Вісник ЛДУБЖД
Bulletin of Lviv State University of Life Safety
https://journal.ldubgd.edu.ua/index.php/Visnuk
ISSN 2078-4643 (print), ISSN 2708-1389 (online)

UDC 004.9:656.1
DOI: 10.32447/20784643.28.2023.06

Л. П. Гащук, П. М. Гащук
Львівський державний університет безпеки життєдіяльності, м. Львів, Україна
ORCID: https://orcid.org/0000-0002-5522-2757 – Л. П. Гащук
https://orcid.org/0000-0002-2345-4879 – П. М. Гащук
reivenor2099@gmail.com

ОКРЕМІ АСПЕКТИ ІНТЕРПРЕТАЦІЇ ЗМІСТУ ТЕОРЕМИ ГІПОТЕЗ У СИТУАЦІЯХ НЕВИЗНАЧЕНОСТІ

Розглядаються різні аспекти об'єктивного прояву і суб'єктивного сприйняття невизначеності. Проведено короткий аналіз змісту, що його зазвичай вкладають у поняття «невизначеність», «апріорність», «апостеріорність». Основна мета роботи — критично оцінити можливість і доцільність вивчення в навчальних і застосування в наукових дисциплінах теореми Т. Бейза (Th. Bayes) як знаряддя долати невизначеність у процесі пізнання й ухвалення рішень. Виклад матеріалу спирається на абстрактні мотивуючі приклади.

Зроблено такі загальні висновки.

Якщо з'явилися справді віри гідні апріори, то це означає, що цілком напевне раніше провадилися ґрунтовні дослідження, на тлі яких бейзівський висновок — це виконане нашвидкуруч бліде і не завжди переконливе продовження цих досліджень. Тому не дивно, що був період, коли теорема Бейза майже зникла із засобів статистичних досліджень.

В середині ХХ століття підхід Бейза знову виринув серед інструментарію для пошуку допустимих і мінімаксних оцінок в типових задачах оцінювання невідомого параметра ймовірнісного розподілу. Бейзіанство тепер набуло популярності ще й в алгоритмах навчання штучного інтелекту на основі великих обсягів даних, що засновані на людському досвіді.

Та насправді це все радше прояв моди, ніж визнання особливих потенцій теореми гіпотез. Те, що штучний інтелект облюбував собі саме технологію бейзіанства, нічого особливого не означає. Його раніше чи пізніше природний інтелект легко наверне на інший, розумніший шлях — знати б найкращий шлях.

Загалом теорема Т. Бейза не заслуговує стати предметом ретельного вивчення та активного поширення як надзвичайно ефективного інструментарію дослідження й пізнання.

Ключові слова: невизначеність, суб'єктивізм, об'єктивізм, апріорна ймовірність, апостеріорна ймовірність, теорема гіпотез/причин, формула Т. Бейза

L. P. Hashchuk, P. M. Hashchuk
Lviv State University of Life Safety, Lviv, Ukraine

CERTAIN ASPECTS OF THE INTERPRETATION OF THE CONTENTS OF HYPOTHESES THEOREM IN UNCERTAINTY SITUATIONS

Various aspects of objective manifestation and subjective perception of uncertainty are considered. A brief analysis of the content, which is usually included in the concepts of "uncertainty", "a priori", "a posteriori", was carried out. The main goal of the work is to critically assess the possibility and expediency of studying T. Bayes' theorem in educational and scientific disciplines as a tool to overcome uncertainty in the process of cognition and decision-making. The presentation of the material is based on abstract motivating examples.

The following general conclusions were made.

If truly credible a priori appeared, then this means that thorough research was certainly conducted before, against the background of which Bayesian inference is a hastily performed pale and not always convincing continuation of these studies. Therefore, it is not surprising that there was a period when Bayes' theorem almost disappeared from the tools of statistical research.

In the middle of the 20th century, the Bayesian approach again emerged among the tools for finding admissible and minimax estimates in typical problems of estimating an unknown parameter of a probability distribution. Bayesianism has now also gained popularity in artificial intelligence learning algorithms based on large volumes of data based on human experience.

But in fact, this is more a manifestation of fashion than a recognition of the special potential of the theorem of hypotheses. The fact that artificial intelligence chose Bayesian technology for itself does not mean anything special. Sooner or later, his natural intelligence will easily convert him to another, more reasonable path - he would know the best way.

In general, the theorem of T. Bayes does not deserve to be the subject of careful study and active dissemination as an extremely effective tool for research and cognition.

Key words: uncertainty, subjectivism, objectivism, a priori probability, a posteriori probability, hypotheses/causes theorem, T. Bayes formula

Вступ. Мета роботи — критично оцінити можливість і доцільність вивчення в навчальних дисциплінах і застосування в наукових дослідженнях теореми Т. Бейза (Т. Байєса, Th. Bayes) як знаряддя долати невизначеність у процесі висування і ухвалення рішень. Ключову роль тут відіграє поняття ймовірнісності як у трактуванні класичної теорії ймовірності, так і у трактуванні математичної статистики. Дослідження призначене радше для студентів-курсантів і ад'юнктів-аспірантів, які хотіли б зорієнтуватись у виборі пріоритетів у засобах пізнання.

Теорія ймовірностей та класична статистика, — це особлива пара ніби протиставних одна одній наук. Теорія ймовірностей оперує радше теоретично постульованим (а не означуваним) поняттям імовірності і всюди в своїх задачах передбачає наявність відомого ймовірнісного закону. Виходячи з цього, вона намагається дізнатись і розповісти, як будуть вести себе результати якихось конкретних спостережень. А от у статистичних задачах усе навпаки: відомими є результати конкретних спостережень (така собі вибірка), а треба з'ясувати, який конкретно ймовірнісний закон проявляв себе під час ведення спостережень. До слова, спочатку статистика була просто «державною арифметикою» («статистика» є похідним радше від латинського *status* — держава, аніж від латинського *status* — стан, становище).

В теорії ймовірностей використовують низку характеристик ймовірнісних законів, серед яких — моменти, семіінваріанти, характеристичні функції… Теорія ймовірностей — це точна наука, кореневу систему якої складають аксіоми (див., приміром, [1, 2]).

В математичній статистиці для аналізу вибірки загалом чи її фрагментів також використовують числові функції — так звані статистики. Статистиками є найрізноманітніші середні (чисті чи зважені, тобто модульовані вагами): арифметичне, квадратичне, геометричне, гармонічне… (кількість різновидів середніх може бути нескінченною). До статистик можна зарахувати й часові ряди разом із засобами інтерполяції, згладжування, екстраполяції. Статистичним поняттям є, приміром, фізична (що не має стосунку до інформації) ентропія. Вона застосовна суто до масових (в жодному разі не до одиничних) явищ.

Найскладнішим є питання чи ймовірнісність — це відображення реальності, чи вона цілком належить розумові. За Ст. Джевонсом: «ймовірність в речах, які ймовірні, чи у розумі, який вважає їх імовірними?». Як би хто не відповідав на це запитання, ймовірність потрапляє в контекст наших міркувань значно частіше, ніж це усвідомлюємо.

Привабливою в статистиці є можливість частотного (частісного, частотнісного) означення й вимірювання ймовірності (Р. Мізес намагався просунути такий підхід і в теорію ймовірностей, але там емпіричного та й будь-якого іншого її означення не дуже потребують). К. Поппер (K. Popper) відстоював думку, що ймовірність не є суто фізичною властивістю самого по собі об'єкта, а характеризує внутрішню налаштованість (propensity) експериментальної ситуації сприяти тому чи іншому прояву властивостей об'єкта, раз він у неї втрапив. Отже йдеться про схильність ситуації (що її разом формують об'єкт, ситуативні умови та суб'єкт) до вияву певних подій з тими чи іншими характерними частотами. Властивості чи стани об'єкта — не побачити без того, щоб дати можливість їм проявитися в певній ситуації «на очах» суб'єкта. Якими б не були вразливими для критики згадані поняття, насправді вони дуже корисні.

Тоді як ймовірність події є постійною величиною, відносна частота настання події коливається від серії до серії випробувань. Запроваджене Р. Мізесом частотне означення ймовірності не дуже сприяло бездоганному розкриттю природи випадковості — воно типово феноменологічне. Але спосіб (хай наближеного) обчислення ймовірності на основі деяких емпіричних спостережень — неабияка цінність.

Суб'єктивність і об'єктивність. Колись давно Г. В. Лейбніц критикував підхід Аристотеля і його прихильників за те, що він і вони пов'язували ймовірнісні судження із загальноприйнятими думками (поглядами і гадками авторитетів). І наводив приклад Коперника, коли той був одинокий зі своєю думкою з відомого приводу, а насправді ця думка була ймовірнішою за іншу панівну в той час.

Але дуже згодом математики Ф. П. Рамсей, Л. Дж. Севідж, Бруно де Фінетті взялися за відродження «суб'єктивної» («людської», «пересоналістської») ймовірності. У них ймовірність трактується як вимірник корисності чи привабливості, як вимірник довіри чи раціональної (розумової) віри. Знову запанував прагматизм, озброєний тепер аксіологією. Зрештою, насправді ніби й не існує можливості чисто випадкову об'єктивну у всіх сенсах ситуацію ні розпізнати і показати, ні змоделювати і використати (заперечувати будь-яку визначеність доводиться хіба що генеруванням квазі- чи псевдовипадкових чисел, що доволі непросто). Саме Леонард Дж. Севідж розбурхав найбільшу суперечку в сучасній статистичній науці [3]. Його теорія, пов'язана саме з персоналістичним тлумаченням ймовірності, кинула виклик панівній у той час частотнісній школі.

До арсеналу цих учених-суб'єктивістів потрапила так звана теорема, так званий принцип Бейза. Теорема (деколи — закон, правило, принцип, формула) Томаса Бейза (1701—1761) вперше опублікована зусиллями Річарда Прайса у 1763 році [4] (зважмо, у нашому 2023-му відтоді минуло 260 років). Мартін Купер [5] та Шарон Мак-Грейн [6] прихильники того, що послуга Річарда Прайса була настільки вагомою, що теоремі пасувало б дати ім'я Бейза-Прайса. Більш звичне тепер формулювання теореми у своїй праці 1812 року «Аналітична теорія ймовірностей» подав П'єр-Симон Лаплас. Гарольд Джеффріс підвів під нього аксіоматичну основу. А ще Шарон Мак-Грейн переконувала, що це така собі теорія, яка ніколи не помре, що саме завдяки їй зламали код *Enigma* під час Другої світової війни та вислідили російські підводні човни в пізніші часи, що вона вийшла тріумфатором із двох століть суперечок. А чи так воно? Та як би воно не було, відкрите математиком і проповідником XVIII століття правило стало наріжним каменем сучасної статистики [7].

Існують величезні прогалини (чи не прірви?) між сутнісно інтуїтивними (евристичними, а тому й дуже суб'єктивними) початковими уявленнями про перманентності, принципово об'єктивними даними первинних систематичних досліджень, емоціональними суб'єктивними висновками (підказками, домислами) теореми Бейза, суттєво об'єктивним реальним світом ухвалених рішень. Виникає питання, а що в цьому ланцюгу суцільних складнощів робить примітивна теорема Бейза? У будь-якому курсі «Теорія ймовірності» ця теорема багато місця не займає [8, 9]. Але на межі теорії ймовірностей і теорії ухвалення рішень вона раптом набуває якоїсь вибухової ваги.

