

# Байесовская стратегия оценки достоверности и метод байесовских сетей

И. Д. Баранов

Финансовый университет при Правительстве Российской Федерации (Финуниверситет), Financial University  
Baranov\_id@yandex.ru

**Аннотация.** На сегодняшний день использование байесовского подхода с целью его применения к процессам различных видов выступает довольно актуальной задачей оценки статистических, а также математических моделей. Байесовская методология является альтернативным вариантом классическому статистическому подходу, данная методология позволяет давать исследуемым моделям наиболее полную и точную оценку.

**Ключевые слова:** байесовский подход; неопределенность; модель; алгоритм; решение; нечеткие множества

Байесовский подход предоставляет вероятность получения приемлемых результатов в ситуациях, в которых применение классических методов является крайне ограниченным (к примеру, применение классических статистических методов ограничено в ситуации с коротким диапазоном выборки данных). Байесовская методология открывает новые, широкие возможности использования методов математического моделирования, а созданные на основе генерирования случайных чисел алгоритмы оценивания помогают нахождению решения различных задач при помощи современных вычислительных процедур.

## I. СУТЬ БАЙЕСОВСКОГО ПОДХОДА

Байесовские методы разработаны вследствие многочисленных попыток ученых определить проблемы статистического анализа поведения различных процессов и найти их решение с помощью применения основы байесовской методологии – теоремы Байеса. Использование данной теоремы имеет ряд предпосылок, основная из которых – наличие определенных соотношений между вероятностями явлений, имеющих различный характер и спецификации любого явления на нужном уровне [1]. Байесовская методология отличается от других подходов тем, что еще до получения данных исследователь определяет уровень своего доверия к возможным моделям и впоследствии представляет ее в виде определенных вероятностей. После того, как исследователем получены данные, с использованием теоремы Байеса он находит еще одно множество вероятностей, которые являются пересмотренными степенями доверия к возможным моделям с учетом полученной исследователем новой информации. Одним из ключевых преимуществ байесовского подхода является использование любой начальной (априорной) информации относительно параметров модели. Такая информация

выражается в виде априорной вероятности или функции плотности вероятности. Затем начальные вероятности «пересматриваются», с помощью выборочных данных, которые находят свое отображение в виде апостериорного распределения оценок параметров или переменных модели.

Необходимо выделить следующие особенности байесовского подхода: абсолютно все параметры и величины принято считать случайными, а именно точное значение параметров неизвестно исследователю, из чего следует то, что параметры являются случайными с точки зрения незнания исследователя; методы Байеса используются даже при нулевом объеме выборки. В этом случае значения априорного и апостериорного распределений равны; для оценки неизвестных переменных используют апостериорные распределения, то есть, найти решение задачи по оцениванию определенной величины, означает определить апостериорное распределение этой величины; основным инструментом подхода является формула (теорема) Байеса, а также такие правила, как *sum-rule* (если  $A_1, \dots, A_k$  – взаимоисключающие события, то одно из них происходит всегда) и *product-rule* (любую совместную плотность всегда можно разбить на множители).

Наряду с вышеперечисленными достоинствами байесовской методологии необходимо выделить ее недостатки. Начиная с 1930 гг. байесовская парадигма довольно часто подвергалась резкой критике и практически не находила применения по следующим причинам:

- в байесовских методах предполагается, что априорное распределение известно до начала наблюдений и не предлагается конструктивных способов его выбора;
- принятие решения при использовании байесовских методов в нетривиальных случаях требует колоссальных вычислительных затрат, связанных с численным интегрированием в многомерных пространствах;
- Фишером была показана оптимальность метода максимального правдоподобия, а следовательно – бессмысленность попыток придумать что-то лучшее.

На сегодняшний момент (начиная с 1990 гг.) ученые наблюдают возрождение методологии Байеса, методы которой оказались полезными для поиска решений многочисленных и достаточно серьезных вопросов и проблем в сфере машинного обучения и статистики.

