UDC [519.226:519.233.3:519.246]:004 DOI: https://doi.org/10.15407/geotm2025.175.094

RELATIVE PROBABILITIES OF HYPOTHESES IN THE BAYESIAN APPROACH

Lapshyn Ye., Shevchenko O.

M.S. Poliakov Institute of Geotechnical Mechanics of the National Academy of Sciences of Ukraine

Abstract. Bayesian networks are widely used in various fields of science and industry, in particular geotechnical mechanics, ecology. They are applied to support decision making under risk conditions, under which is understood as the product of the probability of danger seriousness of the consequences if it is implemented. A Bayesian network allows you to establish cause-and-effect relationships between events and determine the probability of a particular situation when receiving new information about a change in the state of any network variable. It allows you to consistently refine the probabilities of a complete group of events based on the results of observations based on the intermediate results of a stochastic experiment. In other words, Bayes' theorem provides the possibility of determining the posterior probabilities of hypotheses taking into account the priors based on additional experience results. It is known that Bayes' theorem implies: the ratio of the posterior probability of a hypothesis to the prior probability is equal to the ratio of the probability of an event under a given hypothesis to the total probability of the event. However, there is no information on how much the posterior probability of hypotheses changes compared to the prior probability with the appearance of new data. Such a need arises when planning experiments, the conduct of which is associated with large financial and time costs, and most importantly, is associated with unacceptable risks. As a result of the research, an analysis of the influence of additional data on the probabilities of events that depend on hypotheses on the ratio of the posterior and prior ones was carried out. The ratio of probabilities (relative posterior probability of hypotheses) shows how many times the posterior probability will change compared to the prior. The dependence of the ratio on probabilities is shown on graphs in 3d and 2d form. The relative posterior probability of hypotheses allows you to identify the hypothesis that reacts most strongly to additional data and, on this basis, determine the least expensive experiments to refine the prior probability. To demonstrate the simplicity of calculations that allow you to rank hypotheses in terms of sensitivity to new data, a model example is given. As a result of the research, the following was obtained. The relative posterior probability of hypotheses is proposed, which shows how many times the posterior probability will change compared to the prior when receiving additional data related to the occurrence of events caused by the hypotheses. The relative posterior probability of hypotheses is equal to the ratio of the conditional probability of the occurrence of an event under a given hypothesis to the probability of this event. The relative posterior probability of hypotheses allows us to rank hypotheses in terms of sensitivity to new data. The greatest sensitivity was found for an event probability of less than 0.1.

Keywords: hypothesis, event, probability, a priori, a posteriori.

1. Introduction

Bayesian networks are widely used in various fields of science and industry, in particular, geotechnical mechanics [1] and ecology [2]. They are applied to support decision making under risk conditions, under which is understood as the product of the probability of danger seriousness of the consequences if it is implemented [1, 3, 4]. A Bayesian network allows one to establish cause-and-effect relationships between events and determine the probability of a given situation occurring when new information is received about a change in the state of any network variable [5]. Bayesian networks are based on Bayes' theorem, which is sometimes called a law, rule, principle, or formula for the probabilities of hypotheses [6]. It "…allows for a consistent refinement of the probabilities of a complete group of events based on the results of observations on the basis of intermediate results of a stochastic experiment" [7]. In other words, Bayes' theorem provides the ability to determine the a posteriori probability of hypotheses based on additional experimental results, taking into account the a priori. The paper [8] states: "Bayes' theorem is a remarkable theorem in probability theory. As technology advances and society progresses, the advantages of Bayes' theorem gradually become apparent, allowing us to better utilize existing re-

Received: 15.04.2025 Accepted: 15.12.2025 Available online: 18.12.2025

sources to make more accurate judgments." The philosophical aspect of the theorem is discussed in detail in the monograph [9]. It is known [10] that Bayes' theorem states that the ratio of a hypothesis's posterior probability to its prior probability is equal to the ratio of the probability of an event under that hypothesis to the total probability of the event. However, there is no information on how significantly the posterior probability of hypotheses changes compared to the prior probability when new data emerges. This need arises when planning experiments that are associated with significant financial and time expenditures and, most importantly, are associated with unacceptable risks.

In this regard, the *aim of this study* is to analyze the influence of additional data on the probabilities of hypothetical events on the ratio of posterior to prior probabilities.

Object of the study: the influence of additional data on the probabilities of hypothetical events on the ratio of posterior to prior probabilities.