Теорему можна викласти дуже простою мовою [10], а можна вдатися до вишукано точного апарату [11, 12]. Її застосовують до подолання невизначеності в математиці, медицині, бізнесі і фінансах, інженерії, машинному навчанні тощо [13—15]. Інколи вдаємося до неї навіть у нашому повсякденному житті, коли намагаємося ухвалити рішення на підставі щойно отриманої нової інформації (нових даних).

В роботі [16], приміром, здійснено огляд наслідків і перспектив застосування бейзівського аналізу в царині дослідження підприємництва. Загальний висновок статті оптимістичний: широке використання бейзівського аналізу дозволить поглибити розуміння феномену підприємництва, сприяючи ретельнішому вивченню процесу ухвалення рішень у ситуаціях високого рівня невизначеності та впливу попереднього досвіду підприємницької діяльності на схвалювані нові рішення.

Відповідно до [17] апостеріор завжди існує і однозначно визначається апріором, даними поточних спостережень та статистичною моделлю. Було б добре використовувати фактичні попередні знання [17] чи якісь надійні еталонні апріори [18], але часто намагаються керуватись і невідповідними пріорами, отримуючи невідповідні постеріори. А далі залучають мистецтво інтерпретації. Звісно, інколи можна віднайти хороші можливості вдатись до конвергенції попередніх даних [19, 20].

Один одному часто протиставляють бейзівський підхід до статистичного аналізу і підхід Р. А. Фішера, засновника сучасної статистики, що запропонував ідею довірчого висновку в першій половині 20 століття. Але довірчий висновок не завжди можна було якісно припасувати до багатопараметричних статистичних задач. Та починаючи десь з 2000 року, почала визрівати технологія узагальненого довірчого висновку (Generalized Fiducial Inference — GFI), яка полягає в ретельному перенесенні випадковості з масиву даних у простір параметрів за допомогою такого собі оберненого рівняння, що генерує дані, без використання теореми Бейза [21, 22]. Тож цілком можливо (якщо не брати до уваги запити штучного інтелекту), що бейзівський підхід знову відійде в тінь.

Як слушно зазначено в [23], корисна теорія потребує логічних правил побудови та розумної інтерпретації ймовірностей. Імовірності обчислюються на засадах формалізації здогадки про те, що майбутнє, ймовірно, буде подібним до минулого. Імовірності об'єктивно випливають із подібності, але самі подібності є суб'єктивними судженнями індивідів [23 — 25]. А тому на статистику й індукцію справді можна покладатись як на мистецтво вгадувати [26]. І всі розуміють, що виявляти будь-які експертні знання доводиться «суб'єктивно, але науково» [27]. Звісно, теореми залишатимуться вірними за будь-якої інтерпретації ймовірності, якщо задовольнятимуться формальні аксіоми. Статистика і теорія ймовірностей, філософія і теорія ухвалення рішень можуть і повинні розглядатись у єдності [28].

Про знайоме й очевидне: засадничі поняття і терміни. Позначати ймовірності настання подій (чи просто ймовірності подій) A і B прийнято через $P(A)$ і $P(B)$. Символ P зазвичай використовують саме для позначення ймовірностей, оскільки слово «ймовірність» латиницею часто записується з букви P — *probabilitas* латиною, *probability* англійською, *probabilidad* іспанською, *probabilité* французькою, *probabilitá* італійською…

Сумою подій A і B називають подію $C = A + B = A \cup B$, яка полягає в настанні хоча б якоїсь з подій A і B (тобто в настанні чи A, чи B, чи разом A з B). І взагалі, сумою декількох подій називають подію, яка полягає в настанні хоча б однієї з цих подій.

Добутком подій A і B називають подію $C = AB = A \cap B$, яка полягає в сумісному настанні A і B (тобто в настанні разом і A, і B). І взагалі, добутком декількох подій називають подію, яка полягає в сумісному настанні всіх цих подій.

Дві події є незалежними, якщо ймовірність настання однієї з них не залежить від того, настала чи не настала інша. Подію A називають залежною від події B, якщо ймовірність події A змінюється через те, що настала подія B. Імовірність події A, визначену за умови настання події B, позначають як $P(A|B)$ і називають умовною. Її формально означують через відношення

$$P(A \mid B) = \frac{P(AB)}{P(B)}.$$

Тож умову незалежності події A від події B можна записати у вигляді

$$P(A \mid B) = P(A),$$

а умову залежності, звісно, — у вигляді

$$P(A \mid B) \neq P(A).$$

Легко бачити, що

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{P(B)\,P(A \mid B)}{P(A)}. \tag{1}$$

Ця остання формула сприймається як цілком зрозуміла і майже очевидна. Але… (і про це згодом).

Імовірність суми (загалом сумісних) подій можна обчислити за формуло:

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i} P(A_i) - \sum_{i,j} P(A_i A_j) + \sum_{i,j,k} P(A_i A_j A_k) - \ldots + (-1)^{n-1} P(A_1 A_2 \ldots A_n) \tag{2}$$

(суми охоплюють різні значення індексів: по одному i; по два i, j; по три i, j, k;…).

Натомість ймовірність добутку подій в загальному випадку можна обчислити за формулою формуло:

$$P\left(\prod_{i=1}^{n} A_i\right) = \sum_{i} P(A_i) - \sum_{i,j} P(A_i + A_j) + \sum_{i,j,k} P(A_i + A_j + A_k) - \ldots + (-1)^{n-1} P\left(\sum_{i=1}^{n} A_i\right). \tag{3}$$

Імовірність добутку хоч яких несумісних подій, звісно, дорівнює нулю. Співвідношення (2) і (3) — це так звані формули додавання ймовірностей.

Дуже важливою є також теорема (формула) множення ймовірностей, що визначає ймовірність добутку n подій як добуток

$$P(A_1 A_2 \ldots A_n) = P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i \,|) =$$

$$= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 A_2)\ldots P(A_n \mid A_1 A_2 \ldots A_{n-1}).$$

Зрозуміло, додавання й множення подій та додавання й множення ймовірностей — різні речі. Цікаво, що якщо події $A_1, A_2, \ldots, A_n$ не-незалежні, то можна орієнтуватись на те, що

$$-\left(1 - \frac{1}{n}\right)^{n} \leq P(A_1 A_2 \ldots A_n) - P(A_1)P(A_2)\ldots P(A_n) \leq (n-1)n^{-n/(n-1)}.$$

Зазвичай теореми додавання та множення ймовірностей доводиться використовувати разом. Наслідком теорем додавання та множення ймовірностей є так звана формула повної ймовірності, що випливає з таких міркувань. Хай треба визначити ймовірність деякої події A, яка стається разом з якоюсь однією з подій

$$H_1,\ H_2,\ \ldots,\ H_n$$

які створюють повну групу несумісних подій. Ці останні події доречно називати гіпотезами. Ймовірність події A в такому разі визначається за формулою власне повної ймовірності

$$P(A) = \sum_{i=1}^{n} P(A \mid H_i) P(H_i), \tag{4}$$

тобто ймовірність події A обчислюється як сума усіх добутків ймовірності $P(A|H_i)$ події A унаслідок реалізації гіпотези H_i на ймовірність цієї гіпотези.

Все щойно викладене — це засади будь-якого курсу «Теорія ймовірності»* [8, 9]. Доречним буде запитання: ну що може бути простішим, і про що тут є ще якийсь сенс говорити? Але все-таки є одна похідна формула, до якої треба ставитися по-особливому.

«Майже очевидне» у цілком очевидному. Хай Ω — так звана множина елементарних подій, A — алгебра на множині Ω, елементами якої є деякі підмножини множини Ω — не обов'язково елементарні події, P — задана на алгебрі A ймовірнісна міра. Трійка (Ω, A, P) — це ймовірнісний простір. Стосовно окресленого ймовірнісного простору (Ω, A, P) можна висунути таке твердження.

Хай у просторі (Ω, A, P) задана повна система подій $\mathrm{X} \subset \mathrm{A}$, така що для будь-яких двох подій $X \in \mathrm{X}$ і $Y \in \mathrm{A}$ існує умовна ймовірність $P(Y \mid X)$ (імовірність події Y за умови настання події X). Тоді для будь-яких подій $A, B \in \mathrm{A}$, таких що $P(B) \neq 0$ і існує умовна ймовірність $P(B \mid A)$, існує водночас й (обернена) умовна ймовірність $P(A \mid B)$ (чому б не згадати формулу (1)), визначувана за так званою формулою Бейза

$$P(A \mid B) = \frac{P(A)\, P(B \mid A)}{\sum\limits_{X \in \mathrm{X}} P(X)\, P(B \mid X)} .$$

Висновок Бейза легко безпосередньо тлумачити як теорему гіпотез. Хай існує повна група несумісних гіпотез H_1, H_2, …, H_n, ймовірності $P(H_1)$, $P(H_2)$, …, $P(H_n)$ яких заздалегідь ніби достеменно відомі. Проведено дослід, унаслідок якого спостережена якась подія A. Виникають питання, чи позначиться це на ймовірностях гіпотез, чи не потрібно їх переоцінити?

Отож ідеться про те, щоб оцінити умовну ймовірність $P(H_i \mid A)$ кожної гіпотези H_i. За теоремою множення ймовірностей

$$P(AH_i) = P(A)P(H_i \mid A) = P(H_i)P(A \mid H_i)\ ,\ i = \overline{1, n}\ ,$$

а отже цінним стає співвідношення

$$P(A)P(H_i \mid A) = P(H_i)P(A \mid H_i)\ ,\ i = \overline{1, n}\ ,$$

з якого випливає рівність

$$P(H_i \mid A) = \frac{P(H_i)P(A \mid H_i)}{P(A)}\ ,\ i = \overline{1, n}\ .$$

*Тут справді «ймовірність» в однині. Українською розумніше називати дисципліни «Теорія автомобіля», а не «Теорія автомобілів», «Будова автомобіля», а не «Будова автомобілів»...

А беручи до уваги формулу повної ймовірності (4), останній вираз можна подати у формі

$$P(H_i \mid A) = \frac{P(A \mid H_i)P(H_i)}{\sum\limits_{i=1}^{n} P(A \mid H_i)P(H_i)}\ ,\ i = \overline{1, n}\ . \qquad (5)$$

Отже спочатку ймовірність гіпотези H_i дорівнювала $P(H_i)$, а потім набула значення $P(H_i \mid A)$ відповідно до (5). Це власне і є формула Бейза. Вона показує, як за *апріорними* ймовірностями $P(H_i)$ (ймовірностям подій H_i до того, як подія A відбулася) знайти *апостеріорні* ймовірності (після того, як подія A відбулася). Якщо трактувати події H_i як причини, то формула Бейза виражатиме «теорему про ймовірності причин» — гарне тлумачення.