## II. ТЕОРЕМА БАЙЕСА: ФОРМУЛА, ПРИМЕНЕНИЕ

Формула (теорема) Байеса является одной из фундаментальных теорем теории вероятностей и позволяет установить вероятность определенного события А, если имеет место возникновение другого взаимозависимого с ним статистического события В. Говоря иначе, применение теоремы Байеса дает возможность исследователю наиболее точно пересчитать вероятность, учитывая как информацию, полученную им ранее, так и новые данные более поздних наблюдений. Байесовская теорема может быть получена из фундаментальных аксиом теории вероятностей, а именно из условной вероятности. Формуле Байеса присущи определенные особенности, главная из которых состоит в том, что для использования теоремы на практике необходимо значительное количество вычислений и расчетов, вследствие этого оценки байесовской парадигмы нашли активное применение только в период после революции, произошедшей в сфере сетевых и компьютерных технологий.

Перед формулированием байесовской теоремы вероятности, используемые в ней, подвергались многочисленным вероятностным интерпретациям. В одной из них устанавливался факт того, что вывод формулы зависит напрямую от использования особого подхода к статистическому анализу. При применении байесовской интерпретации вероятности теорема Байеса показывает, как степень личного доверия способна значительно меняться после определенного числа возникших событий. В этом состоят основные выводы Байеса, ставшие фундаментальными для байесовской статистики. Однако теорема нашла свое применение не только в байесовском анализе, но и активно используется в большом количестве иных расчетов.

Теорема Байеса получила свое название в честь её автора Томаса Байеса (1702–1761) – английского математика, первым предложившего применение сформулированной им теоремы для корректировки убеждений, опираясь на свежеполученную информацию [4]. Сущность байесовской парадигмы состоит в том, что при нахождении исследователем новой информации она дает основу для измерения вероятностей, которые обусловлены связанными между собой событиями. Формула Байеса имеет следующий вид [1]:

$$P_A(H_i) = \frac{P(H_i) * P_{Hi}(A)}{\sum_{i=1}^n P(H_i) * P_{Hi}(A)},$$

где  $P(H_i)$  – вероятности гипотез до опыта (априорные вероятности);  $P_{Hi}(A)$  – условные вероятности возникновения события А при выборе i-й гипотезы;  $P_A(H_i)$  – условная вероятность i-й гипотезы после возникновения события А (апостериорная или послеопытная вероятность).

С помощью данной теоремы удастся изменить значение вероятности на основе более поздних полученных сведений. Интерпретация байесовского подхода имеет следующий вид: пусть к началу реорганизации определенного объекта имеются n гипотез  $H_1, H_2, \dots, H_n$  о возможных его состояниях. На основе статистических данных за прошлые годы можно приписать им априорные вероятности  $P(H_1), P(H_2), \dots, P(H_n)$ . Затем проводится эксперимент (реализуется проект), в результате которого может наступить или не наступить событие А. Опытным путем определяются условные вероятности возникновения события А при выборе i-й гипотезы  $P_{Hi}(A)$  как частоты наблюдения события А. При условии наступления события А делается переоценка веры в справедливость каждой гипотезы посредством замены вероятностей  $P(H_i)$  на вероятности  $P_A(H_i)$ .

Необходимо отметить, что в случае отсутствия статистических данных об априорных вероятностях наблюдений или гипотез о возникновении события А байесовскую методологию применить невозможно, так как подобная «формализация» теряет физический смысл [6].

## III. БАЙЕСОВСКАЯ СТРАТЕГИЯ ОЦЕНКИ ДОСТОВЕРНОСТИ ВЫВОДОВ В ЭКСПЕРТНЫХ СИСТЕМАХ

Байесовская стратегия оценки выводов находит все более частое применение в науке, экономике и промышленности. Для детального изучения байесовской стратегии необходимо рассмотреть пример оценки достоверности гипотезы на ее основе, а также обосновать использование теоремы Байеса для экспертных систем.