2. Methods

The following methods were used in this study: an analytical review of literature sources; probability theory; and mathematical statistics. Internet resources were used to analyze the current state of the art.

To achieve this goal, the following **tasks** were solved:

– how much the posterior probability changes compared to the prior probability when new data appears is established;

– what the posterior probability is determined;

– what the relative a posteriori probability of hypotheses allows us to execute is established;

– the sensitivity of the relative a posteriori probability of hypotheses for an event probability less than 0.1 is determined.

3. Theoretical and experimental parts

There is a complete group of hypotheses A_i $(i=1,...,N)$ and statistically related events B. The a priori probabilities of hypotheses are known $P(A_i)$. Bayes' theorem allows one to determine the probabilities of hypotheses given that an event has occurred B [7, 11, 12]

$$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{P(B)}, \tag{1}$$

where $P(A_i|B)$ – probability of a hypothesis A_i upon the occurrence of an event B (posterior probability); $P(B|A_i)$ – probability of an event occurring B if the hypothesis is true A_i $P(A_i)$ – a priori probability of a hypothesis A_i;

$$P(B)=\sum_{i=1}^{N}P(B|A_i)P(A_i) \qquad (2)$$

– total probability of an event occurring B, which $P(B)\neq 0$.

Dividing the left and right sides of Bayes' formula by $P(A_i)$ and denoting relationships P_i^*, we receive

$$P_i^*=\frac{P(A_i|B)}{P(A_i)}=\frac{P(B|A_i)}{P(B)}. \qquad (3)$$

Taking into account (1) and (3), we obtain

$$P(A_i|B)=P_i^*P(A_i). \qquad (4)$$

Attitude P_i^* let's call it the relative a posteriori probability of hypotheses (RPPH) A_i upon receipt of additional data, which shows how many times the posterior probability will change $P(A_i|B)$ compared to the a priori $P(A_i)$.

Addiction P_i^* from probabilities $P(B)$ and $P(B|A_i)$ is shown in Figure 1: a) in 3D form, and b) in 2D form. The graph of the function (see Fig. 1b) is a hyperbola with domain $P(B)=(0,1]$. The branches are located in the first quadrant – positive inverse proportionality coefficient.

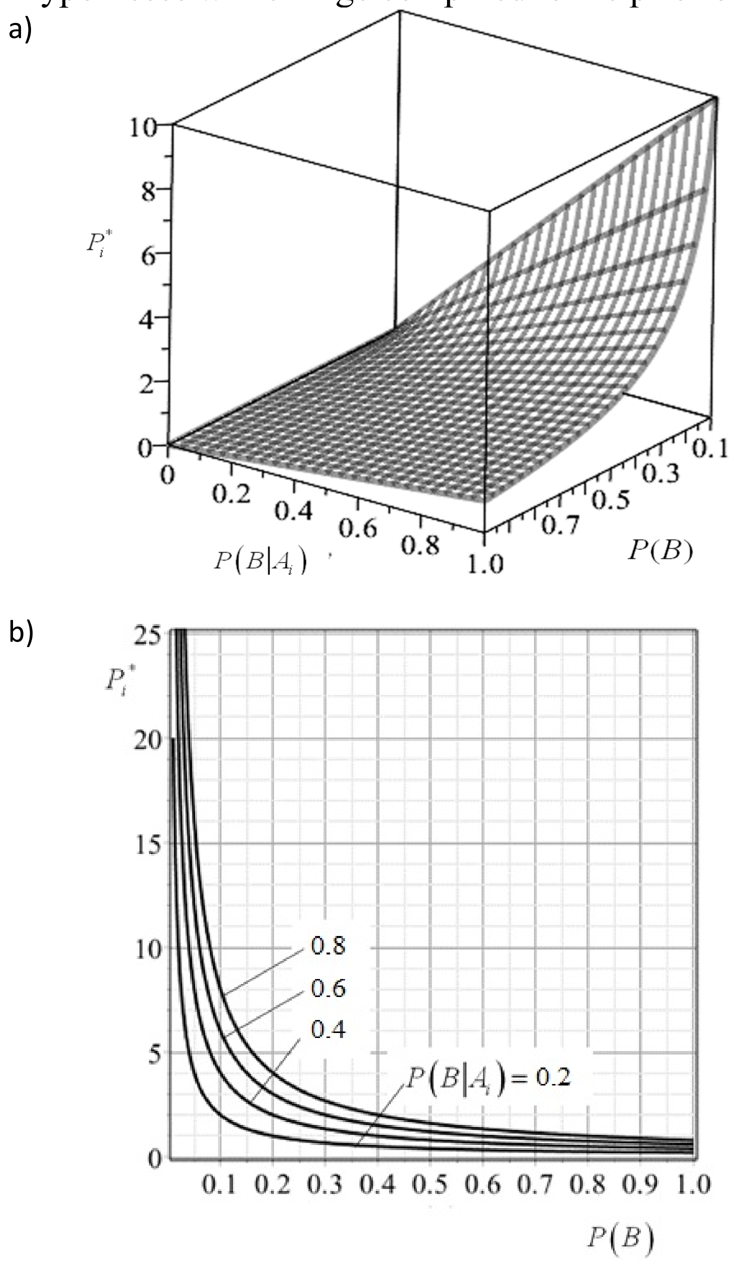
The domain of a function can be divided into three intervals. The first, $0<P(B)\leq 0.1$; the second, $0.1<P(B)\leq 0.5$; the third, $0.5<P(B)<1$. On the first interval, compared to the third, with a decrease in probability $P(B)$ RPPH is rapidly increasing (derivative $\partial P_i^*/\partial P(B)$ on the first interval it significantly, more than 10 times, exceeds the derivative on the third interval). The second interval is transitional.