У науці якраз завзято й оперують гіпотезами. Гіпотези в певному сенсі завжди мають якусь ймовірнісність, тобто — це не тверді знання. І саме ймовірність часто править за міру переконаності. У термінах теорії ймовірностей на такій собі апостеріорній ймовірності обов'язково позначаються отримувані нові емпіричні дані, а також і ймовірності апріорні (давні). А оновлювати свої переконання нібито дає змогу якраз Теорема Бейза.

За Томасом Бейзом (Th. Bayes; прізвище Bayes транслітерують як Байєс, Бейєс, Баєс, а у нас — Бейз) те, наскільки у щось вірять після отримання емпіричних даних (апостеріорно), залежить не тільки від самих цих емпіричних даних, але й від того, наскільки у це вірили від самого початку (апріорно). Це трохи пригнічує, оскільки хочеться сподіватись на об'єктивність науки і завжди покладатись на отримувані емпіричні дані: наука — це ж бо процес набуття істотно нової інформації (радше істотно нових даних). Аж тут раптом доводиться зважати ще й на суб'єктивні попередні уявлення й теорії — суб'єктивні тому, що вони не спираються на істотно нові об'єктивні дані (їх ще не було). Та не слід забувати, що сам процес навчання, що передує емпіричному досвідові — це формування попередніх уявлень і теорій (теорії тепер частіше називають моделями), вагомість яких неможливо переоцінити.

Завжди щось є наслідком чогось. Формула повної ймовірності, приміром, — наслідок теореми додавання ймовірностей і теореми множення ймовірностей. А от формула Томаса Бейза (теорема про гіпотези чи причини) — наслідок теореми множення ймовірностей і формули повної ймовірності. Тож формула Бейза не є чимось самоцінним, але поти, поки її не намагаються наповнити непересічним змістом. Інакше кажучи, сама по собі теорема Бейза незаперечна (в математичному сенсі), але у практичних застосуваннях виникають нездоланні проблеми із розкриттям нею фізичної суті конкретних ситуацій. Через це теорема досі є джерелом суперечок в царині статистики, а провалля між сповідниками бейзівського та антибейзівського світоглядів стає дедалі нездоланнішим. І це при тому, що вона майже тривіальна, бо є, як зазначено, примітивним наслідком так званої теореми множення ймовірностей і формули повної ймовірності.

Відомо, що Бейз застосовував свою теорему у випадку неперервних ймовірнісних розподілів на інтервалі $(0,1)$. Якщо, приміром, в $n+m$ дослідах подія, що має невідому ймовірність p, настала n разів, то ймовірність того, що p належить підінтервалу (a,b) інтервалу $(0,1)$, дорівнює

$$\frac{\int_a^b x^n(1-x)^m\,dx}{\int_0^1 x^n(1-x)^m\,dx}. \tag{6}$$

В основу цього твердження лягла ідея, що якщо нема ніякої попередньої інформації про p, то чого б не припустити, що апріорна ймовірнісна густина параметра p однакова на всьому інтервалі $(0,1)$. Відомий такий показовий приклад: якщо $n=1$, $m=0$, $a=1/2$, $b=1$, то ймовірність того, що $p\in(1/2,1)$ дорівнює

$$\frac{\int_{1/2}^1 x(1-x)^0\,dx}{\int_0^1 x(1-x)^0\,dx}=\frac{\int_{1/2}^1 x\,dx}{\int_0^1 x\,dx}=\frac{3}{4}. \tag{7}$$

Але виявляється, мало охочих довіряти результатам (6), (7) (бо, перш за все, чого б це покладатись на рівномірний апріорний розподіл?).

Усі можливі альтернативи відразу, з першого погляду, що не дивно, справді здаються ніби рівноправними. Аби визрів якийсь пріоритет, треба «придивитись» до цих альтернатив, вдаючись, приміром, до якогось змістовного критерію (чи простого вимірника).

В статистичному аналізі проводять один-декілька дослідів, спостережень, випробувань, саме після чого альтернативи перестають сприйматись як рівноправні. Тому не дивно, що так званий підхід-метод Бейза визнає апріорі рівноправність скінченної кількості альтернатив, з яких тільки апостеріорі вибудовується пріоритетність. Якщо невизначеність описується в термінах теорії ймовірностей, то ситуацію у разі скінченної кількості альтернатив, як виглядає, є підстави описати рівномірним розподілом ймовірностей. Саме цей принцип ніби слугує основою для апостеріорного аналізу.

С. Бернстейн та Р. Мізес ще десь перед 1920 р. зазначали, що за деяких умов багаторазове застосування теореми Бейза дає послідовність апостеріорних ймовірнісних розподілів, що збігається до певного істинного розподілу, незалежно від того, яким був апріорний розподіл. А тому ніби прийнятий апріорний розподіл ймовірностей на асимптотиці не позначається.

Руйнівним у світогляді бейзівців (прихильників і послідовників теорії Бейза) є насправді незнання правдивого апріорного розподілу. Тому за відсутності апріорних даних припускають, що $P(H_i)$ є однаковими, а це навряд чи правильно. Рівноймовірність стає спочатку ознакою цілковитого незнання (на противагу, скажімо, знанню про досконалість монети, якою бавляться реально і в статистиці).

Мотивуючі приклади. Говард Райфа [29] від імені Варда Едвардса наводить такий тест на суб'єктивне сприйняття ймовірності. У двох однакових мішечках містяться покерні фішки: в одному (його умовно названо зеленим і позначено як *G* — *Green*) — 70 зелених і 30 білих; у другому (його умовно названо білим і позначено як *W* — *White*) — 70 білих і 30 зелених. Усі фішки цілком однакові, якщо тільки не брати до уваги колір. З невідомо якого (випадково поданого) мішечка витягували і повертали кожного разу знову у мішечок по черзі 12 фішок. Серед цих фішок 8 виявились зеленими, а 4 — білими. У якій мірі можна бути впевненим у тому, що мішечок, з яким проводили експеримент, є зеленим?

Звісно, ймовірність подання для експерименту чи зеленого — $P(G)$, чи білого — $P(W)$ мішечка цілком однакова — ½. Дослід мав би змінити цю попередню оцінку: все-таки з'явились нові дані. Виглядає так, що мішечок радше зелений. Але як поміряти це «радше»?

Хай фішки з'являлись у порядку *GGWGWGGWGWGG* (надалі — це подія A; послідовність насправді не важлива, якщо фішки знову повертають у мішечок). Тож умовні ймовірності появи події A, коли мішечок зелений — $P(A|G)$, або коли білий — $P(A|W)$, можна обчислити як

$$P(A|G)=0,7\cdot 0,7\cdot 0,3\cdot\ldots\cdot 0,7=0,7^8\cdot 0,3^4,$$
$$P(A|W)=0,3\cdot 0,3\cdot 0,7\cdot\ldots\cdot 0,3=0,7^4\cdot 0,3^8.$$

За формулою Бейза ймовірність $P(G|A)$ того, що A — результат експериментування з зеленим мішечком, становить

$$P(G|A)=\frac{P(A|G)\cdot P(G)}{P(A)}=\frac{P(A|G)\cdot P(G)}{P(A|G)\cdot P(G)+P(A|W)\cdot P(W)}=$$
$$=\frac{(0,7^8\cdot 0,3^4)\cdot\frac{1}{2}}{(0,7^8\cdot 0,3^4)\cdot\frac{1}{2}+(0,3^8\cdot 0,7^4)\cdot\frac{1}{2}}=\frac{0,7^4}{0,7^4+0,3^4}=0,9674$$

Це ж ледь не достовірна подія.

Впадає у вічі надмірна штучність ситуації. По-перше, достовірно відомо, що різних мішечків є тільки по одному, але потім чомусь забули, де який з них. По-друге, хтось достеменно з'ясував вміст кожного з мішечків, але чомусь лінькі було поставити мітки, щоб за потреби розпізнати їх. Можливо, тільки заради задачі — чому б ні!

А якщо зелених мішечків — 25, а білих — 75? В такому разі

$$P(G\mid A)=\frac{(0,7^8\cdot 0,3^4)\cdot\dfrac{1}{4}}{(0,7^8\cdot 0,3^4)\cdot\dfrac{1}{4}+(0,3^8\cdot 0,7^4)\cdot\dfrac{3}{4}}=\frac{0,2401}{0,2644}=0,9081\cdot$$

Отже ситуація не дуже вже й заплутана, бо «зелене» істотно помітне.

А навіщо повертати фішки у мішечок? Якщо не повертати фішки, то подію $GGWGWGGWGWGG$ характеризуватимуть числа

$$P(A\mid G)=\frac{70}{100}\cdot\frac{69}{99}\cdot\frac{30}{98}\cdot\frac{68}{97}\cdot\frac{29}{96}\cdot\frac{67}{95}\cdot\frac{66}{94}\cdot\frac{28}{93}\cdot\frac{65}{92}\cdot\frac{27}{91}\cdot\frac{64}{90}\cdot\frac{63}{89}=\frac{63\cdot\ldots\cdot 70\cdot 27\cdot\ldots\cdot 30}{89\cdot\ldots\cdot 100},$$

$$P(A\mid W)=\frac{30}{100}\cdot\frac{29}{99}\cdot\frac{70}{98}\cdot\frac{28}{97}\cdot\frac{69}{96}\cdot\frac{27}{95}\cdot\frac{26}{94}\cdot\frac{68}{93}\cdot\frac{25}{92}\cdot\frac{67}{91}\cdot\frac{24}{90}\cdot\frac{23}{89}=\frac{67\cdot\ldots\cdot 70\cdot 23\cdot\ldots\cdot 30}{89\cdot\ldots\cdot 100},$$

$$P(G\mid A)=\frac{\dfrac{63\cdot\ldots\cdot 70\cdot 27\cdot\ldots\cdot 30}{89\cdot\ldots\cdot 100}\cdot\dfrac{1}{2}}{\dfrac{63\cdot\ldots\cdot 70\cdot 27\cdot\ldots\cdot 30}{89\cdot\ldots\cdot 100}\cdot\dfrac{1}{2}+\dfrac{67\cdot\ldots\cdot 70\cdot 23\cdot\ldots\cdot 30}{89\cdot\ldots\cdot 100}\cdot\dfrac{1}{2}}=$$

$$=\frac{63\cdot 64\cdot 65\cdot 66}{63\cdot 64\cdot 65\cdot 66+23\cdot 24\cdot 25\cdot 26}\approx 0,979678388407846$$

Результат ще переконливіший на користь «зеленого», але більша складність обчислень, взагалі кажучи, не окуповується.

Звісно, результат одиничного досліду з витягування фішок може бути будь-яким іншим. Приміром, поспіль 30 разів може з'являтись зелена фішка:

$$P(G\mid A)=\frac{0,7^{30}\cdot\dfrac{1}{2}}{0,7^{30}\cdot\dfrac{1}{2}+0,3^{30}\cdot\dfrac{1}{2}}\approx 0,999999999991.$$

А якщо вона з'явилась 31 раз, то… Однак стільки нема в білому мішечку.