Для начала необходимо привести более расширенный вариант формулы Байеса с учетом некоторого события Е, связанного с событиями  $H_1, H_2, \dots, H_n$ . Вероятности события Е известны при том условии, что какое-то из событий  $H_1, H_2, \dots, H_n$  наступило:  $P(E/H_1), P(E/H_2), \dots, P(E/H_n)$ . Предположим, нам известно, что событие Е произошло. В этом случае вероятность того, что какое-либо из событий  $H_i$  ( $i=1, \dots, n$ ) наступило, определяется по следующей формуле [2]:

$$\begin{aligned} P(H_i|E) &= \frac{P(E|H_i) * P(H_i)}{P(E|H_1) * P(H_1) + P(E|H_2) * P(H_2) + \dots + P(E|H_n) * P(H_n)} \\ &= \frac{P(EH_i)}{P(E)} \end{aligned}$$

События  $H_1, H_2, \dots, H_n$  – гипотезы, а событие Е называется свидетельством. Вероятности гипотез  $P(H_i)$  без учета свидетельства (т.е. без учета того, произошло событие Е или нет) называются доопытными (априорными), а вероятности  $P(H_i/E)$  – послеопытными (апостериорными). Величина  $P(EH_i)$  – совместная вероятность событий Е и  $H_i$ , т.е. вероятность того, что произойдут оба события вместе. Величина  $P(E)$  – полная (безусловная) вероятность события Е. В экспертных системах (компьютерные системы, которые способны частично заменить специалиста-эксперта в разрешении

конкретной проблемной ситуации) теорема Байеса может использоваться для оценки вероятностей заключений продукционных правил на основе данных о достоверности их посылок. Заключение (выводы) в этом случае соответствуют гипотезам в теореме Байеса, а посылки – свидетельствам. Обычно посылка правила в экспертных системах содержит несколько условий. Вероятности  $P(H_i)$  и  $P(E/H_i)$  определяются на основе статистических данных с использованием основных формул, применяющихся в теории вероятностей (формулы умножения вероятностей, формулы сложения вероятностей). В качестве наглядного примера рассмотрим экспертную систему, которая помогает оценить условия труда рабочих в определенной организации. В следующей таблице представлены данные о 5 тыс. рабочих (у 315 обнаружено заболевание, возникшее в связи с их профессиональной деятельностью):

ТАБЛИЦА I ПРИМЕР ОЦЕНКИ ДОСТОВЕРНОСТИ ГИПОТЕЗЫ

Условие	Значение	Кол-во ситуаций обнаружения проф. заболеваний	Кол-во рабочих, у которых нет проф. заболевания
Контакт с вредными веществами	Постоянно	262	237
	Часто	37	517
	Периодически	14	1011
	Отсутствует	2	2920
Физические нагрузки	Большие	168	927
	Средние	111	1847
	Нет	36	1911
Нервное напряжение	Есть	202	1532
	Нет	113	3153
Общие заболевания	Есть	196	2011
	Нет	119	2674

а. Источник: [2]

Необходимо дать оценку вероятности появления болезни у рабочего, работа которого почти каждый день связана с контактом с вредными веществами; также его деятельность связана с физическими нагрузками среднего уровня, нервное напряжение отсутствует, общих заболеваний у рабочего также нет.

В приведенной выше таблице за гипотезы берется уровень состояния здоровья рабочих:  $H_1$  – наличие профессиональной болезни,  $H_2$  – отсутствие профессиональной болезни.

За свидетельство принимается совокупность 4х факторов, описывающих работу: присутствие контакта с вредными для здоровья веществами, физические нагрузки, присутствие нервного напряжения, наличие общих заболеваний (свидетельства  $E_1$ ,  $E_2$ ,  $E_3$ ,  $E_4$ ). Совокупность всех 4х факторов принимается за событие  $E$ .