RPPH allows us to identify the hypothesis that responds most strongly to additional data and, based on this, determine the least expensive experiments to refine the prior probability.

To demonstrate the simplicity of the calculations that allow us to rank hypotheses by sensitivity to new data, let's consider a model example.

Model example. Pollution exceeding the maximum permissible concentration has been detected in a reservoir. The water discharges are carried out by three mining and processing plants, which we will consider as hypotheses.

Probabilities related to a plant will be written with a subscript $i=1, 2, 3$. Let us denote: $P(A_i)$ – the likelihood of discharge of contaminated water by enterprises; $P(B|A_i)$ – conditional probabilities of the presence of a toxic substance in the discharge of enterprises. It is necessary to determine by how many times the posterior probabilities of hypotheses will change compared to the prior ones.



a) and b) three-dimensional and two-dimensional graphics

Figure 1 – Addiction P_i^* from probabilities $P(B)$ and $P(B|A_i)$

Given:

$$P(A_1) = 0.4, \quad P(A_2) = 0.2, \quad P(A_3) = 0.4; \tag{5}$$

$$P(B|A_1) = 0.03, \quad P(B|A_2) = 0.02, \quad P(B|A_3) = 0.04,$$

here (5) are the a priori probabilities of the hypotheses.

The total probability of a toxic substance entering a water body during the operation of three enterprises is calculated using formula (2)

$$P(B) = 0.03 \cdot 0.4 + 0.02 \cdot 0.2 + 0.04 \cdot 0.4 = 0.032.$$

Using formula (3) we determine RPPH

$$P_1^* = \frac{P(B|A_1)}{P(B)} = \frac{0.03}{0.032} = 0.9375,$$

$$P_2^* = \frac{P(B|A_2)}{P(B)} = \frac{0.02}{0.032} = 0.6250,$$

$$P_3^* = \frac{P(B|A_3)}{P(B)} = \frac{0.04}{0.032} = 1.2500.$$

RPPH indicate that the probability of hypothesis 1 decreases $\left(P_1^* = 0.9375 < 1\right)$, the probability of hypothesis 2 also decreases $\left(P_2^* = 0.6250 < 1\right)$, and the probability of hypothesis 3 increases $\left(P_3^* = 1.2500 > 1\right)$, moreover, the increase in the probability of hypothesis 3 occurs more than 1.9 times more intensively than the decrease in the probability of hypothesis 2.

We also calculate the posterior probabilities of the hypotheses (the probability of discharges by enterprises given that the maximum permissible concentration has been exceeded). For this, we use formula (4)

$$P(A_1|B) = 0.9375 \cdot 0.4 = 0.375,$$

$$P(A_2|B) = 0.6250 \cdot 0.2 = 0.125,$$

$$P(A_3|B) = 1.250 \cdot 0.4 = 0.500.$$

Thus, the proposed RPPH shows whether the posterior probability will increase or decrease compared to the prior one.