Поспіль 30 разів може з'являтись біла фішка:

$$P(G\mid A)=\frac{0,3^{30}\cdot\dfrac{1}{2}}{0,7^{30}\cdot\dfrac{1}{2}+0,3^{30}\dfrac{1}{2}}=\frac{0,3^{30}}{0,7^{30}+0,3^{30}}\approx 0,000000000009$$

Але тільки, якщо вона з'явилась 31 раз, достовірно стає відомо, що мішечок білий (так багато білого нема в зеленому мішечку).

Подія $GGWGWGGWGWGG$ є гарантовано значущою для «зеленого» висновку тільки у масовій перспективі, але не в одиничній. А загалом ситуація більше тривіальна, ніж… повчальна. Вона надто штучна.

Вдамося ще до такого прикладу (з підручника «Теорія ймовірностей» для вищих технічних закладів освіти авторства О. Вентцель).

За об'єктом, що може перебувати в одному з двох станів S_1 чи S_2 та випадково час від часу переходити з одного в інший, спостерігають два інформаційні центри C_1 і C_2. У ході тривалого стеження з'ясовано, що 30 % часу об'єкт перебуває в стані S_1. Спостережний центр C_1 схильний передати хибну інформацію про поточний стан об'єкта у 2 % випадків, а центр C_2 — у 8 % випадків. Якось центр C_1 повідомив, що об'єкт перебуває в стані S_1, а от центр C_2, — що в стані S_2. Кому повірити?

Вірити треба, мабуть, тому повідомленню, ймовірність бути істиною якого більша. Отже треба критично оцінити дві гіпотези: H_1 — об'єкт перебуває в стані S_1, ймовірність якої $P(H_1)=0,3$; H_2 — об'єкт перебуває в стані S_2, ймовірність якої $P(H_2)=0,7$. Спостережена подія A у цьому випадку полягає в тому, що: центр C_1 бачить об'єкт в стані S_1, а центр C_2 — в стані S_2.

Щоби відбулась подія A за припущення H_1, центр C_1 мав би повідомити істину, а центр C_2 схибити, а за припущення H_2 усе має бути навпаки. Тож

$$P(A\mid H_1)=0,98\cdot 0,08,$$
$$P(A\mid H_2)=0,92\cdot 0,02<P(A\mid H_1).$$

Формула Бейза в такому разі повідомить, що істинним є радше стан S_1: характеризуватимуть

$$P(H_1\mid A)=\frac{P(H_1)P(A\mid H_1)}{P(H_1)P(A\mid H_1)+P(H_2)P(A\mid H_2)}=\frac{0,3\cdot 0,98\cdot 0,08}{0,3\cdot 0,98\cdot 0,08+0,7\cdot 0,92\cdot 0,02}=0,645\gg 0,300. \quad (8)$$

Чи не дивно, сам лише факт проведення акції оцінювання стану об'єкта — подія, сенс якої залишився насправді прихованим, — ще більше посилила повагу до спостережного центру C_1 і водночас істотно змінила уявлення про об'єкт (стан S_1 став переважним)? Як би мали сприйняти цей результат ті, хто у процесі тривалого (!) стеження з'ясував, що 30 % часу об'єкт перебуває в стані S_1, а 70 % часу — у стані S_2?

Приймімо умовно, що десять спостережень — це результат (не дуже, але хай) тривалого стеження: у трьох випадках об'єкт перебував у стані S_1 і сім разів його виявили у стані S_2. І ось відбулося ще одне спостереження, і хай при цьому умовно (нібито) об'єкт виявили в стані S_1 (насправді це невідомо). Якщо раніше можна було казати, що $P(H_1) = 3/10$, то тепер потрібно визнати, що

$$P(H_1) = \frac{3+1}{10+1} = \frac{4}{11} \approx 0,36 > 0,30 .$$

Але ж це далеко не $P(H_1 \mid A) = 0,645$. А якщо стан S_1 зафіксований у 30 випадках зі ста, то матимемо:

$$P(H_1) = \frac{30+1}{100+1} = \frac{31}{101} \approx 0,307 > 0,300 .$$

Тож чим ретельніше стежили б за об'єктом, тим скромнішими були б результати на користь гіпотези H_1.

Отже зрозуміло: ретельні/рутинні дослідження не дуже підсилюють гіпотезу H_1. А от варто тільки вдатись до парадного алгоритму Бейза як ця гіпотеза відразу стає дуже-дуже вагомою. Якби схильність передавати істинну (хибну) інформацію про поточний стан об'єкта в обох центрів була однакова, то справджувалась би рівність $P(A \mid H_1) = P(A \mid H_2)$. В такому разі відповідно до загальної частини формули (8) виникло співвідношення $P(H_1 \mid A) = P(H_1)$.

Припустімо тепер, що в новому спостереженні на користь стану S_1 формально зараховано тільки частка 0,3 від ніби його появи (відповідно до ймовірності настання). То тепер

$$P(H_1) = \frac{3+0,3}{10+1} = \frac{30+0,3}{100+1} = 0,30 ,$$

що й слід було очікувати. Отже, якщо попередні дослідження справді були тривалими (ретельними), то вони надалі обов'язково будуть стійкими (не догоджатимуть принципові Бейза).

Ще один приклад (належний згадуваній О. Вентцель).

Двом стрільцям дали можливість по одному разу поцілити в одну і ту саму мішень. Вправність стрільців різна: ймовірність влучання першого з них — $P(S_1) = 0,8$, а другого — $P(S_2) = 0,4$. В мішені виявили лише одне влучання. Треба знайти ймовірність того, що це влучання належить першому стрільцю. (Що ця ймовірність мала б нам ще повідомити, коли відомо, що перший стрілець вправніший?)

Перед випробуванням можлива така система можливостей (гіпотез):

H_1 — жоден зі стрільців не поцілить;

H_2 — обидва поцілять;

H_3 — перший стрілець поцілить, а другий ні;

H_4 — другий стрілець поцілить, а перший ні.

Ймовірності правдивості цих гіпотез такі:

$$P(H_1) = (1 - P(S_1))(1 - P(S_2)) = 0,2 \cdot 0,6 = 0,12 ;$$
$$P(H_2) = P(S_1)P(S_2) = 0,8 \cdot 0,4 = 0,32 ;$$
$$P(H_3) = P(S_1)(1 - P(S_2)) = 0,8 \cdot 0,6 = 0,48 ;$$
$$P(H_4) = (1 - P(S_1))P(S_2) = 0,2 \cdot 0,4 = 0,08 .$$

Випробування істотно позначилося на умовній ймовірності події A/«Влучання у мішень» за різних гіпотез:

$$P(A \mid H_1) = P(A \mid H_2) = 0 ;$$
$$P(A \mid H_3) = P(A \mid H_4) = 1$$

(після випробування гіпотези H_1 і H_2 стають цілком хибними, а гіпотези H_3 і H_4 стають цілком вірогідними). А от відповідно до теореми Бейза ймовірності ніби вірогідних гіпотез становитимуть

$$P(H_3 \mid A) = \frac{0,48 \cdot 1}{0,48 \cdot 1 + 0,08 \cdot 1} = \frac{6}{7} > \frac{8}{10} ,$$
$$P(H_4 \mid A) = \frac{0,08 \cdot 1}{0,48 \cdot 1 + 0,08 \cdot 1} = \frac{1}{7} < \frac{4}{10} .$$

Очевидно, що отриманий висновок аж-ніяк не зорієнтований на те, що сталось насправді. Хай знову «умовно ретельними» є попередні дослідження, коли кожному стрільцю дали можливість показати свою вправність десятьма пострілами (окремо один від одного). Тоді власне було з'ясовано, що перший стрілець поцілив вісім разів, а другий — чотири. Якщо б у новому одинадцятому випробуванні поцілив тільки перший стрілець, то мала б відбутись така переоцінка ймовірностей:

$$P(S_1) = \frac{9}{11} = 0,818 , \; P(S_2) = \frac{4}{11} = 0,364$$

(при цьому $P(H_3) = 0,52$ і $P(H_4) = 0,07$).

Але ж могло статись все навпаки і в такому разі:

$$P(S_1) = \frac{8}{11} = 0,727 , \; P(S_2) = \frac{5}{11} \approx 0,455$$

(при цьому $P(H_3) = 0,40$ і $P(H_4) = 0,12$).

Але чим ретельнішими (численнішими) були б попередні дослідження, тим меншою була б різниця між наведеними даними.

Отже можна розрізнити два штиби суджень: скромні від тих, хто рутинно добуває дані, і гучні від тих, хто ніби відразу набув інформацію і має обов'язок ухвалювати рішення. Цікаво, що в останній задачі ніби й зрозуміло, що скоріш за все поцілив перший стрілець, але дуже хочеться будь-що виміряти ймовірність цього факту.

Звісно, бувають приклади, коли «Бейз — на коні».

Наприклад, існує віддавна дуже гарний приклад того, що ймовірність події може зазнати суттєвої переоцінки після проведення досліду. Цей приклад трактують як парадокс.

Припустімо, Ви — учасник шоу, якому пропонують вгадати за якими з трьох дверей знаходиться автомобіль. Він має стати Вашим, якщо вгадаєте. За двома іншими дверима — порожньо (існує варіант шоу, коли за іншими дверима розміщують якісь дрібніші призи, але тут це не принципово). Ви вибираєте, скажімо, двері *3*, і ведучий, який знає, щó там за кожними дверима, відчиняє для Вас якісь інші, скажімо, двері *2*, за якими порожньо (чи якийсь дрібний приз), мал. 1. Далі він Вас запитує: «Чи не хочете змінити свій вибір (тобто вибрати двері *1*)?». То чи у Ваших інтересах відмовитись від свого попереднього вибору?

Рисунок 1 — Зміна ситуації невизначеності

З першого погляду здається, що інформація, отримана при відчиненні дверей *3*, несуттєва. Тож здається, що незалежно від Ваших наступних дій, ймовірність виграти машину залишається на рівні 1/3? На початку 1990-х років ця задача привернула загальну увагу, і навіть багато професорів математики з університетів Ivy League (Ivy League — це вісім провідних університетів у східній частині США) вважали, що Ви не можете вплинути на ймовірність виграшу, змінивши вибір дверей.

Але насправді, як не дивно, можете! У разі зміни вибору ймовірність стає рівною 2/3. Нескладно збагнути, що якщо ви змінюєте свій вибір, то програєте тільки у випадку, коли машина знаходилась саме за дверима *1*, тому ймовірність програшу та сама, тобто 1/3, як і ймовірність виграшу, коли Ви не змінюєте свій вибір. Та не слід забувати, що тут вигулькнула ще й невипадкова (але не дуже) інформація від ведучого.

Чи «працює» тут теорема Бейза?

Позначмо через A_1, A_2, A_3 гіпотези, що автомобіль знаходиться відповідно чи за першими, чи за другими, чи за третіми дверима, а от через D_1, D_2, D_3 — ті відповідно перші, другі, треті двері, які ніби міг би відчинити ведучий після того, як став відомим вибір учасника шоу (за цими дверима автомобіля, звісно, нема).