По формуле Байеса определим вероятности, которые нужны для дальнейших расчетов. Априорные вероятности гипотез (вероятности нахождения профессиональных заболеваний и их отсутствие без учета рабочих условий):

$$P(H_1)=315/5000=0,063; P(H_2)=120/173=0,937.$$

Наблюдаемым свидетельством выступает совокупность 4х событий, которые наблюдаются совместно: частый контакт с вредными веществами, средний уровень

физической нагрузки, отсутствие нервного напряжения и общих заболеваний. При условии, что эти события считаются независимыми, определяются условные вероятности свидетельства с помощью формулы произведения вероятностей:

$$P(E/H_i) = P(E_1, E_2, E_3/H_i) = P(E_1/H_i) P(E_2/H_i) P(E_3/H_i), i=1,2.$$

Затем определим величины, которые нужны для дальнейшего использования формулы произведения вероятностей:

$$P(E_1/H_1)=37/315=0,117; P(E_2/H_1)=111/315=0,352;$$

$$P(E_3/H_1)=113/315=0,359; P(E_4/H_1)=119/315=0,378;$$

$$P(E_1/H_2)=517/4685=0,11; P(E_2/H_2)=1847/4685=0,394;$$

$$P(E_3/H_2)=3153/4685=0,673; P(E_4/H_2)=2674/4685=0,571.$$

$P(E_1/H_1)$  – вероятность того, что сотрудник часто контактирует с вредными веществами, при том условии, что в будущем у него будет обнаружено профессиональное заболевание. Данная величина отображает, с какой частотой у сотрудников, часто контактирующих с опасными веществами, находят профессиональное заболевание.

Подставив найденные величины в формулу произведения вероятностей, получаем:  $P(E/H_1) = 0,117 \cdot 0,352 \cdot 0,359 \cdot 0,378 = 0,006$ ;  $P(E/H_2) = 0,11 \cdot 0,394 \cdot 0,673 \cdot 0,571 = 0,017$ .

$P(E/H_1)$  – величина, показывающая вероятность рабочих условий при том условии, что в будущем у сотрудника обнаружат профессиональное заболевание. Найдем апостериорную вероятность обнаружения профессионального заболевания при фактических рабочих условиях:

$$P(H_i|E) = \frac{P(E/H_1) \cdot P(H_1)}{P(E/H_1) \cdot P(H_1) + P(E/H_2) \cdot P(H_2)} = \frac{0,006 \cdot 0,063}{0,006 \cdot 0,063 + 0,017 \cdot 0,937} = 0,022.$$

Полученная апостериорная вероятность выступает наиболее четкой оценкой вероятности обнаружения у сотрудника организации профессионального заболевания, чем априорная вероятность  $P(H_1)$ , которая была получена на основе данных без учета рабочих условий. Стоит заметить, что рассчитанная апостериорная вероятность (0,022) меньше, чем априорная вероятность (0,063). Из этого можно сделать вывод, что наблюдаемые свидетельства (частый контакт с вредными веществами, физическая нагрузка среднего уровня, отсутствие нервного напряжения и общих заболеваний) полностью подтверждают гипотезу, которая говорит о том, что профессиональное заболевание у сотрудника не появится.

#### IV. БАЙЕСОВСКОЕ МОДЕЛИРОВАНИЕ (МЕТОД БАЙЕСОВСКИХ СЕТЕЙ)

Понятием «байесовские сети» обозначают графические структуры, которые служат для представления вероятностных отношений между

значительным числом переменных, а также структуры, необходимые для осуществления вероятностного вывода на основе представленных переменных [4].

Альтернативным названием метода байесовских сетей выступает понятие «байесовская классификация». Изначально байесовскую классификацию применяли для формализации экспертных знаний в экспертных системах. В настоящее время байесовская классификация используется также в качестве одного из методов Data Mining. Понятие «Data Mining» обозначает добычу данных, их глубинный анализ, также данное понятие применяется в целях обозначения системы методов нахождения в статистических данных неизвестных ранее, практически полезных знаний, нужных для принятия решений в различных областях деятельности людей. В большей степени простым вариантом метода, использующего байесовскую классификацию, является «наивная» классификация или, иными словами, наивно-байесовский подход. С использованием данного подхода решаются задачи классификации. «Наивная» классификация является довольно легким для понимания методом классификации. Название «наивная» берет начало из предположения о том, что признаки являются взаимно независимыми. Следует выделить перечень основных свойств наивной классификации [5]:

1. Использование всех переменных и нахождение абсолютно всех связей между ними.
2. Относительно переменных существует два предположения: все переменные имеют одинаковую значимость; все переменные независимы статистически.

