ввиду того, что множество $\left\{E_{i_1}\dots E_{i_\mu}\right\}$ является подмножеством $\left\{E_1\dots E_\mu\right\}$, учитывается произвольный порядок поступления свидетельств.

Соответствующая записанному соотношению рекуррентная формула при начальном значении

$$P\left(H_k \middle| E_{i_1}\right) = \frac{P(H_k)P\left(E_{i_1} \mid H_k\right)}{\sum_{k=1}^{n}\left(P(H_k)\left[\sum_{i}P(E_i \mid H_k) + \sum_{i,j}P(E_i \mid H_k)P(E_j \mid H_k) + \dots + \prod_{i=1}^{m}P(E_i \mid H_k)\right]\right)}$$

имеет вид:

$$P\left(H_k \left| \prod_{i=i_1}^{i_{\mu+1}} E_i \right.\right) = P\left(H_k \left| \prod_{i=i_1}^{i_\mu} E_i \right.\right) \times P\left(E_{i_{\mu+1}} \mid H_k\right)$$

Такая формула в отличие от прототипов, ввиду корректного нормирования, является существенно более адекватной изначальным статистическим данным. При этом в каждом шаге рекуррентного процесса используется всего одна расчётная операция умножения.

Заключение. Известные модификации формулы Байеса, используемые для вычисления апостериорных условных вероятностей несовместных гипотез с учётом того или иного количества мультипликативно накопленных свидетельств, не применимы для определения апостериорных условных вероятностей, учитывающих изменяемое количество последовательно накапливаемых свидетельств, так как, ввиду неодинакового нормирования, приводят к результатам, не соответствующим теоретико-вероятностным закономерностям. Сравнение апостериорных условных вероятностей гипотез при изменяемом количестве свидетельств осуществимо в случае корректного нормирования с использованием не разных нормирующих делителей, а одного общего делителя, равного сумме значений тех нормируемых вероятностей, которые подлежат сравнению либо совместной обработке.

Библиографический список

1. Gnedenko, B. V. Elementary Introduction to theory of Probability / B. V. Gnedenko, A. Ya. Khinchin. — San Francisco, London : Freeman and Co, 1961. — 139 p.

2. Naylor, C.-M. Build Your Own Expert System / C.-M. Naylor. — Wilmslow : Sigma Technical Press, 1983. — 248 p.

3. Романов, В. П. Интеллектуальные информационные системы в экономике : учеб. пособие. — 2-е изд., стер. / В. П. Романов. — Москва : Экзамен, 2007. — 496 с.

4. Змитрович, А. И. Интеллектуальные информационные системы / А. И. Змитрович. — Минск : НТООО «ТетраСистемс», 1997. — 496 с.

5. Вентцель, Е. С. Теория вероятностей : учеб. для вузов. — 9-е изд., стер. / Е. С. Вентцель. — Москва : Академия, 2003. — 576 с.

6. Андронов, А. М. Теория вероятностей и математическая статистика : учеб. для вузов / А. М. Андронов, Е. А. Копытов, Л. Я. Гринглаз. — Санкт-Петербург : Питер, 2004. — 481 с.

7. Баврин, И. И. Теория вероятностей и математическая статистика : учеб. для вузов / И. И. Баврин. — Москва : Высш. шк., 2005. — 160 с.

8. Пугачёв, В. С. Теория вероятностей и математическая статистика : учеб. пособие. — 2-е изд., испр. и доп. / В. С. Пугачёв. — Москва : Физматлит, 2002. — 496 с.

9. Справочник по теории вероятностей и математической статистике / В. С. Королюк [и др.]. — Москва : Наука, 1985. — 640 с.

10. Тимошенко, Е. И. Теория вероятностей : учеб. пособие / Е. И. Тимошенко, Ю. Е. Воскобойников. — Новосибирск : Новосиб. гос. архитектур.-строит. ун-т, 2003. — 98 с.

Материал поступил в редакцию 11.12.2012.

References

1. Gnedenko, B. V., Khinchin, A. Y. Elementary Introduction to theory of Probability. San Francisco, London : Freeman and Co, 1961, 139 p.

2. Naylor, C.-M. Build Your Own Expert System. Wilmslow : Sigma Technical Press, 1983, 248 p.

3. Romanov, V. P. Intellektualnyye informatsionnyye sistemy v ekonomike : uchebnoye posobiye [Intelligent information systems in economy : tutorial.] 2nd ed., reimpress. Moscow : Ekzamen, 2007, 496 p. (in Russian).

4. Zmitrovich, A. I. Intellektualnyye informatsionnyye sistemy [Intelligent information systems.] Minsk : NTOOO«TetraSistems», 1997, 496 p. (in Russian).

5. Ventsel, E. S. Teoriya veroyatnostey: uchebnik dlya vuzov. [Probability theory: textbook for universities.] 9th ed., reimpress. Moscow : Akademiya, 2003, 576 p. (in Russian).

6. Andronov, A. M., Kopytov, E. A., Gringlaz, L. Y. Teoriya veroyatnostey i matematicheskaya statistika : uchebnik dlya vuzov. [Probability theory and mathematical statistics: textbook for universities.] Sankt-Peterburg : Piter, 2004, 481 p. (in Russian).

7. Bavrin, I. I. Teoriya veroyatnostey i matematicheskaya statistika : uchebnik dlya vuzov. [Probability theory and mathematical statistics : textbook for universities.] Moscow : Vysshaya shkola, 2005, 160 p. (in Russian).

8. Pugachev, V. S. Teoriya veroyatnostey i matematicheskaya statistika : uchebnoye posobiye. [Probability theory and mathematical statistics: tutorial.] 2nd revised and enlarged ed. Moscow : Fizmatlit, 2002, 496 p. (in Russian).

9. Korolyuk, V. S., et al. Spravochnik po teorii veroyatnostey i matematicheskoy statistike. [Handbook on Probability Theory and Mathematical Statistics.] Moscow : Nauka, 1985, 640 p. (in Russian).

10. Timoshenko, E. I., Voskoboynikov, Y. E. Teoriya veroyatnostey : uchebnoye posobiye. [Probability theory: tutorial.] Novosibirsk : NGASU, 2003, 98 p. (in Russian).

ON CORRECTNESS OF BAYES FORMULA*

A. I. Dolgov

The inadmissibility of the comparison and cooperative processing of the conditional probabilities calculated with the use of the known variations of Bayes formula considering a well-defined amount of the accumulated evidences is established. This statement is justified as follows. Due to the unequal normalization, the results leading to the conclusions mismatching the probability-theoretical law are obtained. The examples of the concrete initial data show that in the case of using the traditional formulas, the multiplicative effect is missing: reducing the posteriori conditional probabilities of each of the hypotheses under the multiplicative accumulating evidences is unfeasible. The Bayesian formula modifications are offered. They provide a well-defined quantity of sequentially accumulated evidences constructed through the common divisor which is equal to the sum of the values of all normalizable probabilities liable to comparison or cooperative processing. In contrast to the well-known formulas, due to the correct normalization, the results significantly more adequate to the initial statistics are obtained.

Keywords: *conditional probabilities, hypotheses, accumulated and accumulating evidences, normalization.*

* The research is done within the frame of the independent R&D.