Тож Ви вибрали треті двері, а ведучий вимушений був відчинити другі двері. Тому, щоб з'ясувати чи варто змінювати вибір, необхідно визначити величину $P(A_1 \mid D_2)$ — ймовірність того, що автомобіль знаходиться за першими дверима (що їх Ви не вибирали) за умови, що відчинені саме другі двері (де автомобіля точно нема).

Формула Бейза (5) в цьому випадку має вигляд

$$P(A_1 \mid D_2) = \frac{P(D_2 \mid A_1)\, P(A_1)}{P(D_2)} = \frac{P(D_2 \mid A_1)\, P(A_1)}{\sum_{i=1}^{3} P(D_2 \mid A_i)\, P(A_i)}.$$

Зрозуміло, що

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3},$$

оскільки Ви не знаєте, де автомобіль. Натомість

$$P(D_2 \mid A_1) = 1,$$

оскільки ведучий знає, що автомобіль — за першими дверима, а Ви вибрали треті двері, то в нього нема варіантів окрім, як відчинити другі двері. При цьому $P(D_2 \mid A_3) = \frac{1}{2}$ (якщо б автомобіль знаходився за третіми дверима, то можна було б відчинити або перші, або другі). До того ж, у виразі

$$P(D_2) = \sum_{i=1}^{3} P(D_2 \mid A_i)\, P(A_i) = P(D_2 \mid A_1)\, P(A_1) + P(D_2 \mid A_2)\, P(A_2) + P(D_2 \mid A_3)\, P(A_3)$$

другий доданок дорівнює нулю ($P(D_2 \mid A_2)\, P(A_2) = 0$), оскільки у разі припущення (гіпотези), що за другими дверима — автомобіль, відчиняти ці ж самі двері ведучому не випадало б (була б зруйнована інтрига).

$$P(D_2) = P(D_2 \mid A_1)\, P(A_1) + P(D_2 \mid A_3)\, P(A_3) = 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2},$$

$$P(A_1 \mid D_2) = \frac{1 \cdot \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Це доводить, що змінювати вибір дверей таки вигідно (ймовірність надибати автомобіль зростає вдвічі). Але це вірно статистично: цим треба було б керуватись, коли б довелося багаторазово брати участь у цьому шоу. Висновок підтверджено імітаційним моделюванням на комп'ютері та натурним експериментом.

Після того, як ведучий відчинив певні двері, учаснику шоу може здатися, що ймовірність знаходження автомобіля за будь-якими з двох закритих дверей становить $\frac{1}{2} > \frac{1}{3}$. Це справді було б так, коли б ведучий відчинив саме ті двері, які вже обрали Ви (тож вибір все-дно довелося б змінити), або ж відчинив би які-небудь двері перед тим, як Ви зробили свій вибір. Ситуація змінилася б, якщо б, приміром, довелося вибирати нові двері за допомогою підкидання досконалої монети. Інколи таки доводиться підкидати правильну монету (чи якось по-іншому рандомізувати ситуацію), аби ухвалити рішення.

Загалом ситуація здається контрінтуїтивною. Хоча, можна відстежити етапи зміни цієї ситуації невизначеності за допомогою мал. 1: перший етап — повна невизначеність, а тому $P(A_1) = P(A_2) = P(A_3) = 1/3$; другий етап — зроблено припущення, що автомобіль за дверима *3*, а отже імовірність того, що автомобіль — за іншими обома дверима (ніби одночасно і першими, і другими) дорівнює $P(A_1A_2) = 2/3 = 2P(A_3)$; третій етап — відчинено другі двері, за якими порожньо, а отже можна зробити висновок: $P(A_1) = P(A_1A_2) = 2/3 = 2P(A_3)$.

Уявімо тепер, що дверей є сто і тільки за одними — автомобіль. Учасник шоу обрав якісь одні, а ведучий відчинив 98 інших. Тут зрозуміло, що робити. А якщо ведучий відчинив тільки одні двері? Чи дуже конструктивною може бути «правильна» поведінка?

Ось ще одна задача про сенс зміни вибору (парадокс двох конвертів). Є два конверти, мал. 2, що містять боргові зобов'язання на деякі суми. Один конверт вручили *A*, інший — *B*. Їм сказали, що величина боргового зобов'язання в одному із конвертів вдвічі більша, але не сказали, в якому саме. З ймовірністю 1/2 у *A* знаходиться менш цінний конверт. Чи є сенс для *A* обмінятися конвертами? Ця задача має відомі аналоги: про сенс обміну краватками, з яких одна дорожча (про що наперед невідомо), чи гаманцями з неоднаковою (прихованою) кількістю грошей.

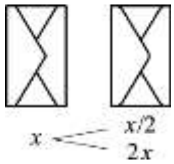

Рисунок 2 — Зміна ситуації невизначеності

Інтуїція підказує, що обмін не буде засобом гарантовано отримати зиск. Тут інтуїція справді не підведе. Але, як кажуть, з протилежного боку трава зеленіша…

В окресленій ситуації *A* може міркувати собі так. Хай в моєму конверті зобов'язання на суму x. В конверті, який у *B*, боргове зобов'язання вдвічі менше від того, що в моєму конверті, з ймовірністю 1/2 і вдвічі більше з ймовірністю 1/2, мал. 2. Отже, в середньому матимемо

$$y = \frac{1}{2}\left(\frac{1}{2}x\right) + \frac{1}{2}(2x) = \frac{5}{4}x > x,$$

тому мені вигідно обмінятися конвертами. З тих самих причин *B* теж хоче обмінятися конвертами. Обидва ніби виграють. Але ж та сама підстава обмінятися конвертами існуватиме далі. Багатократний обмін міг би вести до зростання середнього аж до нескінченності, мал. 2. Щось тут не так!

Тут, мабуть, наплутано з апріорністю і апостеріорністю (*a priori* і *a posteriori* — принципово різні речі). Припустімо, що x — менша сума зобов'язання, і вона в конверті, що у A. Тому в B може бути тільки конверт із сумою $2x$ (іншого просто бути не може). Ймовірність цієї ситуації — 1/2, як і ситуації, коли в A — конверт з сумою $2x$, а тому в B — конверт з сумою x. Якщо A вирішить обміняти конверт, то в першій ситуації з імовірністю 1/2 він матиме зиск $+x$, але в другій з імовірністю 1/2 — збитки $-x$. Отже в середньому

$$y = \frac{1}{2}(+x) + \frac{1}{2}(-x) = 0 .$$

Отож інтуїція підказує цілком правильно. А роздумування, виявляється, шкодить. Та переважна більшість ймовірнісних висновків є-таки контрінтуїтивними. Відразу спадає на думку відомий приклад дивних подій, що відбуваються майже напевно.

Хай одна подія відбувається з ймовірністю $p = 0,9900$, а інша — з імовірністю $p = 0,9999$. Це майже однакової ваги події, що відбуваються майже напевно. Але чи так воно у всіх сенсах?

Припустімо, що йдеться про незалежні події, які можуть відбутись у будь-який день року з імовірністю $p = 0,9900$. Тоді ймовірність того, що вони відбуватимуться щодня впродовж року (365 днів поспіль) буде меншою за $P \approx 0,03$ — це майже неймовірна, неможлива подія. Але якщо $p = 0,9999$, то $P \approx 0,97$ — це майже певна подія.

В прикладі про шоу ситуація була за крок до усунення проблемності. А тому формула Бейза звела її до статистично визначального алгоритму (розв'язком є саме алгоритм, а не параметри): якби була можливість брати участь багато разів у такого штибу шоу, то слід було б вдаватись до єдиного конструктивного алгоритму гри, який обов'язково вестиме до успіху. Але в цю гру відразу перестануть грати, оскільки організатори шоу не мають інтересу завжди програвати. І взагалі, ігри, у яких хтось приречений програвати, не повинні культивуватись. Навіть Природа, як тепер розуміє Людство, не повинна бути завжди в програші. Якось має нарешті усталитися постійний розвиток (про сталий розвиток не можна тільки говорити). Але це не заклик сповідувати суто азарт.

Натомість в інших задачах конкурентними є алгоритм нагромадження даних, у якому перед ведуть принципи об'єктивності, та алгоритм Бейза, у якому править радше суб'єктивізм. А насправді усім править довгострокова ймовірність, на яку власне спирається алгоритм нагромадження даних.

Хай маємо ситуацію з п'ятьма цілком однаковими глеками (урнами) *1*, *2*, …, *5*, у яких містяться по 10 однакового розміру кульок білого (W — *Wight*) чи/та чорного (B — *Black*) кольору. У таблиці вказана кількість білих кульок у кожному глеку, решта — чорні.

Параметри ситуації	**Глек**				
	1	***2***	***3***	***4***	***5***
Кількість білих кульок	0	2	5	8	10
Імовірність p_{Wi} витягти білу кульку	0	0,2	0,5	0,8	1
Апріорна ймовірність p_i вибору i -го глека	0,2	0,2	0,2	0,2	0,2
Апостеріорна ймовірність P_i вибору i -го глека, якщо у ньому виявили білу кульку	0	0,08	0,2	0,32	0,4
Апостеріорна ймовірність P_i' вибору i -го глека, якщо у ньому виявили білу і чорну кульки	0	0,28	0,44	0,28	0
Нова апостеріорна ймовірність P_i'' вибору i -го глека, якщо у ньому виявили білу і чорну кульки	0	0,11	0,44	0,45	0

Вміст кожного глека відомий, але ззовні не позначений. Треба наугад вибрати глек і спробувати з'ясувати, скільки в ньому могло б бути білих і скільки чорних кульок. Апріорна ймовірність надибати глек з тим чи іншим складом кульок очевидна: $p_i = 1/5 = 0,2$, $i = 1, 2, ..., 5$. Зорієнтуватись, з яким вмістом глек насправді вибрано, неможливо, і було б добре порушити невизначеність експериментом.

Хай з обраного глека витягли кульку і вона виявилась білою. Тож достовірно з'ясовано, що це точно не глек *1*, а один з глеків *2*, *3*, *4*, *5*. А от поміряти ці «точні» дані ймовірністю можна, вдаючись до формули Бейза (див. таблицю)

$$P_i = \frac{p_i p_{Wi}}{p_1 p_{W1} + p_2 p_{W2} + p_3 p_{W3} + p_4 p_{W4} + p_5 p_{W5}} = \frac{p_i p_{Wi}}{1/2} = 2 p_i p_{Wi},$$

де $p_{Wi} = P(W \mid i)$ — умовна ймовірність того, що біла кулька належить глеку з *i*-м складом кульок (подібно, $p_{Bi} = P(B \mid i) = 1 - P(W \mid i)$ позначатиме умовну ймовірність того, що чорна кулька належить *i*-му глеку).