Следует выделить ряд преимуществ байесовских сетей: в модели устанавливаются связи между всеми переменными, что предоставляет возможность легко подвергать обработке случаи, где отсутствуют значения некоторых параметров; байесовские сети довольно легки в интерпретации и дают возможность при прогностическом моделировании проводить анализ по сценарию «что, если»; байесовский метод дает возможность совмещать закономерности, которые были выведены из статистических данных, и знания экспертов, которые были получены фактически; применение байесовских сетей на практике позволяет избежать чрезмерного усложнения модели. Помимо перечисленных выше преимуществ, наивно-байесовский подход обладает такими недостатками, как [5]: перемножать условные вероятности корректно только тогда, когда все входные переменные действительно статистически независимы; хотя часто данный метод показывает достаточно хорошие результаты при несоблюдении условия статистической независимости, но теоретически такая ситуация должна обрабатываться более сложными методами, основанными на обучении байесовских сетей; невозможна непосредственная обработка непрерывных переменных – требуется их преобразование к интервальной шкале, чтобы атрибуты были дискретными; однако такие преобразования иногда могут приводить к потере значимых закономерностей; на результат классификации в наивно-байесовском подходе влияют только индивидуальные значения входных

переменных, комбинированное влияние пар или троек значений разных атрибутов здесь не учитывается. Это могло бы улучшить качество классификационной модели с точки зрения ее прогнозирующей точности, однако, увеличило бы количество проверяемых вариантов.

Байесовская классификация нашла широкое применение на практике. Не так давно ее было предложено использовать для персональной фильтрации спама. Первый фильтр был создан Полем Грахемом. Для правильной работы алгоритма необходимо выполнение двух требований.

Первое требование – необходимо, чтобы классифицируемый объект обладал достаточным количеством признаков. Этому идеально удовлетворяют все слова писем пользователя, за исключением совсем коротких и редко встречающихся. Второе требование – постоянное переобучение и пополнение набора «спам – не спам». Данные условия корректно работают в локальных почтовых клиентах, так как поток «не спама» у конечного клиента довольно постоянен, а если изменения присутствуют, то они не носят стремительного характера.

Однако для всех клиентов сервера точное определение потока "не спама" представляет значительную сложность, так как одно и то же письмо, выступающее для одного клиента спамом, для другого клиента спамом не является. Словарь получается слишком объемного размера, также не присутствует четкого разделения на спам и «не спам», вследствие этого качество классификации, в конкретном случае решение задачи фильтрации писем, значительно понижается. Использование методологии Байеса в формировании статистических выводов дает возможность совсем по-иному воспринимать и исследовать оцениваемые модели. Он позволяет оперировать не только полученными оценками, а также соответствующими вероятностными распределениями, применять имеющиеся в разных формах априорные знания исследователя относительно оценок параметров модели. Это дает возможность получать большие объемы исходной информации и точнее описывать структуру и другие характеристики исследуемой модели.

Научный руководитель статьи от Финансового университета при Правительстве РФ доц. каф. «Системный анализ в экономике» Звягин Л.С.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Айвазян С.А. Байесовский подход в эконометрическом анализе - М.: Синергия, 2008. 332 с.
- [2] Звягин Л.С. Применение байесовских интеллектуальных технологий при моделировании управлением операционными рисками// Экономика. Управление. Право. 2011. № 11-1. С. 33-42.
- [3] Звягин Л.С. Системный анализ в исследовании природных систем// Международная конференция по мягким вычислениям и измерениям. 2014. Т. 2. С. 227-233.
- [4] Шапиро Л.Д. Экономико-математическое моделирование – Томск: Изд-во Томск. ун-та, 1987. 247 с.
- [5] Mayo, M., & Mitrovic, A. Optimising ITS behavior with Bayesian networks and decision theory. International Journal of Artificial Intelligence in Education, 2001. 241 p.