4. Conclusions

Thus, the following findings were established in the course of the research. Bayes' formula, also known in the literature as Bayes' theorem, law, principle, and formula for the probability of hypotheses, is widely used in various fields, including science, medicine, economics, risk analysis, and others. It allows one to determine the posterior probabilities of hypotheses, taking into account the prior probabilities, based on additional experimental results. However, there is no information on how significantly the posterior probability of hypotheses changes compared to the prior probability when new data becomes available. This need arises when planning experiments, the implementation of which is associated with significant financial and time expenditures. In accordance with the stated goal of the study, an analysis was performed of the influence of additional data on the probabilities of hypotheses-dependent events on the ratio of the posterior and prior probabilities of the latter, and the following problems were solved:

1. A relative posterior probability of hypotheses is proposed, characterizing how much the posterior probability changes compared to the prior probability when new data emerges.

2. The relative posterior probability of hypotheses is equal to the ratio of the conditional probability of an event occurring under a given hypothesis to the total probability of that event.

3. The relative posterior probability of hypotheses allows for ranking hypotheses by sensitivity to new data.

4. The greatest sensitivity of the relative posterior probability of hypotheses is found for events with probabilities less than 0.1.

A model example is provided.

The results can be used in planning experiments that require significant financial and time expenditures.

Conflict of interest

Authors state no conflict of interest.

REFERENCES

1. Mishra, R.A., Uotinen, L. and M. Rinne (2021), "Bayesian network approach for geotechnical risk assessment in underground mines", *The Journal of the Southern African Institute of Mining and Metallurgy*, vol. 121, pp. 287–294. https://doi.org/10.17159/2411-9717/581/2021

2. McCann, Marcot, R.K. and Ellis, R. (2006), "Bayesian belief networks: applications in ecology and natural resource management", *Canadian Journal of Forest Research*, no. 36(12), pp. 3053-3062. https://doi.org/10.1139/x06-238

3. Bulat, A.F., Bunko, T.V. Kokoulin, I.E. Yashchenko, I.O., Shevchenka, V.H. and Zhalilov, A.Sh. (2015), "About the unified concept of technogenic safety in coal mines", *Geo-Technical Mechanics*, no. 124, pp. 1–15. available at: https://nasplib.isofts.kiev.ua/items/16f90ae0-e46a-4b3b-9a02-6f07151a1e05 (Accessed 9 April 2025).

4. Bulat, A.F., Krukovsky, A.P., Bunko, T.V., Kokoulin, I.Ie., Yashchenko, I.A.and Zhalilov, A.Sh. (2017), "Assessing the unsafety of mines based on the theory of risks", *Geo-Technical Mechanics*, no. 135, pp. 1–15. available at: https://nasplib.isofts.kiev.ua/items/d1a485fc-48b0-43c9-936b-93daba7b697c (Accessed 9 April 2025).

5. Bidyuk, P.I. and Korshevnyuk, L.O. (2010), *Proektuvannia kompiuternykh informatsiinykh system pidtrymky pryiniattia rishen* [Design of computer information systems for decision support: Textbook], NNC "IPSA" NTUU "Kyiv Polytechnic Institute", Kyiv, Ukraine. available at: https://duikt.edu.ua/ua/lib/1/category/96/view/2207 (Accessed 9 April 2025).

6. Haschuk, L.P. and Haschuk, P.M. (2023), "Some aspects of interpreting the content of the hypothesis theorem in situations of uncertainty", *Bulletin of the Lviv State University of Life Safety*, no. 28, pp. 55–72. available at: https://journal.ldubgd.edu.ua/index.php/Visnuk/article/view/2625 (Accessed 9 April 2025).

7. Kartashov, M.V. (2008), *Imovirnist, protsesy, statystyka* [Probability, processes, statistics: Handbook[, *Publishing and printing center* "*Kyiv University*", Kyiv, Ukraine. available at: https://probability.knu.ua/userfiles/kmv/VPS_Pv.pdf (Accessed 9 September2025).

8. Zhang, Y. (2023), "The Application of Bayesian Theorem. Highlights in Science", *Engineering and Technology*, no. 49, pp. 520–526, https://doi.org/10.54097/hset.v49i.8605

9. Sprenger, J. and Hartmann, S. (2019), *Bayesian Philosophy of Science*, Oxford and New York: Oxford University Press, available at: https://www.laeuferpaar.de/Papers/BookFrame_v1.pdf (Accessed 9 April 2025).

10. Roush, S. (2005), *Tracking Truth: Knowledge, Evidence, and Science*, Oxford: Oxford University Press, UK. https://doi.org/10.1093/0199274738.001.0001

11. Gnedenko, B.V. (2010), *Kurs teorii ymovirnostei: pidruchnyk* [Course in Probability Theory: Textbook], Publishing and Printing Center "Kyiv University", Kyiv, Ukraine. available at: http://www.nbuv.gov.ua/node/338 (Accessed 9 April 2025).