Сукупно кількість білих і чорних кульок в глеках однакова. Тож імовірність натрапити чи на білу, чи на чорну кульку що умовно, що ніби реально у два етапи (з початковим вибором глека) однакова — дорівнює ½. Припустімо, що в ході двох дослідів поспіль з поверненням кульок у глек випала біла і чорна кульки (послідовність їх появи значення не має). У такому разі формула Бейза дасть нові оцінки (див. таблицю):

$$P_i' = \frac{p_i p_{Wi} p_{Bi}}{p_1 p_{W1} p_{B1} + p_2 p_{W2} p_{Bi2} + p_3 p_{W3} p_{Bi3} + p_4 p_{W4} p_{Bi4} + p_5 p_{W5} p_{Bi5}} =$$

$$= \frac{p_{Wi} p_{Bi}}{p_{W1} p_{B1} + p_{W2} p_{Bi2} + p_{W3} p_{Bi3} + p_{W4} p_{Bi4} + p_{W5} p_{Bi5}} = \frac{p_{Wi} p_{Bi}}{0{,}57}$$

Але якщо замінити апріори на апостеріори, то виникнуть нові апостеріори (див. таблицю):

$$P_i'' = \frac{P_i' p_{Wi} p_{Bi}}{P_1' p_{W1} p_{B1} + P_2' p_{W2} p_{Bi2} + P_3' p_{W3} p_{Bi3} + P_4' p_{W4} p_{Bi4} + P_5' p_{W5} p_{Bi5}} = \frac{P_i' p_{Wi} p_{Bi}}{0{,}114}.$$

Звісно, кожен наступний експеримент на засадах бейзівської теорії якось наближатиме висновок до істини. Але те саме зробить звичне рутинне дослідження, суть якого висловлює теза «треба брутально перевірити вміст вибраного глека». Щоби переконатись, що вибрано, приміром, глек з першим чи п’ятим набором кульок достатньо провести дев’ять виймань кульок без їх повернення. Власне подібним і займається наука, в рамках якої є цілковите розуміння того, що пошук істини неперервний і нескінченний процес, у якому реальне задоволення від пошуку важить більше за намагання привласнити недосяжну істину.

Доречно наголосити, що в цьому прикладі на кожному проміжному кроці бейзівського оцінювання ситуації невизначеності співіснуватимуть у тій чи іншій пропорції достовірність та ймовірність. Приміром, після того як витягли білу і чорну кульки, стало достеменно відомо: 1) глек не містить набори кульок *1* і *5*, але може йтися про один з наборів *2*, *3*, *4* — достовірні дані; 2) глек може містити набір *2* з імовірністю 0,11, *3* з імовірністю0,44, *4* з імовірністю 0,45 — імовірнісні дані.

Наведені приклади (як і багато інших) мали б слугувати за спрощені моделі реальних ситуацій невизначеності, але виглядають вони занадто штучними, неправдивими.

Обмірковування. Бейзівський принцип отримування висновків у науці мусить спиратись на попередні ідеї, уявлення, переконання. А як інакше? Та водночас не оминути занепокоєння, що послідовне багатократне (в якомусь сенсі успішне, результативне) застосування формули Бейза (коли після кожного досліду апостеріорні ймовірності перераховують і використовують далі як апріорні) цілком нівелює первинні апріорні ймовірності — вони ніби цілком не потрібні. То можна нічого не знати, бо й так навчить принцип Бейза?

Звісно, є підстави вважати, що за певних умов багатократне застосування теореми Бейза дасть послідовність апостеріорних розподілів, що збігається до істинного розподілу, а отже апріорний розподіл в асимптотичному вимірі ніби ролі не відіграє. Але ж асимптотичність рутинного накопичування інформації (корисних даних, з яких зародиться інформація) — це саме те, чим наука невпинно займається, намагаючись пізнати сутність через спостереження явища. Бейзівські ж міркування — це маніпулювання ймовірностями за ніби законними правилами з метою посилити свої переконання, не переобтяжуючи себе рутинними дослідженнями. І невідомо, з якої миті слід сподіватись на об’єктивність набутого знання.

Застосовувати принципи бейзіанства можуть хіба що тверді прихильники індуктивного бачення процесу пізнання. Натомість К. Поппер, уникаючи Аристотелевої теорії пізнання здорового глузду, простими засобами доводив, що викладене колись людською мовою знання, більше не є частиною нас самих, а стає окремою сутністю, яка розвивається завдяки критичному відбору [30]. Він стверджував, що наука вимушена покладатися на методологію, засновану на фальсифікованості, оскільки жодна кількість експериментів не зможе гарантувати правильність висунутої теорії, а от спростувати її зможе навіть одне спостереження [31]. В літературі це твердження поширилося у формі сентенції про чорного лебедя: скільки б не довелося зустрічати білих лебедів, це не утвердить нашого переконання, що всі без винятку лебеді білі, але варто зустріти одного чорного… Універсальні твердження не висновуються з конкретних висловлень, але цілком можуть суперечити окремому висловленню — асиметрія суджень.

Розумні люди, звісно, розрізняють середовище симетрії і середовище асиметрії. Симетричність там, де очевидною є, приміром, «бездоганність» монети, якою бавляться підкиданням її. Схоже, надмірно популярний нормальний закон розподілу значень випадкової величини також належить до симетричного середовища. Симетрія провокує плутанину у сприйнятті ймовірності і сподівання (очікування). Але якщо у якійсь нечіткій справі ймовірність малого зиску — 99/100, а великого збитку — 1/100, то очікувані шанси на успіх у цій справі істотно залежать від співвідношення великого збитку і малого зиску (не так від великої ймовірності зиску, як від відношення добутків відповідних ймовірностей на зиск та збиток, від асиметрії шансів). Рідкісні події насправді дуже часто бувають навіть дуже вагомими. То чи здатен принцип Бейза шукати дуже вагомі, але рідкісні уточнення наявної інформації? Тут могла б зарадити дедукція.

Виникає також питання, а чи може суб'єктивна інформація, що міститься, приміром, в апріорних уявленнях, взагалі бути основою для об'єктивних ухвалень? Бруно де Фінетті твердив, що ймовірність не існує об'єктивно, вона є лише спробою матеріалізувати наші (суб'єктивні, звісно) уявлення. Індивідуальна чи суб'єктивна ймовірність вимірює нашу впевненість, що та чи інша (очікувана) подія настане. Тож можна об'єктивно говорити про суб'єктивну ймовірність, навіть якщо випадковість не об'єктивна.

А як ставитися до вражаючої думки, що свободи волі і, як наслідок, відповідальності взагалі не існує. Український філософ М. Попович [32] зауважував: «… вибір рішення людиною — це ілюзія, бо він є результатом якихось причин, що їх ще ми не знаємо і не знає сама ця людина, і ніякої вільної активності, ніякого перетворення світу людиною насправді не існує». А може пристати на бік Філострата: «Боги знають майбутнє, люди — те, що відбувається, мудреці — те, що наближається»? Тож навіщо чимось науці перейматись?! Посилання на повчальні приклади й цитування авторитетів, які якось колись такими стали, у разі висновування принципово загального — це практика наївної емпіричної індукції, примітивного суб'єктивізму. Узвичаєна ж мудрість звертає увагу більше на те, що можна пояснити негайно, та ще й красиво та «двома словами».

Звісно, кращими за старі (учорашні) дані є нові (сьогоднішні)!? Вибудовуючи тренди, ми послуговуємося старими даними, на основі яких формуємо свою інформацію. Старе ніби стає основою майбутнього, закарбовуючи суб'єктивність минулого.

Суб'єктивною ймовірністю оперує згадуваний уже Л. Севідж, поєднуючи її з теорією корисності, розробленою Нейманом і Моргенштерном. А Говард Райфа впевнено твердив, що визнавати переваги наслідків схвалюваних рішень треба керуючись певною шкалою корисності, а оцінювати невизначеність — у термінах певної шкали ймовірностей. Він стояв на позиціях ідеології бейзівців (такої собі меншини суб'єктивістів, але не самого Бейза), закликаючи оперувати поняттям корисності і поняттям суб'єктивної ймовірності. А от більшість учених переконані, що і об'єктивна випадковість, і об'єктивна ймовірність існують реально, вони закладені у підвалини пізнаваного нами світу. Макс Борн, приміром, безапеляційно ввів об'єктивну ймовірність у квантову фізику.

Яким довгим мав би бути шлях до об'єктивного знання через систему апріорів-апостеріорів? Чи потрібно і чи вигідно користати з кожного апостеріору? Чи кожен крок від апостеріору до нового апостеріору несе тільки зиск? Чи треба вичікувати якогось особливо «хорошого» апріору, нехтуючи усіма проміжними? Насправді відповідати на ці запитання не дуже й цікаво, бо вони визначають шлях суто індуктивного пізнання. Не могла цивілізація займатися лише збиральництвом та мисливством, треба було щось культивувати. Не можна тільки «індукувати», треба обов'язково «дедукувати».

Тож між прихильниками різних підходів до формування засад теорії статистичних рішень і статистичного висновування розгорілася незгасна суперечка і пролягла неподоланна прірва цілком об'єктивно. Скоріш за все вона корисна, але ні до чого конструктивного не призведе.

Будь-яке спостереження деформує будь-яку ситуацію. Ситуація стає в тій чи іншій мірі спотвореною (викривленою), відмінною від тієї, що була перед спостереженням, перед тим, як суб'єкт зацікавився нею. Це щось подібне на квантову невизначеність у фізиці (згадаймо принцип Гейзенберга). Тож на усьому, до чого проявить інтерес людина, обов'язково позначиться її суб'єктивність. Кожна ситуація, таким чином, природно стає і об'єктивною, і суб'єктивною. Людина перейматиметься навіть тим, «несподівана це чи випадкова» подія, і загострить на цьому свою увагу.

Коли (після випробування Бейза) стверджують, що в певній ситуації з ймовірністю P правдивою є гіпотеза H, то можуть мати на увазі різне. Твердження можна розуміти так: існує дуже багато (було б добре) однакових ситуацій і саме в $100P$ % з них гіпотеза H вірогідно справдиться, якщо усі ситуації для тестування (верифікації) вибрані цілком випадково. В такому разі теорема Бейза — простий факт, схожий на закон великих чисел і вимогу рафінування (калібрування) умов усіх досліджень. Але частіше мають на увазі ступінь своєї власної віри в те, що са́ме оця конкретна ситуація на $100P$ % «налаштована справджувати» гіпотезу H.

Імовірності у разі реалізації Бейзового принципу отримування висновків обчислюються на засадах формалізації здогадки про те, що майбутнє, радше, буде подібним до минулого. Але саме в такому разі кількість випробувань Бейза для висновування апостеріорів не повинна істотно поступатись кількості подібних досліджень, на основі яких побудовані так би мовити опорні апріори. І тільки тоді технологія Бейза мала б шанс зробити внесок в розвиток науки. А от на апостеріор унаслідок кожного окремого досліду покладатися ризиковано. Тут доречно провести таку паралель.

Азартні гравці переконані, що якщо правильна (досконала) монета дуже часто випадала тризубом догори, то необхідно зростає ймовірність того, що в подальшому ця монета частіше випадатиме догори номіналом. І від розуміння цього, їм здається, можна отримати якийсь зиск. Невідповідність результатів підкидань законові великих чисел (законові Бернуллі), звісно, відразу мала б стати підставою забракувати монету. Але хіба можна знехтувати той факт, що монета радше досконала, але не так, щоби мати пам'ять і здатність обліковувати події, які трапляються з нею? До слова, пам'ять — це необхідна умова розуму.

Неправильно було б вважати, що невизначеність — це щось, вивчення чого передбачає обов'язково класичний ймовірнісний підхід. Приміром, невизначеність може існувати й у формі нечітких множин, які, взагалі кажучи, незвідні до ймовірнісних мір (Л. Заде [33]). А хіба заборонено припускати, що в ситуації невизначеності діє такий закон (принцип) з такою ймовірністю, а інший з іншою ймовірністю — ця ситуація, мабуть, «задуже ймовірнісна»?

З-посеред найзнаменитіших робіт Л. Заде є підстави в контексті цієї статті виділити *«Основи нового підходу до аналізу складних систем і процесів ухвалення рішень»* [34]. Описаний в цій роботі підхід має три принципові особливості: 1) використання замість чи на додаток до звичних числових змінних таких собі «лінгвістичних» змінних у формі речень на природній чи штучній мові; 2) відображення зв'язків між змінними нечіткими умовними операторами; 3) опис складних взаємин за допомогою нечітких алгоритмів. У статті *«Нечіткі множини як основа теорії можливості»* [35], з якої починався перший номер міжнародного журналу «Нечіткі множини й системи», Л. Заде запропонував підхід до обчислення невизначеностей, що спирається на неадитивну міру можливості, і зокрема, інтерпретацію нечіткої множини як функції розподілу повноважень (теорія можливостей). На відміну від нечіткої множини, що виражає неточність оцінки деякого атрибута, міра можливості описує невизначеність, неповноту інформації, пов'язану з появою тієї чи іншої чіткої події. Сьогодні міра можливості та двоїста до неї міра необхідності правлять за основні засоби моделювання невизначеності в інтелектуальних системах.

Тож теорема гіпотез-причин не може визнаватись універсальним засобом усунення невизначеності чи хай навіть зменшення її рівня. Не можна захоплюватися лише А. Зельнером чи В. Т. Моррісом [28, 36], треба пам'ятати й про К. Поппера [30, 31].

Резюме. Отже йшлося про таке.

Якщо H_1, $H_2,\ldots$ — події, що попарно не перетинаються та мають додатні ймовірності, і одна з яких відбувається завжди (чи принаймні з імовірністю 1), то

$$P(H_k \mid A) = \frac{P(H_k)\,P(A \mid H_k)}{P(H_1)P(A \mid H_1) + P(H_2)P(A \mid H_2) + \ldots}.$$

Це формула Бейза, яка ніби розкриває алгоритм за апріорними ймовірностями $P(H_k)$ (імовірностями подій H_k перед тим, як подія A відбулася) знаходити апостеріорні ймовірності (ймовірності тих самих подій вже після того, як подія A відбулася). Спадає на думку розглядати події H_k як причини, а відтак Формула Бейза мала б розкривати зміст такої собі теореми про ймовірності причин.

Але дуже часто апріорні ймовірності подій H_k невідомі. А тому, якщо вже й користуються формулою Бейза, то за відсутності інформації про причини вважають зазвичай, що усі ймовірності $P(H_k)$ однакові. Але багатьом не безпідставно це здається цілком неприйнятним. І справді, коли таке стається це означає, що задача/проблема погано обумовлена, погано вивчена, погано підготовлена, аби нею перейматись.

Є люди, які ніколи не приймуть ідеї випадковості, а є ті, яких ця ідея завжди мучитиме через непереборне бажання запізнати структуру безструктурності, в існуванні якої вони переконані. Для когось важливим є такий хід висновування: 1) висувається гіпотеза H; з H випливає, що F не може мати місце; виявилось, що F — таки факт; отже H — хибна (це доведення до абсурду — від супротивного, а часом від неймовірного). І завжди будуть ті, хто покладатимуться на теорему Бейза. Як і ті, хто впевнені: судова система не зобов'язана довести, що хтось злочинець (або не-злочинець), але зобов'язана ретельно дбати, щоби судочинство відбулось строго за правилами (це іронія). І звісно, багато важитиме дисципліноване ставлення до ризику пізнати не-те та інтроспекція (занурення в себе).

Формула Томаса Бейза (виразник теореми про гіпотези чи причини) — формальний наслідок поєднання теореми множення ймовірностей і формули повної ймовірності. Вона занадто проста, аби стати інструментом глибокого пізнання світу. Інакше кажучи, сама по собі теорема Бейза незаперечна в математичному сенсі, але у практичних застосуваннях виникають клопоти із внесенням змісту і сенсу в конкретні ситуації. З нею, як бачимо, виникає багато питань, на які годі знайти відповідь. Нічого дивного: щирий математик зосереджений на тому, що у нього в голові, а щирий (і не тільки) фізик (і не тільки) — на тому, що його оточує. Так інколи закріплюють за кимось абстрактне, а за іншими реальне.

А якщо у когось з'явилися справді віри гідні апріорні ймовірності, то це означає, що колись, цілком напевне, провадилися ґрунтовні дослідження, на тлі яких бейзівський висновок — це виконане нашвидкуруч часто бліде й не дуже переконливе продовження цих досліджень. Тому не дивно, що був період, коли теорема Бейза майже зникла із засобів статистичних досліджень. І тільки в середині XX століття підхід Бейза знову виринув серед інструментарію для пошуку допустимих і мінімаксних оцінок в типових задачах оцінювання невідомого параметра ймовірнісного розподілу. Бейзіанство тепер набуло також популярності, приміром, у процесах оживлення штучного інтелекту та формування у нього здатності навчатися на основі великих обсягів даних, що засновані на людському досвіді.

Але те, що штучний (погано прогнозований і людиноконкурентний) інтелект облюбував собі саме технологію бейзіанства, нічого особливого не означає. Він просто «любить» завантажувати свою пам'ять, несумірну з людською. Його раніше чи пізніше природний інтелект наверне на інший, розумніший, шлях — знайти б найкращий.

Неприйняття суб'єктивізму в науці помітне зовсім не через те, що хочуть, аби в результатах пізнання будь-що не відображався суб'єкт, що пізнає. Для того, аби залучити до процесу створення теорії дедукцію (а без неї насправді не виникне теорія), треба вийти за межі проблеми і подивитися на неї здалека, а це може зробити розумом й інструментально, звісно, тільки суб'єкт (хто ж іще?). Тільки через дедукцію ми знаємо, приміром, про час і енергію, яких в природі насправді не існує (де тут об'єктивність?). Дедукція дарує вірогідні знання, але не здатна в своїх аксіомах беззастережно чітко окреслити надійні передумови (але, як без них і отже як без індукції?). Ймовірні знання дарує індукція, яка саме з участю суб'єкта всередині проблеми прокладає шлях від емпіричного до теоретичного, але не здатна його завершити (потрібна дедукція!). А от цінним для суб'єкта є те, що він побачить на межі, де в тій чи іншій ситуації пізнання стикаються окреслене ним дедуктивне й індуктивне, загальне й конкретне. Ця межа завжди буде рухомою саме завдяки суб'єктові, і не потрібно й неможливо, щоб вона раптом і ненароком нерухомо зависла.

Попри це загрозливим є суб'єктивізм, обумовлений/зумовлений тим, що хтось чомусь нав'язує дивитись на світ саме через такі, а не інші «криві окуляри». На прикладах і на основі результатів нескладного аналізу можна було бачити хіба що примарні перспективи ефективного застосування технології бейзіанства. Тож, приміром, у рамках дисциплін «Дослідження операцій», «Теорія ймовірності» (за Б. Паскалем — «Математика випадку»), «Теорія транспортних процесів і транспортних систем», «Теорія ухвалення рішень» тощо-тощо — усюди, де доводиться вимушено стикатись з невизначеністю, або цілеспрямовано трактувати пізнаване як невизначене — нема потреби просувати бейзіанство як щось обов'язкове до вивчення, розуміння, застосування. Тож можна заощадити інтелектуальні ресурси та фонд робочого часу будь-кого схильного вивчати і пізнавати. Тому власне доцільно застерегти собі право не дуже заглиблюватись у сенс апостеріорної ймовірності та бейзівського висновку. Теорема Т. Бейза настільки важлива, наскільки це бачить класична «Теорія ймовірності», а не деякі спеціальні монографії (де вона перебільшено потужна).

Список літератури:

1. Renyi A. On a new axiomatic theory of probability // Acta Mathematica Academiae Scientiarum Hungarica, 1955, *6*(3), 285–335. https://doi.org/10.1007/BF02024393

2. Renyi A. Foundations of probability. — San-Francisco: Holden-Day. Inc., 1970. — xvi+366 p. ISBN: 0816271143, 9780816271146

3. Savage L. J. The foundations of statistics. — John Wiley & Sons, Inc., 1954, 2nd Revised ed. — Dover Publications, 1972. — 294 p. ISBN: 978-0486623498

4. Bayes Mr, Price Mr. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr Bayes, F. R. S. Communicated by Mr Price, in a Letter to John Canton, M. A. and F. R. S. // Philosophical Transactions of the Royal Society of London. 1763, 53, 370—418. Downloaded from https://royalsocietypublishing.org/ on 12 May 2023

5. Hooper M. Richard Price, Bayes' theorem, and God. // Significance, 2013, 10 (1), 36–39. doi:10.1111/j.1740-9713.2013.00638.x. (англ.)

6. McGrayne S. B. The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy. —

Yale University Press, 2011. — 288 p. https://www.jstor.org/stable/j.ctt1np76s
ISBN-13 9780300169690

7. Stone J. V. Bayes' Rule: A Tutorial Introduction to Bayesian Analysis (5.5.2013 ed.). — Sebtel Press, 2013. — 174 p.
DOI: 10.13140/2.1.1371.6801

8. Васильків І. М. Основи теорії ймовірностей і математичної статистики: навчальний посібник. — Львів: ЛНУ імені Івана Франка, 2020. — 184 с. ISBN 978–617–10–0354–5

9. Теорія ймовірностей [Електронний ресурс]: навч. посіб. / Уклад.: О. В. Барабаш, А. П. Мусієнко, О. В. Свинчук. — Київ: КПІ ім. Ігоря Сікорського, 2021. — 193 с.

10. Stylianides N., Kontou E. Bayes Theorem and its Recent Applications // Mathematics Research Journal, 2020, MA3517, 1–6.