12. Oppermann, A. (2023), *Bayes' Theorem: The Holy Grail of Data Science. Intuitive derivation of the Bayes' Theorem*. available at: https://medium.com/data-science/bayes-theorem-the-holy-grail-of-data-science-55d93315defb (Accessed 9 April 2025).

About the authors

Lapshyn Yevhen, Doctor of Technical Sciences (D.Sc.), Senior Researcher, Principal Researcher in Department Ecology of Natural Resources Development, M.S. Poliakov Institute of Geotechnical Mechanics of the National Academy of Sciences of Ukraine (IGTM of the NAS of Ukraine), Dnipro, Ukraine, 19481948les@gmail.com, ORCID ***0000-0002-5443-5566***

Shevchenko Oleksandr, Doctor of Technical Sciences (D.Sc.), Senior Researcher, Senior Researcher in Department of Geomechanical Basis of Open-Pit Technology, M.S. Poliakov Institute of Geotechnical Mechanics of the National Academy of Science of Ukraine (IGTM of the NAS of Ukraine), Dnipro, Ukraine, igtm.aishevchenko@gmail.com ***(Corresponding author)***, ORCID ***0000-0003-2630-0186***

ВІДНОСНІ ВІРОГІДНОСТІ ГІПОТЕЗ ПРИ БАЙЄСІВСЬКОМУ ПІДХОДІ

Лапшин Є., Шевченко О.

Анотація. Баєсівські мережі широко використовуються в різних галузях науки та промисловості, зокрема геотехнічній механіці, екології. Вони застосовуються для підтримки прийняття рішень в умовах ризику, під яким розуміється добуток ймовірності небезпеки на серйозність наслідків при його реалізації. Баєсівська мережа дозволяє встановлювати причинно-наслідкові зв'язки між подіями та визначати ймовірність певної ситуації при отриманні нової інформації про зміну стану будь-якої мережевої змінної. Вона дозволяє послідовно уточнювати ймовірності повної групи подій на основі результатів спостережень на основі проміжних результатів стохастичного експерименту. Іншими словами, теорема Баєса забезпечує можливість визначення апостеріорних ймовірностей гіпотез з урахуванням апріорних на основі додаткових результатів досвіду. Відомо, що з теореми Баєса випливає: відношення апостеріорної ймовірності гіпотези до апріорної ймовірності дорівнює відношенню ймовірності події за заданою гіпотезою до загальної ймовірності події. Однак немає інформації про те, наскільки змінюється апостеріорна ймовірність гіпотез порівняно з апріорною ймовірністю з появою нових даних. Така потреба виникає при плануванні експериментів, проведення яких пов'язане з великими фінансовими та часовими витратами, а головне, пов'язане з неприйнятними ризиками. В результаті дослідження було проведено аналіз впливу додаткових даних на ймовірності подій, що залежать від гіпотез, від співвідношення апостеріорних та апріорних. Співвідношення ймовірностей (відносна апостеріорна ймовірність гіпотез) показує, у скільки разів зміниться апостеріорна ймовірність порівняно з апріорною. Залежність співвідношення від ймовірностей показано на графіках у 3d та 2d формі. Відносна апостеріорна ймовірність гіпотез дозволяє виявити гіпотезу, яка найсильніше реагує на додаткові дані, та на цій основі визначити найменш витрати на експерименти для уточнення апріорної ймовірності. Для демонстрації простоти розрахунків, що дозволяють ранжувати гіпотези за чутливістю до нових даних, наведено модельний приклад. В результаті дослідження було отримано наступне. Запропоновано відносну апостеріорну ймовірність гіпотез, яка показує, у скільки разів зміниться апостеріорна ймовірність порівняно з апріорною при отриманні додаткових даних, пов'язаних з виникненням подій, спричинених гіпотезами. Відносна апостеріорна ймовірність гіпотез дорівнює відношенню умовної ймовірності настання події за заданою гіпотезою до ймовірності цієї події. Відносна апостеріорна ймовірність гіпотез дозволяє ранжувати гіпотези за чутливістю до нових даних. Найбільша чутливість була виявлена для ймовірності події менше 0,1.

Ключові слова: гіпотеза, подія, ймовірність, апріорна, апостеріорна.