11. McCullagh P. What is a statistical model? // The Annals of Statistics, 2002, 30(5), 1225–1310. https://doi.org/10.1214/aos/1035844977

12. Taraldsen G., Lindqvist B. H. Conditional probability and improper priors. Communications in Statistics — Theory and Methods, 2016, 45(17), 5007–5016.
doi:10.1080/03610926.2014.935430

13. Puga J. L., Krzywinski M., Altman N. Bayes' theorem // Nature Methods, 2015, 277 – 278. doi:10.1038/nmeth.3335

14. Berrar D. Bayes' Theorem and Naive Bayes Classifier // Tokyo, Japan: Data Science Laboratory, Tokyo Institute of Technology, 2018, 1–18. doi:
DOI: 10.1016/B978-0-12-809633-8.20473-1

15. Thompson D. R., Martin C. R. Bayes' theorem and its application to cardiovascular nursing // European Journal of Cardiovascular Nursing, 2017, 16(8), 659–661.
doi:https://doi.org/10.1177/1474515117712317

16. Lohrke F. T., Carson Ch. M., Lockamy A. Bayesian analysis in entrepreneurship decision-making research: Review and future directions // Management Decision, 2018, 56(2), 972–986. DOI:10.1108/MD-12-2016-0948

17. Taraldsen G., Tufto J., Lindqvist B. H. Improper priors and improper posteriors // Scandinavian Journal of Statistics, 2022, 49(3), 969–991. https:// doi.org/10.1111/sjos.12550

18. Berger J. O., de Oliveira V., Sanso B. Objective Bayesian analysis of spatially correlated data // Journal of the American Statistical Association, 2001, 96(456), 1361–1374.
https://www. jstor .org /stable/3085905

19. Bioche C., Druilhet P. Approximation of improper priors // Bernoulli, 2016, 22(3), 1709–1728. DOI: 10.3150/15-BEJ708

20. Bord S., Bioche C., Druilhet P. A cautionary note on Bayesian estimation of population size by removal sampling with diffuse priors // Biometrical Journal, 2018, 60(3), 450–462.
DOI: 10.1002/bimj.201700060

21. Hannig J., Iyer H., Lai R. C. S., Lee T. C. M. Generalized fiducial inference: A review and new results // Journal of the American Statistical Association, 2016, 111(515), 1346–1361.
https://doi.org/10.1080/01621459.2016.1165102

22. Taraldsen G., Lindqvist B. H. Fiducial theory and optimal inference // Annals of Statistics, 2013, 41(1), 323–341. DOI: 10.1214/13-AOS1083

23. Hartigan J. Bayes theory. — Springer, 1983, 2011. — 158 p. ISBN: 978-1461382447

24. Lindley D. V. Understanding uncertainty (Revised ed.). — John Wiley & Sons, Inc, 2014. — 432 p. ISBN: 978-1-118-65012-7

25. Lehmann E. L., Romano J. P. Testing statistical hypotheses. (3-rd ed). — Springer Science+Business Media, LLC, 2005. — xiv+785 p. ISBN 0-387-98864-5

26. Finetti B. D. Probability, induction and statistics: The art of guessing. — Wiley, 1972. — 266 p. ISBN 0471201405, 9780471201403

27. O'Hagan A. Expert knowledge elicitation: Subjective but scientific // The American Statistician, 2019, 73(Suppl 1), 69–81.
https://doi.org/10.1080/00031305.2018.1518265

28. Zellner A. Bayesian Analysis in Econometrics and Statistic: The Zellner View and Papers (Economists of the Twentieth Century). — Edward Elgar Publishing, 1997. — 553 p. ISBN-13: 978-1858982205

29. Raiffa H. Decision analysis. Introductory Lectures on Choices Under Uncertainty Reading Mass. — Addison-Wesley, 1968. — xxiii+309 p. ISBN-13: 978-0201062908

30. Popper K. R. Objective Knowledge: An Evolutionary Approach. — Oxford, England: Oxford University Press, 1972. — X+395 p. DOI 10.2307/2184085 ISBN-13 : 978-0198750246

31. Popper K. The Logic of Scientific Discovery. — London, New York: Routledge Classics, 2005. — xiii+513 p. ISBN 0-203-99462-0

32. Попович М. В. Логіка і наукове пізнання. — Київ: Наукова думка, 1971. — 156 с.

33. Zadeh L. Fuzzy Sets // Information and Control, 1965, 8, 338—353.
https://doi.org/10.1016/S0019-9958(65)90241-X

34. Zadeh L. A. Outline of a new approach to the analysis of complex systems and decision processes // IEEE Trans. on Systems, Man and Cybernetics, 1973, SMC-3(1), 28–44. DOI: 10.1109/TSMC.1973.5408575

35. Zadeh L. A. Fuzzy sets as a basis for a theory of possibility // Fuzzy Sets and Systems, 1978, 1, 3–28. DOI:
https://doi.org/10.1016/0165-0114(78)90029-5

36. Morris W. Th. Management Science: A Bayesian Introduction. — N. Y.: Prentice-Hall Inc., Englewood Cliffs, — 1968, 226 p. ISBN-13: 978-0135491393

References:

1. Renyi, A. (1955). On a new axiomatic theory of probability. Acta Mathematica Academiae Scientiarum Hungarica, 6(3), 285–335. https://doi.org/10.1007/BF02024393

2. Renyi, A. (1970). Foundations of probability. San-Francisco, Holden-Day. Inc., xvi+366 p. ISBN: 0816271143, 9780816271146

3. Savage, L. J. (1954, 1972). The foundations of statistics. John Wiley & Sons, Inc., 1954, 2nd Revised ed. Dover Publications 1972, 294 p. ISBN: 978-0486623498

4. Bayes, Mr; Price, Mr (1763). An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr Bayes, F. R. S. Communicated by Mr Price, in a Letter to John Canton, M. A. and F. R. S. Philosophical Transactions of the Royal Society of London. 53, 370–418. Downloaded from https://royalsocietypublishing.org/ on 12 May 2023

5. Hooper, M. (2013). Richard Price, Bayes' theorem, and God. Significance, 10 (1), 36–39. doi:10.1111/j.1740-9713.2013.00638.x.

6. McGrayne, S. B. (2011). The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy. Yale University Press. 288 p. jstor.org/stable/j.ctt1np76s ISBN-13 9780300169690

7. Stone, J. V. (2013). Bayes' Rule: A Tutorial Introduction to Bayesian Analysis (5.5.2013 ed.). Sebtel Press, 174 p. DOI: 10.13140/2.1.1371.6801

8. Vasyl'kiv I. M. (2020). Bases of Theory of Chances and Mathematical Statistics: Train Aid. Lviv. Lviv University Press,184 p.ISBN 978–617–10–0354–5

9. Barabash, O. V., Musijenko, A. P., and Svynchuk, O. V. (2021). Theory of Chances [eBook]: Train Aid. Kyiv, Kyiv Polytechnic Institute Press, 193 p.

10. Stylianides, N., and Kontou, E. (2020). Bayes Theorem and its Recent Applications. Mathematics Research Journal, MA3517, 1–6.

11. McCullagh, P. (2002). What is a statistical model? The Annals of Statistics, 30(5), 1225–1310. https://doi.org/10.1214/aos/1035844977

12. Taraldsen, G., and Lindqvist, B. H. (2016). Conditional probability and improper priors. Communications in Statistics — Theory and Methods, 45(17), 5007–5016. doi:10.1080/03610926.2014.935430

13. Puga, J. L., Krzywinski, M., and Altman N. (2015). Bayes' theorem. Nature Methods, 277 – 278. doi:10.1038/nmeth.3335

14. Berrar, D. (2018). Bayes' Theorem and Naive Bayes Classifier. Tokyo, Japan: Data Science Laboratory, Tokyo Institute of Technology. 1–18. doi:DOI: 10.1016/B978-0-12-809633-8.20473-1

15. Thompson, D. R., and Martin, C. R. (2017). Bayes' theorem and its application to cardiovascular nursing. European Journal of Cardiovascular Nursing, 16(8), 659–661. doi:https://doi.org/10.1177/1474515117712317

16. Lohrke, F. T., Carson, Ch. M., and Lockamy, A. (2018). Bayesian analysis in entrepreneurship decision-making research: Review and future directions. Management Decision, 56(2), 972–986. DOI:10.1108/MD-12-2016-0948

17. Taraldsen, G., Tufto, J., and Lindqvist, B. H. (2022). Improper priors and improper posteriors. Scandinavian Journal of Statistics, 49(3), 969–991. https:// doi.org/10.1111/sjos.12550

18. Berger, J. O., de Oliveira, V., and Sanso, B. (2001). Objective Bayesian analysis of spatially correlated data. Journal of the American Statistical Association, 96(456), 1361–1374. https://www. jstor .org /stable/3085905

19. Bioche, C., and Druilhet, P. (2016). Approximation of improper priors. Bernoulli, 22(3), 1709–1728. DOI: 10.3150/15-BEJ708

20. Bord, S., Bioche, C., and Druilhet, P. (2018). A cautionary note on Bayesian estimation of population size by removal sampling with diffuse priors. Biometrical Journal, 60(3), 450–462. DOI: 10.1002/bimj.201700060

21. Hannig, J., Iyer, H., Lai, R. C. S., and Lee, T. C. M. (2016). Generalized fiducial inference: A review and new results. Journal of the American Statistical Association, 111(515), 1346–1361. https://doi.org/10.1080/01621459.2016.1165102

22. Taraldsen, G., and Lindqvist, B. H. (2013). Fiducial theory and optimal inference. Annals of Statistics, 41(1), 323–341. DOI: 10.1214/13-AOS1083

23. Hartigan, J. (1983, 2011). Bayes theory. Springer. 158 p. ISBN: 978-1461382447

24. Lindley, D. V. (2014). Understanding uncertainty (Revised ed.). John Wiley & Sons, Inc, 432 p. ISBN: 978-1-118-65012-7

25. Lehmann, E. L., and Romano, J. P. (2005). Testing statistical hypotheses. (3-rd ed). Springer Science+Business Media, LLC, xiv+785 p. ISBN 0-387-98864-5

26. Finetti, B. D. (1972). Probability, induction and statistics: The art of guessing. Wiley. 266. ISBN 0471201405, 9780471201403

27. O'Hagan, A. (2019). Expert knowledge elicitation: Subjective but scientific. The American Statistician, 73(Suppl 1), 69–81. DOI: https://doi.org/10.1080/00031305.2018.1518265

28. Zellner, A. (1997). Bayesian Analysis in Econometrics and Statistic: The Zellner View and Papers (Economists of the Twentieth Century). Edward Elgar Publishing. 553 p.
ISBN-13: 978-1858982205

29. Raiffa, H. (1968). Decision analysis. Introductory Lectures on Choices Under Uncertainty Reading, Mass.: Addison-Wesley, xxiii + 309 pp. ISBN-13: 978-0201062908

30. Popper, K. R. (1972). Objective Knowledge: An Evolutionary Approach. Oxford, England: Oxford University Press. X+395 p. DOI 10.2307/2184085 ISBN-13 : 978-0198750246

31. Popper, K. (2005). The Logic of Scientific Discovery. London, New York: Routledge Classics, xiii+513 p. ISBN 0-203-99462-0

32. Popovych M. V. (1971). Logic and scientific cognition. Kyiv, Naukova dumka, 156 p

33. Zadeh, L. (1965). Fuzzy Sets. Information and Control. 8, 338—353.
https://doi.org/10.1016/S0019-9958(65)90241-X

34. Zadeh, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. on Systems, Man and Cybernetics, SMC-3(1), 28–44.
DOI: 10.1109/TSMC.1973.5408575

35. Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 3–28. https://doi.org/10.1016/0165-0114(78)90029-5

36. Morris, W. Th. (1968). Management Science: A Bayesian Introduction. N. Y.: Prentice-Hall, Inc., Englewood Cliffs, 226 p. ISBN-13: 978-0135491393

Науково-методична стаття.
Надійшла до редакції 01.08.2023.
Прийнято до публікації 01.12.2023.